

Aer70

ESTUDOS EM AVALIAÇÃO EDUCACIONAL • N. 2 JUL./DEZ. 1990 •

FUNDAÇÃO CARLOS CHAGAS • SÃO PAULO

A partir de 2006 passa a ser QUADRIMESTRAL

ISSN 0103-6831

e-ISSN 1984-932X

1. Avaliação 2. Políticas Educacionais 3. Qualidade do Ensino

I. Fundação Carlos Chagas II. Departamento de Pesquisas

Educacionais/FCC

INDEXADO EM

BAMP – Biblioteca Ana Maria Poppovic (*Brasil, FCC*)

www.fcc.org.br

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

http://pergamum.inep.gov.br/pergamum/biblioteca/index.php?resolution2=1024_1

Capes – 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/oi_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

VERSÃO IMPRESSA

Abril de 2018

Tiragem: 350 exemplares

E-MAILS

eae@fcc.org.br (*contato*)

publicacoesfcc@fcc.org.br (*aquisição e assinaturas*)

ESTUDOS EM AVALIAÇÃO EDUCACIONAL

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

As normas para a publicação estão no final do fascículo.

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

Fabiana Silva Fernandes

Gabriela Miranda Moriconi

Vandré Gomes da Silva

EDITORES ASSOCIADOS

Daniel Abud Seabra Matos

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*)

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 Vasconcellos de Luna

(*Pontifícia Universidade Católica de São Paulo, São Paulo, São Paulo, Brasil*)

Yara Lúcia Esposito

(*Fundação Carlos Chagas, São Paulo, São Paulo, Brasil*)

CHEFE DE EDIÇÕES

Adélia Maria Mariano da Silva Ferreira

ASSISTENTES EDITORIAIS

Camila Maria Camargo de Oliveira

Marcia Caxeta

AUXILIAR DE EDIÇÕES

Camila de Castro Costa

NORMALIZAÇÃO BIBLIOGRÁFICA

Biblioteca Ana Maria Poppovic

REVISÃO ESTATÍSTICA

Miriam Bizzocchi

Raquel da Cunha Valle

PROJETO GRÁFICO

Casa Rex

DIAGRAMAÇÃO

Líquido Editorial - Claudio Brites

IMPRESSÃO

Paym Gráfica e Editora Ltda.

SUMÁRIO

ARTIGOS

- Contextualização curricular numa rede de escolas portuguesas:
promessa ou oportunidade perdida?
*Contextualización curricular en una red de escuelas portuguesas:
¿promesa o oportunidad perdida?*
*Curriculum contextualization in a network of Portuguese schools:
a promise or a missed opportunity?*
Maria do Céu Roldão, Sílvia de Almeida.....8
- Impacto da bonificação educacional em Pernambuco
Impacto de la bonificación educacional en Pernambuco
Impact of educational bonus in Pernambuco
Clayton Sirilo do Valle Furtado, Tufi Machado Soares.....48
- A (des)igualdade de conhecimento no ciclo de alfabetização
La (des)igualdad de conocimiento en el ciclo de alfabetización
The (in)equality of knowledge in the literacy cycle
Naira da Costa Muylaert Lima, Alicia Maria Catalano de Bonamino,
Luis Antonio Fajardo Pontes.....78
- Avaliação da educação infantil na Austrália:
contribuições para o Brasil
Evaluación de la educación infantil en Australia: contribuciones para Brasil
Evaluation of early childhood education in Australia: contributions for Brazil
Cláudia Oliveira Pimenta.....108
- Avaliação da aprendizagem profissional no estágio
supervisionado: estudo exploratório na educação infantil
*Evaluación del aprendizaje profesional en el prácticum supervisado:
estudio exploratorio en la educación infantil*
*Assessment of professional learning in supervised practicum:
exploratory study in early childhood education*
Sara Barros Araújo, Ana Pereira Antunes.....134
- Dados ausentes em avaliações educacionais:
comparação de métodos de tratamento
*Datos ausentes en evaluaciones educacionales:
comparación de métodos de tratamiento*
*Missing data in educational assessment:
a comparison of data treatment methods*
Luis Gustavo do Amaral Vinha, Jacob Arie Laros.....156

Austrália, Brasil e Canadá: impacto das avaliações no ensino de Ciências <i>Australia, Brasil y Canadá: impacto de las evaluaciones en la enseñanza de Ciencias</i> <i>Australia, Brazil and Canada: the impact of assessment on Science teaching</i> Paulo Sérgio Garcia, Xavier Fazio, Debra Panizzon, Nelio Bizzo.....	188
Avaliação institucional na educação básica: os desafios da implementação <i>Evaluación institucional en la educación básica: los desafíos de la implementación</i> <i>Institutional evaluation in basic education: implementation challenges</i> Luiz Síveres, José Roberto de Souza Santos.....	222
Práticas letivas de sala de aula de matemática nos anos iniciais <i>Prácticas lectivas de Matemáticas en aulas de primaria</i> <i>Mathematics classroom practices in the elementary school</i> Isabel Cristina Rodrigues de Lucena, Josete Leal Dias, Antônio Manuel Águas Borralho.....	254
Indicadores educacionais de matrículas de alunos com deficiência no Brasil (1974-2014) <i>Indicadores educacionales de matrículas de estudiantes con discapacidad en Brasil (1974-2014)</i> <i>Education indicators related to the enrollment of pupils with disabilities in Brazil (1974-2014)</i> Andressa Santos Rebelo, Mônica de Carvalho Magalhães Kassar.....	276
INSTRUÇÕES A COLABORADORES.....	308

ARTIGOS

CONTEXTUALIZAÇÃO CURRICULAR NUMA REDE DE ESCOLAS PORTUGUESAS: PROMESSA OU OPORTUNIDADE PERDIDA?

MARIA DO CÉU ROLDÃO

SÍLVIA DE ALMEIDA

RESUMO

Na década de 1990, as políticas educativas internacionais e o campo teórico dos estudos curriculares conferiram centralidade ao debate sobre o currículo, conduzindo, nos países de tradição centralista, à passagem do paradigma do currículo uniforme, prescritivo, para um paradigma de contextualização curricular. Em Portugal, a introdução deste paradigma pelas políticas curriculares determinou a necessidade de prescrever às escolas a elaboração de um projeto curricular. A partir de então, a literatura tem demonstrado uma conceção das dinâmicas de construção, implementação e avaliação de projetos curriculares como documentos burocráticos, obedecendo a uma lógica da normatividade decretada, elaborados sobretudo para os painéis de avaliação externa das escolas, incapazes deste modo de implementar mudanças nas práticas curriculares. Neste texto, apresentam-se os resultados da avaliação de 12 projetos curriculares de uma rede colaborativa de escolas portuguesas com recurso à análise de conteúdo.

PALAVRAS-CHAVE CONTEXTUALIZAÇÃO CURRICULAR • AVALIAÇÃO CURRICULAR • PROJETOS CURRICULARES • POLÍTICA EDUCACIONAL.

CONTEXTUALIZACIÓN CURRICULAR EN UNA RED DE ESCUELAS PORTUGUESAS: ¿PROMESA O OPORTUNIDAD PERDIDA?

RESUMEN

En la década de los 90, las políticas educativas internacionales y el campo teórico de los estudios curriculares debatían por encima de todo el currículo, lo que condujo en los países de tradición centralista al paso del paradigma del currículo uniforme y prescriptivo a un paradigma de contextualización curricular. En Portugal, la introducción de este paradigma por las políticas curriculares determinó la necesidad de prescribir a las escuelas la elaboración de un proyecto curricular institucional. Desde entonces la literatura ha demostrado una concepción de las dinámicas de construcción, implementación y evaluación de proyectos curriculares como si fueran documentos burocráticos, obedeciendo a una lógica de la normatividad decretada, elaborados sobre todo para los paneles de evaluación externa de las escuelas, incapaces de este modo de implementar cambios en las prácticas curriculares. En este texto, se presentan los resultados de la evaluación de 12 proyectos curriculares de una red colaborativa de escuelas portuguesas realizada teniendo en cuenta el análisis de contenido.

PALABRAS CLAVE CONTEXTUALIZACIÓN CURRICULAR • EVALUACIÓN CURRICULAR • PROYECTOS CURRICULARES • POLÍTICA EDUCATIVA.

CURRICULUM CONTEXTUALIZATION IN A NETWORK OF PORTUGUESE SCHOOLS: A PROMISE OR A MISSED OPPORTUNITY?

ABSTRACT

In the 1990's, international educational policies, as well as the theoretical field of curriculum studies, gave centrality to the debate on curriculum. This led countries with a centralist tradition to shift from the paradigm of a uniform and prescriptive curriculum towards a paradigm based on curriculum contextualization. Following the latter, new curricular policies were introduced in the Portuguese system which determined that schools should develop their own curriculum plan. Since then, the literature on the subject has shown that the construction, implementation and evaluation dynamics of these curriculum plans have been conceived as bureaucratic documents, obeying a logic of normativity and being mostly designed to suit external evaluation panels that proved unable to implement change in curriculum practices. In this paper, the method of content analysis was used to present the results of the evaluation of 12 curriculum projects implemented in a collaborative network of Portuguese schools.

KEYWORDS CURRICULUM CONTEXTUALIZATION • CURRICULUM ASSESSMENT • CURRICULUM PROJECTS • EDUCATIONAL POLICIES.

INTRODUÇÃO

Em Portugal, à semelhança de outros países europeus e no âmbito da Organização para a Cooperação e Desenvolvimento Económico (OCDE), as políticas educativas e curriculares desenvolvidas a partir da década de 1990 têm vindo a sublinhar o papel da escola como espaço central da ação educativa e o papel do professor como ator principal desse processo, por meio de medidas de contextualização curricular, desde a introdução do projeto de Gestão Flexível do Currículo (PORTUGAL, 1997). Este projeto, de iniciativa ministerial mas lançado para todo o sistema, iniciou uma reflexão no sentido de fazer convergir a prescrição curricular nacional (Currículo Nacional) com a atribuição às escolas do poder de decisão quanto à contextualização e adequação desse currículo às suas situações particulares, corporizada num Projeto Curricular de Escola (PCE). O currículo tem sido assim cada vez mais entendido, na discursividade política, como uma construção participada por vários atores sociais, pelo menos de um ponto de vista retórico, em que se atribui mais autonomia à escola para conceção, desenvolvimento e avaliação dos PCEs em articulação com o currículo nacional.

No contexto português, a literatura tem feito um balanço pouco positivo a respeito das dinâmicas de construção, implementação e avaliação destes projetos curriculares (ROLDÃO, 2005a, 2014; PACHECO; PEREIRA, 2007; MORGADO; MARTINS, 2008; MORGADO, 2011; GONÇALVES; MORGADO, 2013), assinalando as consequências que daí resultam em termos de desenvolvimento da autonomia das escolas e da adequação da ação profissional dos professores, bem como da pretendida melhoria das aprendizagens dos alunos.

Passados 20 anos da introdução do Projeto de Gestão Flexível do Currículo, e com a retoma de políticas de flexibilização curricular que está em curso na ação do atual poder político em Portugal, importa refletir de novo sobre as práticas de gestão curricular nas escolas. Daí a oportunidade para centrar a atenção na importância do PCE como instrumento de renovação e inovação das práticas pedagógicas e de construção da autonomia curricular do professor e da escola, uma vez que é no campo da ação curricular concreta e das práticas docentes que se jogam a qualidade educativa e o sucesso das aprendizagens.

Neste artigo, apresentam-se os resultados de uma análise de conteúdo de 12 PCEs de uma rede colaborativa de escolas portuguesas, que constituem parte dos resultados de uma investigação sobre as práticas de gestão curricular expressas nos seus respetivos projetos.

Tendo como referência a centralidade que este documento orientador assume como instrumento de gestão curricular contextualizada, tanto na literatura teórica como nas políticas curriculares, partimos das seguintes questões: como as escolas que têm um PCE já desenvolvido construíram, ou não, um modelo de gestão curricular? Que componentes e princípios teoricamente sustentados se identificam para construir um modelo de referência orientador da elaboração de PCEs?

ENQUADRAMENTO HISTÓRICO-CONCEPTUAL

AS POLÍTICAS CURRICULARES: DA UNIFORMIDADE À CONTEXTUALIZAÇÃO

As políticas educativas internacionais do final do século XX (SKILBECK, 1990; ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD, 1991, 1993, 1994, 1998, 1999) colocaram o currículo no centro do debate educativo, o que encontrou respaldo no campo teórico (CARMEN; ZABALA, 1991; SACRISTÁN, 1994; ZABALZA, 1992; HUTMACHER, 1992; ROLDÃO, 1995, 1998, 1999b, 1999c, 2000), conduzindo, nos países de tradição centralista, à passagem do paradigma do currículo uniforme, prescritivo, para um paradigma da contextualização curricular.

As medidas curriculares desde então implementadas têm vindo a acentuar a reorientação dos *loci* de decisão curricular, em resultado da complexidade dos contextos e populações que a universalidade da escolarização trouxe para o interior da escola nas sociedades democráticas, acentuando as pressões quer no sentido da contextualização e diferenciação curricular e pedagógica, quer no reforço de um núcleo curricular comum corporizado nos chamados currículos nacionais (OECD, 2013, 2014, 2016).

O debate em torno da responsabilidade acrescida da escola na gestão do seu trabalho curricular tem-se centrado assim quer na articulação deste nível de decisão com o nível central, anteriormente o único nos países de administração centralista como Portugal, quer em questões de organização e gestão contextualizada, nomeadamente concretizada em projetos próprios de cada escola, desejavelmente adequados a cada grupo de alunos e respetivos contextos, no sentido de melhorar o nível e a qualidade das aprendizagens curriculares pretendidas para todos (ROLDÃO, 2008).

O CURRÍCULO ESSENCIAL COMUM (*CORE CURRICULUM*) E O BINÓMIO CURRICULAR

As implicações organizativas deste processo de gestão do trabalho curricular da escola, com iniciativa e responsabilidades claramente acrescidas da instituição e dos professores, têm contudo sido menos discutidas na comunidade de

docentes, largamente marcada por uma cultura prescritiva e pouco socializada numa cultura de iniciativa e de auto e hetero-regulação.

Trata-se, assim, de recordar que o currículo é, na sua génese e mutabilidade, uma construção social (GOODSON, 1988, 2005; YOUNG, 1998). Contudo, gerado no seio de uma instituição especificamente curricular na sua natureza – a escola –, o currículo é também, no seu *modus operandi* (que na realidade lhe dá forma e regula o seu conteúdo e modo de passagem), uma produção organizacional, largamente conformada e imbricada nos mecanismos organizacionais que a escola configura (ROLDÃO, 2000, 2001, 2016; BARROSO, 1995, 1997; PERRENOUD, 1997, 2000).

A década de 1990 trouxe consigo a necessidade de diferenciar dois grandes níveis de decisão curricular: o nível central que estabelece o *core curriculum*; e o nível de decisão contextual, sustentado na maior autonomia das escolas na operacionalização e gestão contextualizada do currículo. A este processo tem sido atribuída a designação de *binómio curricular* (ROLDÃO, 2000, 2003a, 2003b, 2005a). Contudo, sua configuração é bem diversa em países de tradição centralista ou em países com percursos históricos de grande autonomia local. Daí que se deve acautelar a tentação de assimilar genericamente a expressão *currículo nacional* a *core curriculum*. Nos países tradicionalmente centralizados, o currículo nacional era, desde o início, o único que existia. O novo nesses contextos é a autonomia curricular das escolas, que causa perturbação por justamente romper o carácter único do currículo prescrito nacionalmente. Já nos países de grande autonomia local, como os nórdicos e os anglo-saxónicos, ocorre o inverso: o currículo nacional surge como um fator de regulação comum à tradicional diversidade de opções que as escolas podiam adoptar, normalizadas apenas por avaliações externas, aparecendo por isso como mais constrangedor para as escolas.

Malcolm Skilbeck (1994, p. 93) define assim o conceito de *core curriculum*, situando historicamente a sua necessidade política em face da massificação escolar e conseqüente diversidade dos públicos:

- Em síntese, o *core curriculum* consiste naquelas aprendizagens consideradas **básicas e essenciais** para todos: **básicas** porque oferecem um suporte sobre o qual as aprendizagens subsequentes podem construir-se e proporcionam o domínio das ferramentas metodológicas que permitem a cada um prosseguir a sua própria aprendizagem.
- São **essenciais** na medida em que apetrecham os aprendentes para uma participação satisfatória e efectiva na vida social e cultural.
- Trata-se daquela dimensão do currículo total que é genericamente **comum a todas as escolas**, definida em parceria pelos decisores centrais e locais, e depois **reinterpretada por cada escola**.

As dimensões do *core curriculum* referem-se necessariamente ao máximo das aprendizagens comuns a garantir, e não à ideia simplificadora de mínimos que empobreceriam o currículo. Em texto do Australian Curriculum Development Centre (1980¹ apud SKILBECK, 1994, p. 97), estabelece-se que esse denominador comum nacional – o *core curriculum* – deverá conter a prescrição nacional de aspetos comuns quanto a:

- áreas de conhecimento e experiência a abordar;
- processos de aprendizagem a dominar;
- situações e contextos de aprendizagem a experimentar.

A COMPLEXA OPERACIONALIZAÇÃO DO BINÓMIO CURRICULAR

No debate curricular e político sobre a proclamada deslocação de níveis de decisão para a escola e professores, tem faltado situar com clareza os dois fatores que condicionam em absoluto uma tal deslocação: de um lado, a própria natureza dos conteúdos curriculares e do trabalho do seu desenvolvimento (vulgo o *como*, os métodos didáticos, os modos dominantes da prática docente, o acto de *ensinar* e *ensinar o quê*, tal como é vivido nas escolas, bem como os dispositivos – ou ausência deles – estabelecidos para o trabalho do aluno); e, de outro, a incontornável estruturação que é dada ao currículo

¹ AUSTRALIAN CURRICULUM DEVELOPMENT CENTRE. *Core curriculum of the Australian schools: what it is and why it is needed*. Canberra: Curriculum Development Centre, 1980.

pelo esqueleto organizacional da instituição escolar, espantosamente imutável há quase dois séculos: a organização dos tempos, dos espaços, da produção letiva, do ato – aliás atos segmentares – de ensinar (ROLDÃO, 2001; BENAVIDES; BRASLAVSKY, 2007).

Assim, o desiderato político-curricular que permite compreender as políticas que anteriormente designamos como estruturadas em termos de “binómio curricular” orientou-se no sentido de induzir as seguintes mudanças:

- um equilíbrio entre o *core curriculum* – a referência central – e a autonomia das escolas no sentido de promover esse *core curriculum* numa *variedade de formatos contextuais* relacionados com as diversidades dos estudantes e dos seus contextos – adoção de um *binómio curricular* (ROLDÃO, 2003b);
- o abandono quer de um *currículo único e uniforme* para ser “cumprido” (tradição centralista), quer de *uma diversidade de currículos de qualidade desigual*, correspondendo aos níveis e contextos das escolas (tradição descentralizada) em favor de “um *core curriculum* comum para todas as escolas [...] *reinterpretado por essas escolas*”(SKILBECK, 1994, p. 96, grifos nossos);
- a adoção da organização do currículo por *competências* no sentido de viabilizar a diversidade de abordagens.

Entende-se aqui *competência* no sentido de mobilização integrada de conhecimento, experiência e disposições que capacitam o aprendente a confrontar-se com e a responder a situações complexas e contextuais (TARDIF, 1996; LE BOTERF, 1994).

O papel das competências no desenvolvimento do binómio curricular que enquadra a decisão sobre o currículo nesses dois níveis (central comum e local diverso), de forma articulada, traduz-se nas seguintes dimensões que a competência permite desenvolver no interior do currículo:

- competência como organizador curricular do conhecimento – por oposição a simples sequências de temas ou listas de conteúdos;

- competência como mobilizador dos diferentes conhecimentos das disciplinas em torno da construção do currículo como um todo;
- competência como transformação de corpos estáticos de saberes em saberes-em-uso (ROLDÃO, 1995).

Essa mudança de paradigma configurou assim, no caso português, a prescrição de um binómio curricular, que se traduziu na definição de aprendizagens curriculares essenciais comuns – *core curriculum* (SKILBECK, 1990) –, o que ficou expresso no que se denominou por *Currículo Nacional do Ensino Básico – Competências essenciais* (PORTUGAL, 2001b, revogado em 2011), articuladas com a diversidade de PCEs, a carga das escolas e correspondentes a contextos diferenciados (ROLDÃO, 1999b, 2000, 2003a).

AS DIFICULDADES DAS POLÍTICAS DE BINÓMIO CURRICULAR EM FACE DA PERSISTÊNCIA DO FORMATO ORGANIZACIONAL DA ESCOLA - DESENVOLVIMENTOS RECENTES

Passadas duas décadas sobre as políticas anteriormente caracterizadas como binómio curricular, observa-se um duplo efeito, identificado em documentos internacionais (UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION – UNESCO, 2015; OECD, 2014): por um lado, os currículos nacionais em formato *core curriculum* (os dos países com tradição curricular localista, nomeadamente os nórdicos), nas sucessivas revisões após as primeiras versões dos anos 1990, exibem alguma tendência para se tornarem mais extensos e mais detalhados (FINNISH NATIONAL BOARD OF EDUCATION – FNBE, 2016). Contudo, nesses sistemas, nomeadamente os nórdicos, tal reforço, também expressando a pressão da competição em estudos internacionais, de que se destacam as sucessivas edições do Programme for International Student Assessment (PISA), não se traduziu em perda significativa da tradição autonómica das escolas como gestoras efetivas dos seus currículos, mesmo quando mais pesados, na medida em que a cultura escolar e a cultura docente nestas sociedades são fortemente assentes na autonomia da gestão local do currículo, por meio dos poderes

locais e postos em acção pelas escolas e seus professores. Na Finlândia assistimos, desde 2015-2016, a um outro avanço nesse caminho de autonomia gestonária, instituindo para o nível da escolaridade básica (1-9) uma gestão do currículo nacional organizada em torno de projetos curriculares temáticos, geridos pelas equipas de professores, articulados com espaços curriculares disciplinares que permanecem no espaço curricular, mas em menor percentagem (FNBE, 2016).

Por outro lado, nos países de tradição administrativa centralista (Portugal, Espanha, França, Itália), as tentativas de criação de dinâmicas de gestão contextualizada do currículo pelas escolas não ganharam a projecção desejada na operacionalização efetiva de um nível local de decisão curricular situado nas escolas. No caso português, este nível de decisão traduzia-se na produção dos já referidos PCEs (objeto da investigação avaliativa que neste texto se descreve), por sua vez, desenvolvidos em PCT – projetos curriculares de turma (DL 6/2001). Contudo, a estrutura tradicional da organização da gramática escolar (TYACK; CUBAN, 1995; CUBAN, 1990) fez subsumir ou tornar pouco operantes os processos de decisão tentados, que adiante se descrevem. Tal esbatimento foi acompanhado de decisões políticas centrais do ciclo político seguinte na rotação partidária, que desvalorizavam esta linha da autonomia curricular, reforçando a dimensão do controlo central nomeadamente via exames externos em todos os ciclos, políticas que se situaram entre 2011 e 2015.

Mas importa sublinhar que estes patamares de decisão curricular da escola foram largamente postos em causa, mesmo antes da viragem político-partidária do poder central, pelas próprias escolas e professores que transformaram os projetos curriculares (PCE e PCT) em documentos de registo meramente burocráticos, esvaziados da sua finalidade de gestão contextualizada do currículo e do seu desenvolvimento nas práticas de ensino, como se definiam na sua intencionalidade declarada (ROLDÃO, 2005a). A investigação destaca também o papel da ação dos professores no processo curricular, tornando essa uma linha de necessidade de intervenção futura no plano do currículo (OECD, 2014; HATTIE, 2009)

A adoção da designação Currículo Nacional para o currículo essencial comum, que fazia grande sentido por exemplo nos países nórdicos pela sua novidade, em Portugal nunca foi entendida como uma das partes do referido binómio de decisões curriculares, já que o referente único da cultura curricular portuguesa, muito centralista, sempre foi, por definição, o nacional, do qual emanam as prescrições que estão enraizadas profundamente no agir e no pensar das escolas e professores.

Os decisores políticos portugueses atuais – de novo no âmbito da rotação político-partidária dos governos, o que em Portugal, lamentavelmente, condiciona excessivamente as macropolíticas educativas que deviam ser socialmente consensualizadas e desenvolvidas em médio prazo –, defendendo novamente uma política curricular atualizada, convergente com as tendências internacionais, de reforço da dimensão contextual e da necessidade de assegurar melhores níveis de qualidade de ensino e de sucesso das aprendizagens curriculares (OECD, 2015; UNESCO, 2014, 2015; ANANIADOU; CLARO, 2009), recolocaram desde 2015 na agenda política a discussão da questão da gestão curricular e correlativa contextualização e flexibilização do currículo, decorrendo em Portugal neste momento diversos desenvolvimentos dessa retoma em curso, que se defrontam com idênticas resistências da cultura de escola e de professores, que a investigação internacional também tem evidenciado (OECD, 2013, 2014; HATTIE, 2009).

No plano internacional, importa ainda considerar que os desenvolvimentos mais recentes retomam a mesma problemática (UNESCO, 2014, 2015; OECD, 2017), largamente centrada na discussão das dimensões essenciais dos currículos e nas estratégias do seu “emagrecimento” – configurando de novo a discussão do *core curriculum* e das políticas de maior eficácia da gestão contextualizada do currículo pelas escolas, como se documenta nos *working papers* já acessíveis do Projeto *Skills and Education 2030* (OECD, 2017).

No âmbito da OCDE, e em outras sedes de investigação curricular, discute-se neste momento a reconfiguração de políticas curriculares. Constituem temas centrais desse debate:

- a procura do sucesso alargado das aprendizagens;
- a seleção das aprendizagens curriculares essenciais e consequente “emagrecimento curricular”;
- a articulação do currículo com competências transversais;
- a maior adequação do currículo a novas necessidades sociais, políticas e económicas (OECD, 2017; UNESCO, 2015, 2016; ANANIADOU; CLARO, 2009);
- a maior centralidade atribuída à ação e ao profissionalismo docente na construção de um currículo de sucesso (OECD, 2014).

Parte desta permanência, na segunda década do milénio, da problemática da contextualização curricular, originada nos anos 1980 e 1990, pode explicar-se pela imutabilidade relativa das estruturas organizativas em que o currículo se desenvolve. A escola tem mantido até hoje a estrutura organizativa e as lógicas de ensino que a instituíram em épocas passadas, quando os públicos eram muito mais restritos e homogêneos e as finalidades menos ambiciosas para a maioria, embora mais seletivas para um leque reduzido da população. É nessa lógica que toda a orgânica da escola ainda funciona, bem como as correlativas visões da ação dos professores e suas metodologias de ensino dominantes (ROLDÃO, 2012). Situam-se também neste domínio toda a teorização e múltiplas contradições associadas às orientações de política educativa, largamente consensuais no plano discursivo, no sentido da chamada “escola inclusiva” (CLARK, 1999; AINSCOW, 1991).

Emergem na última década movimentos que iniciam abordagens consistentes a esta ruptura no modo organizacional e pedagógico de fazer passar o currículo com a eficácia que o deve legitimar e que não tem tido, tomando como referentes os níveis de insucesso persistentes (ROLDÃO, 2016). Destacam-se o Projeto 2020, desenvolvido por uma rede de colégios na Catalunha, e os desenvolvimentos já referidos do currículo da Finlândia (RED JESUÏTES EDUCACIÓ, 2015; FINNISH NATIONAL BOARD OF EDUCATION – NBF, 2016).

É dessa teia de contradições que aqui nos ocupamos. Considerou-se pertinente desenvolver um estudo centrado

na avaliação de 12 PCEs apresentados por uma rede de escolas no ano letivo de 2016/2017. Neste estudo, está a ser desenvolvida uma avaliação consistente dos PCEs, acompanhada de inquirição dos atores, não só para clarificar de que modos este quadro teórico-curricular e político foi apropriado pelos professores, como também para produzir conhecimento útil para o atual desenvolvimento de renovadas políticas de autonomia curricular das escolas.

AUTONOMIA DAS ESCOLAS: DO PROJETO EDUCATIVO AO PROJETO CURRICULAR

Em Portugal, Espanha e França, desde o final dos anos 1990 até o início do século XXI, conforme anteriormente sublinhamos, a literatura sobre a gestão curricular e PCEs, incluindo modelos de referência para o desenvolvimento curricular nas escolas, foi muito profícua. Importa assim situar a emergência e os entendimentos do omnipresente conceito de projeto, designadamente convocado no dispositivo Projeto Educativo (PE), bem como sobre o conceito de autonomia que lhe está associado, dos quais decorre o PCE, objeto desta investigação. A definição concetual impõe-se pela pluralidade de sentidos que encerram estes conceitos.

Desde o final dos anos 1980, afirma-se, na literatura, que vivemos numa civilização de projetos (BARBIER, 1991) que se tornaram numa espécie de “micro-ideologia da ação” quotidiana (BARROSO, 1992). O termo projeto tem origem nas teorias da administração, a partir do momento em que a planificação e gestão estratégica foram consideradas práticas essenciais do sucesso das empresas.

A planificação estratégica foi criada como um “processo de conceber um futuro desejado, bem como os meios reais para o alcançar” (GODET, 1985, p. 80). A planificação está associada à prospetiva como “uma reflexão para iluminar a ação presente, à luz dos futuros possíveis (GODET, 1989, p. 80). A premência de preparar o futuro levou a que, na primeira fase da planificação, se estabeleça um projeto para a empresa, iniciando com a definição de um diagnóstico para detetar as ameaças e as oportunidades do ambiente estratégico, para identificar as forças e fraquezas dos recursos

da empresa (humanos, técnicos, comerciais e financeiros) (GODET, 1985, p. 171).

A realização do projeto implica a coexistência de uma lógica do desejo que valoriza a dimensão simbólica (missão, visão, valores) e a lógica da ação que sublinha a dimensão operatória (BARROSO, 1992, p. 28) depois da elaboração do diagnóstico, identificando as estratégias de concretização a curto e médio prazos.

As teorias da administração passaram para o mundo educacional. Nos anos 1980, nos países da OCDE e das Comunidades Europeias (BARROSO; SJORSLEV, 1990, p. 121), as autoridades administrativas associaram a autonomia das escolas à elaboração de um projeto, neste caso do PE.

No contexto português, o Decreto-Lei n. 43/89, que consagra a autonomia da escola, relaciona-a com a formulação do PE:

A autonomia da escola concretiza-se na elaboração de um projeto educativo próprio, constituído e executado de forma participada, dentro de princípios de responsabilização dos vários intervenientes na vida escolar e de adequação a características e recursos da comunidade em que se insere. (PORTUGAL, 1989, preâmbulo)

O PE é, assim,

[...] um instrumento de planificação de longo prazo que enquadra a definição e a formulação das estratégias de gestão e do qual decorrem os planos operacionais de médio e de curto prazo. (BARROSO, 1992, p. 30)

O surgimento normativo (2001) do PCE, no âmbito do Decreto n. 6/2001 (PORTUGAL, 2001a), que geria a desejada nova política curricular, teoricamente em binómio, apresentava assim um desfasamento temporal em face da prescrição de produzir, no âmbito da autonomia de gestão da escola, um PE (1989). Este facto duplicou a dificuldade de se entender o PCE como o âmago do PE, e inserido nele. Na verdade, e apesar das normas o aconselharem, o Projeto Curricular de Escola foi visto nas escolas como uma díade documental, e não como um documento integrado.

Das indicações para a elaboração do PCE, destaca-se desde logo a conformidade com o PE, no sentido de “contribuir para a continuidade e coerência da atuação educativa de toda a equipa de professores da escola” (CACHAPUZ et al., 2004, p. 81; COSTA; RAMOS, 2004, p. 88).

O PCE deverá operacionalizar a parte central do PE que são as orientações pedagógico-curriculares, dada a natureza da escola como uma *instituição curricular* no sentido em que sua existência se justifica: por uma *finalidade curricular* – garantir uma passagem/apropriação de saberes tidos como necessários para certo tipo de competências numa dada época e contexto – e pela especificidade processual/metodológica da sua atuação no *desenvolvimento curricular dessa finalidade*, por meio da escolha e estruturação de modos de fazer aprender (métodos de ensino e organização) aquilo que socialmente se considera que deve ser aprendido (ROLDÃO, 2000, p. 4).

Percebe-se, assim, a centralidade do PCE, pois, em conjunto com o PE, constitui um meio fundamental para a construção da autonomia curricular da escola e dos professores, destinado a proporcionar a melhoria da qualidade das aprendizagens e do sucesso educativo dos alunos, ao ser adaptado precisamente aos respetivos contextos. Este documento potencia a gestão curricular das escolas/professores com o objetivo de partilhar e fixar estratégias comuns sobre o *quê, quando e como avaliar* (GIMENO, 1989, p. 14). Reconhece-se, assim, a importância deste instrumento para a autonomia das escolas no sentido que implica o processo de decisão e deliberação curriculares ao nível de cada instituição educativa favorável à melhoria das aprendizagens (MORGADO, 2003, 2011).

Na literatura o projeto curricular é concebido como um instrumento de gestão pedagógica-curricular contextualizador, que reúne pelo menos duas características: reconstrutor do currículo nacional, no sentido de se adequar às características dos contextos locais da sua aplicação; e fomentador de uma cultura de reflexão sobre os processos de ensino, o que apenas é possível por meio do trabalho cooperativo entre os professores (CARMEN; ZABALA, 1991; PACHECO, 1998; ROLDÃO, 1999b; PACHECO; MORGADO, 2002; RODRIGUES; SÁ-CHAVES, 2004; COSTA; RAMOS, 2004).

São estas ideias-chave que atravessam as definições, por exemplo, de Carmen e Zabala (1991, p. 16) ao determinarem o PCE como um

[...] conjunto de decisões articuladas, partilhadas pela equipa docente de uma escola, tendentes a dotar de maior coerência a sua atuação, concretizando as orientações curriculares de âmbito nacional em propostas globais de intervenção pedagógico-didática adequadas a um contexto específico,

ou de Roldão (1999a, p. 44) ao referir que

[...] por projeto curricular entende-se a forma particular como, em cada contexto, se reconstrói e se apropria um currículo face a uma situação real, definindo opções e intencionalidades próprias, e construindo modos específicos de organização e gestão curricular, adequados à consecução das aprendizagens que integram o currículo para os alunos concretos daquele contexto.

A gestão curricular como um processo de tomada de decisões, como já se afirmou anteriormente, inclui dois grandes níveis de decisão curricular: o *nível central*, que se relaciona com o currículo nacional e os modos de ensinar delineados pela administração central; e o *nível institucional*, que diz respeito à contextualização curricular da escola tendo como referência o currículo nacional que se materializa num PCE. Nestes dois níveis de deliberação curricular, é possível identificar seis campos de decisão que são comuns, embora trabalhados com níveis de operacionalização diversos que sistematizamos no Quadro 1.

QUADRO 1 – Níveis e campos de decisão curricular

CAMPOS DE DECISÃO	NÍVEL: ADMINISTRAÇÃO CENTRAL	NÍVEL: INSTITUCIONAL - ESCOLA
1. Filosofia distintiva e oferta educativa	Função social e educativa da escola.	- A partir do diagnóstico e da identidade descrita no PE: fundamentar a filosofia educativa e a oferta educativa.
2. Opções e prioridades	Opções e prioridades curriculares a nível nacional (p.e. melhorar o desempenho na língua materna em 30% no sistema; ou dar prioridade à aprendizagem científica e tecnológica).	- Em que aspetos curriculares - <i>core curriculum</i> - investir mais, em face das características e necessidades da população; - Que conteúdos de aprendizagem não contidos no <i>core curriculum</i> integrar e por quê. Oferta de aprendizagens em campos não cobertos pelo currículo nacional.
3. Organização das aprendizagens	Competências à saída de sistema educativo. Corpo de aprendizagens requerido para o domínio dessas competências que devem ser adquiridas por todos os alunos nas disciplinas (ou outros formatos).	- Articulação entre as atitudes e valores, capacidades e aptidões e conhecimentos definidos no Perfil do aluno: (ex.: redigir relatório científico nas competências - então na sala deve conter um momento de realização do mesmo); - Modos de organização das aprendizagens diversificados: áreas disciplinares e espaços multidisciplinares; (i.e. oficinas de práticas profissionais ou tradicionais).
4. Métodos e estratégias de ensino e avaliação	Metodologias de ensino que se valorizam face às competências socialmente necessárias. Orientações sobre os procedimentos e modalidades de avaliação.	- Estratégias de ensino que a escola privilegia em função dos valores que assume como sua proposta educativa (métodos, técnicas, recursos); - Princípios/Orientações sobre a avaliação das aprendizagens dos alunos (modalidades, instrumentação, campos).
5. Modos de funcionamento e organização da escola e das aulas	Princípios orientadores das opções organizativas das aprendizagens.	Opções organizativas da escola, em termos de enquadramento das atividades de aprendizagem: Critérios pedagógicos de constituição das turmas com organização flexível de tempos e espaços, dentro das turmas (por exemplo, grupos temporários de alunos); variação dos tempos de apoio; Justificação da unidade de tempo de referência e criar dispositivos para a sua flexibilização; Organização de horários e distribuição de serviço docente em termos de maximizar as aprendizagens dos alunos; Recomendações e critérios sobre a comunicação e reuniões com os pais. Princípios enquadramentos das opções organizativas do trabalho dos professores: - Estabelecimento de tempos e espaços para o trabalho colaborativo.
6. Avaliação do resultado de cada uma das opções do projeto curricular	Avaliação das aprendizagens do <i>core curriculum</i> conseguidas a nível nacional através de instrumentos externos adequados. Reformulação do <i>core curriculum</i> quando se revele inadequado, insuficiente ou rígido face às necessidades de formação dos cidadãos.	- Avaliação das aprendizagens propostas no projeto curricular da escola (interna e externa); - Construção de instrumentos de avaliação adequados; - Avaliação do trabalho curricular desenvolvido nas turmas, nos órgãos intermédios e nos órgãos de gestão.

Fonte: Adaptado de Roldão (1999a, p. 60-61).

METODOLOGIA

Privilegiamos uma metodologia de cariz qualitativo e interpretativo com algumas características da investigação-ação para este estudo das lógicas atribuídas pelas escolas à construção e uso dos PCEs – de que aqui se analisa uma parte dos dados.

Numa primeira fase, recolheram-se os PCEs de uma rede portuguesa de escolas. De um total de cerca de 40 escolas, apenas 12 tinham elaborado o PCE, que foi submetido a uma análise de conteúdo.

Numa segunda fase, com o objetivo de aprofundar o estudo sobre as práticas de gestão curricular, foram selecionadas três escolas com projetos curriculares mais desenvolvidos para a aplicação de entrevistas em *focus group*. Os grupos foram constituídos por professores(as) que assumem cargos de lideranças intermédias como os coordenadores de departamento e/ou de ciclo (no caso do 1º ciclo de escolaridade do ensino básico), os coordenadores dos diretores de turma e alguns diretores de turma. As lideranças intermédias têm como uma das suas funções liderar as equipas que, entre outros objetivos, implementam o projeto curricular no quotidiano dos agrupamentos de escolas.

Tendo em conta os resultados da análise de conteúdo, do *focus groups*, duma revisão da literatura sobre gestão curricular, estamos a atualizar a obra *Gestão curricular: fundamentos e práticas* (Roldão, 1999b) no sentido de disponibilizarmos um referencial com orientações para a elaboração de PCEs. Neste texto apresenta-se apenas a avaliação parcial que resulta da análise de conteúdo dos PCEs.

ANÁLISE DE CONTEÚDO

A análise de conteúdo dos projetos curriculares foi realizada com base num sistema de categorias misto, de natureza dedutiva e indutiva, definido com base no conhecimento teórico produzido sobre a problemática em questão e a partir de uma leitura flutuante do *corpus* em estudo. O livro *Gestão curricular: fundamentos e práticas* (ROLDÃO, 1999b) permitiu identificar um conjunto de decisões curriculares (ver Quadro 1) que guiou a construção do sistema de categorias dos PCEs das escolas em análise.

Na identificação das categorias, aplicámos as duas funções de análise de conteúdo enunciadas por Bardin (2006[1977]), ou seja, a função de “administração de prova”, sob a forma de orientações que conduzem a leitura dos documentos, e a função heurística decorrente da leitura daqueles

e do objetivo da nossa investigação. Como unidade de registro, optamos pelo recorte ao nível semântico e, como unidade de contexto, privilegiamos a frase. Já como unidade de enumeração das categorias temáticas, privilegiou-se a presença/ausência.

O sistema integrou um total de 31 categorias (Anexo 1) e a análise foi realizada com o auxílio do programa MAXQDA 12.3.1, que permite o recurso aos métodos mistos, possibilitando a introdução/gestão, para além de um sistema de categorias, de uma base com “variáveis de documento” qualitativas e quantitativas com informações de caráter mais descritivo sobre o *corpus* em análise. A forma mais usual de utilização destas variáveis é no recurso aos métodos mistos (KUCKARTZ, 2014). Assim, é possível “ativar”/selecionar determinadas variáveis quantitativas e analisar os resultados em função delas. Por exemplo, “ativando” a variável “número de alunos da escola”, poderíamos perceber o que distinguia as escolas com mais e menos alunos em relação às categorias em análise.

Neste caso, procedemos a uma aplicação diferente destas “variáveis de documento” (Anexo 2), mobilizando-as como informações relacionadas/adicionais ao sistema de categorias de forma a complementarem a sua análise.

RESULTADOS

CRÍTICA DA FONTE

De forma geral, o registro escrito dos PCEs está muito vinculado aos documentos legais que os enquadram, podendo incluir longas citações da legislação ou reprodução de partes. Os conceitos do discurso das políticas curriculares internacionais, replicados nos normativos legais das políticas curriculares nacionais, como “trabalho colaborativo”, “gestão curricular contextualizada” ou “diferenciação pedagógica” não operacionalizados em medidas concretas, multiplicam-se nos PCEs e parecem remeter para um discurso retórico, possivelmente induzido pela avaliação externa das escolas por não ser suportado, na maioria dos casos, por medidas concretas. Essa ausência de operacionalização pode ainda

significar o “efeito transformador” falacioso da “bondade dos princípios” (BARROSO, 2014) que consubstancia um equívoco recorrente nas práticas educativas, ao se acreditar que determinados conceitos/princípios enunciados nos documentos reguladores conferem aos seus destinatários capacidade de ação.

De referir ainda que nos PCEs acumula-se uma lista de informações/definições que ora enquadram disposições normativas para os docentes ou para a elaboração de documentos pedagógicos, como guiões para elaboração do Projeto Curricular de Turma; ora abrangem especialmente a organização da escola, como a descrição do serviço de Psicologia e Orientação Escolar (Quadro 2). O que inscreve os PCEs entre o registo do regulamento e do relatório, de demonstrar à tutela ou à comunidade educativa, por exemplo, os serviços que a escola dispõe ou os apoios educativos que disponibiliza.

QUADRO 2 – Lista de informações/definições dos Projetos Curriculares de Escola/agrupamento

DISPOSIÇÕES NORMATIVAS PARA OS DOCENTES	DISPOSIÇÕES SOBRE A ORGANIZAÇÃO DA ESCOLA
<ul style="list-style-type: none"> - Critérios de nomeação da coordenação das estruturas de orientação educativa (ex.: critérios de nomeação do diretor de turma, coordenador dos cursos profissionais, etc.). - Perfis de cargos (ex.: perfis dos professores titulares de turma/diretores, de turma/coordenadores, de departamento, etc.). - Cargos da coordenação das estruturas de orientação educativa e tempo da componente letiva e não letiva (cargos e redução da componente letiva). - Guiões (ex.: guião para elaboração do Projeto Curricular de Turma; guião para elaboração do Dossier Pedagógico dos Cursos Profissionais/Vocacionais/Educação e Formação, etc.). 	<ul style="list-style-type: none"> - Estruturas de orientação educativa (ex.: enumeração dos departamentos curriculares e das suas funções). - Serviços (ex.: descrição do funcionamento do Serviço de Psicologia e Orientação Profissional, etc.). - Educação especial (ex.: definição do conceito de NEE; das unidades de apoio, das adequações no processo de avaliação, etc.). - Componente de apoio à família (pré-escolar) (ex.: horário disponível). - Centro para a Qualificação e o Ensino Profissional (ex.: atividades e funções). - Associação de Pais (ex.: a sua existência por escola do agrupamento e atividades). - Descrição de espaços (ex.: descrição da sala de informática). - Assistentes operacionais e técnicos (ex.: descrição das funções). - Exames nacionais (ex.: descrição dos anos dos exames). - Calendário escolar (ex.: descrição do início e fim do ano letivo e períodos de férias). - Substituição de aulas (ex.: procedimentos em caso de permuta). - Mudança de turma ou de estabelecimento de ensino do grupamento (ex.: procedimentos e critérios de aceitação). - Pedidos de transferência provenientes de outros estabelecimentos de ensino (ex.: critérios de aceitação). - Apoios educativos (ex.: descrição e objetivos). - Transição e retenção de ciclo (ex.: critérios).

Fonte: 12 Projetos Curriculares de Escola/agrupamento de uma rede de escolas portuguesas.

FILOSOFIA DISTINTIVA E OFERTA EDUCATIVA

O nível de decisão curricular sobre a filosofia distintiva e oferta da escola está praticamente ausente dos PCEs. A sua maioria não apresenta uma reflexão crítica do diagnóstico do PE na perspectiva da ação pedagógico-curricular da escola em face do contexto e inicia com a descrição das matrizes curriculares por ciclos de ensino ou pelo calendário escolar. Apenas um PCE reproduz duas tabelas do PE – contendo a população escolar por ciclo e estabelecimento de ensino, o número de docentes por ciclo e cargo pedagógico e o número de assistentes técnicos e operacionais – que não são acompanhadas por um texto crítico.

Como consequência, as linhas de adequação da oferta educativa com base no diagnóstico do PE não estão presentes nos PCEs em análise. À exceção de quatro PCEs, a oferta educativa apenas é deduzida a partir da enumeração das matrizes curriculares que em alguns projetos ocupa uma parte substancial do espaço (até um máximo de 14 páginas).

A dimensão simbólica do PCE, designadamente os valores, a visão da escola tendo como referência o PE, está completamente desvalorizada no *corpus* em análise. Assim, ficam por ser colocadas perguntas como: que pretende esta escola alcançar a curto e a médio prazos? Que pretende melhorar na sua imagem e no seu serviço? Qual é o “rosto da escola” em que ela se quer rever no futuro?

Porém, a filosofia e metodologias de ensino que as escolas valorizam como estratégia geral a adotar pelos departamentos curriculares ou pelos professores ao nível da sala de aula são, de alguma forma, concebidas em cinco PCEs que privilegiam a diferenciação curricular como princípio organizador do currículo a ser transmitido.²

Nos restantes PCEs, a diferenciação curricular é ainda invocada, mas em contextos específicos, fazendo depender a sua operacionalização destes, isto é, a partir da avaliação diagnóstica e/ou dos apoios educativos, destacando-se ainda um PCE que, aos dois contextos anteriores, acrescenta a supervisão pedagógica.

² Não citamos excertos de texto exemplificativos da análise dos resultados para garantir o anonimato das escolas que foram objeto do presente estudo.

AS OPÇÕES E PRIORIDADES

As opções/áreas prioritárias selecionadas pelas escolas para a melhoria da ação educativa e desejavelmente articuladas com as dos PEs não constam da maioria dos PCEs, à exceção de quatro, já que dois reproduzem na íntegra as dos PEs. Os demais PCEs focam a sua intervenção no ambiente da aprendizagem, no trabalho colaborativo, na valorização da relação escola/comunidade – no que respeita à constituição de parcerias ou ao envolvimento das famílias na escola –, na formação dos recursos humanos e na melhoria dos resultados escolares. Nesse sentido e na sequência do acima referido, a ausência de articulação entre os PCEs e os PEs é um aspeto a sublinhar, pois apenas seis escolas referem o PE para definirem as prioridades curriculares.

Os aspetos curriculares essenciais a investir mais, em face das características e necessidades da população, estão presentes na maioria dos PCEs (oito) articulados com a necessidade de garantir a consecução das aprendizagens do currículo tido por nuclear. Nesse sentido, valoriza-se o reforço do ensino e das competências a desenvolver tanto em torno de saberes estruturantes, a saber, a Língua Portuguesa ou a Matemática, como no âmbito da cidadania e das TIC (PORTUGAL, 2012).

As medidas de promoção do sucesso escolar proliferam em todos os PCEs e são variadas, porém, na maioria são entendidas como medidas de remediação aparentemente devidas a limitações prévias e deterministicamente percebidas a respeito dos alunos e dos seus contextos, e não a possíveis mudanças ao nível do trabalho pedagógico ou da organização escolar.

Nesse aspeto, os PCEs em análise parecem remeter para alguma dificuldade das respetivas escolas em se conceberem como possíveis instâncias responsáveis pela reprodução de desigualdades educativas, apesar de toda a literatura, desde os *Herdeiros* de Bourdieu e Passeron (1964), demonstrar empiricamente também a sua contribuição para a manutenção das desigualdades sociais traduzidas em desigualdades educativas. De facto, parece faltar nos PCEs uma leitura sociológica das práticas escolares e pedagógicas que poderia

ultrapassar o que parece ser uma concepção das medidas de promoção dos resultados escolares, entendidas em termos meramente do *deficit* cognitivo dos alunos.

Entre as medidas para a promoção do sucesso escolar, destacam-se os apoios educativos, a constituição de grupos de homogeneidade relativa, as atividades de apoio ao estudo (como as salas de estudo) e a diferenciação curricular concebida apenas no âmbito destas medidas de remediação. A medida pedagógica mais referenciada para a melhoria dos resultados escolares é a coadjuvação em sala de aula que está plasmada nos normativos legais (PORTUGAL, 2012).

ORGANIZAÇÃO DAS APRENDIZAGENS

No que respeita à oferta de aprendizagens em campos não cobertos pelo currículo nacional, as escolas dispõem da possibilidade de introduzir, no 1º, 2º e 3º ciclos de escolaridade, uma disciplina de Oferta Complementar e, nos 7º e 8º anos, além desta, a disciplina de Oferta da Escola (Despacho n. 10-A/2015 de 19 de junho de 2015), em articulação com a disciplina de Tecnologias da Informação e da Comunicação (TIC). Apenas dois PCEs não discriminam a disciplina da Oferta Complementar, porém, esmagadoramente selecionam o ensino da Cidadania com as mais variadas denominações (Formação Cívica, Educação para a Cidadania, Educação para a Saúde e Cidadania, Educação e Cidadania, Desenvolvimento de Competências Pessoais e Sociais, Saber Estar, Literacias e Cidadania). As exceções são dois PCEs que, no 1º ciclo, oferecem Iniciação à Programação e Ensino Experimental das Ciências e três que, no 3º ciclo, para além da Cidadania, ainda oferecem Língua e Cultura Alemã, Inglês e Oficina de Escrita e Comunicação, em algum(uns) dos respetivos anos de escolaridade. A justificação mais recorrente, enunciada ou implícita, para a seleção destes saberes é que complementam o currículo prescrito ou os princípios enunciados.

A seleção privilegiada do saber da Cidadania pelas escolas pode dever-se às indicações emanadas da administração central, como sugere o Decreto-Lei n. 139/2012:

[...] pretende-se que a educação para a cidadania enquanto área transversal seja passível de ser abordada em todas

as áreas curriculares, não sendo imposta como uma disciplina isolada obrigatória, mas possibilitando às escolas a decisão da sua oferta nos termos da sua materialização disciplinar autónoma (PORTUGAL, 2012, preâmbulo).

Na Oferta de Escola, a seleção dos saberes recai nas áreas das Expressões (Educação Tecnológica e Artes), à exceção de Estudos e Projetos Europeus (cidadania europeia), Robótica e Produção de Imagem e Animação, em três PCEs respetivamente. Isso significa que apenas dois PCEs selecionam saberes que não constituem o currículo nacional e que, no caso da Robótica, configura uma aproximação ténue da escola aos saberes ligados à área da sua municipalidade, pela proximidade com um Instituto Politécnico que tem uma forte aposta nesta área.

Na seleção de conhecimento para currículo, as escolas parecem privilegiar saberes que decorrem ou complementam o currículo prescrito.

Nas áreas não disciplinares que incluem as atividades de enriquecimento curricular, os PCEs elencam, em sua maioria, uma enorme variedade de projetos, clubes/oficinas oferecidos, não sujeitos à avaliação, que procuram um enriquecimento das áreas de formação pessoal, científica, artística, desportiva e social dos alunos. As justificações para a seleção destas atividades centram-se no complemento ao currículo prescrito, mas inscrevem-se ainda, embora de forma ténue, na contextualização curricular, procurando a adaptação a certas atividades profissionais ou lúdicas da comunidade (Clube de Xadrez, Clube de Robótica) e também como pretexto para a construção de laços identitários ou como forma de apoio a disciplinas com maus resultados escolares.

Em relação às competências a serem adquiridas pelos alunos, poucos são os PCEs a conceber um conjunto de competências à saída do ensino básico (um PCE) ou secundário (dois PCEs), ou até mesmo um conjunto de competências transversais aos ciclos de ensino (um PCE). Porém, para esta situação pode contribuir o facto de Portugal não ter, no nível da administração central, um perfil de competências desde 2011,

quando se revogou o Currículo Nacional (Despacho n. 17169/2011, de 23 de dezembro). No sistema português, o caráter prescritivo da cultura da administração central e a sua naturalização nas escolas são muito fortes, o que explica esta dependência da prescrição central do currículo.

MÉTODOS E ESTRATÉGIAS DE ENSINO E AVALIAÇÃO

A filosofia e metodologias de ensino apresentadas como estratégicas, mas depois operacionalizadas nas disciplinas/anos/ciclos de ensino, que a escola privilegia em função dos valores que assume como sua proposta educativa, estão presentes apenas num PCE que destaca as metodologias ativas, mas equaciona também a expositiva no nível do 1º, 2º e 3º ciclos e ensino secundário.

As técnicas pedagógicas apenas são referidas em três PCEs, que compreendem o trabalho de grupo, a realização de atividades experimentais e o trabalho de projeto.

Os recursos pedagógicos são equacionados também por três PCEs, que sublinham a integração das TICs na sala de aula. O manual como recurso específico do ensino está fora do registo da gestão curricular. Esta omissão sugere algum questionamento a desenvolver em futuras investigações, uma vez que o manual é o recurso pedagógico básico e mais utilizado na prática dos professores e escolas. A sua não referência pode ser entendida como reflexo de uma naturalização que exclui outras possibilidades, ou uma intenção eventualmente involuntária de ocultação do recurso ao manual como principal fonte do currículo.

Assim, nos métodos e estratégias de ensino, os PCEs são altamente deficitários. As práticas de ensino, que constituem um dos elementos principais do currículo e corporizam o seu desenvolvimento, estão praticamente ausentes, configurando uma contradição entre a natureza do projeto curricular e o seu uso.

A avaliação das aprendizagens dos alunos ocupa um espaço central em todos os PCEs, pela importância, reconhecida em alguns destes documentos, da avaliação como um elemento estruturante do processo de ensino-aprendizagem, quer seja como elemento regulador das práticas pedagógicas,

quer seja como fator certificador das aprendizagens ou da orientação do percurso escolar, discurso aliás parafraseado a partir dos normativos legais (PORTUGAL, 2012).

A generalidade dos PCEs estabelecem, sobretudo, como modalidades de avaliação, a diagnóstica, formativa e sumativa e, como campos, os conhecimentos, competências e comportamentos/attitudes/valores, apresentando exemplos de vários instrumentos de avaliação. No que respeita à definição do campo dos comportamentos/attitudes/valores, os PCEs apenas mencionam itens a avaliar, como a participação do aluno ou a assiduidade, e nem sempre definem de forma clara o que se pretende com aqueles conceitos.

Dessa forma, como as práticas avaliativas são suscetíveis de diferentes interpretações, suscitadas por entendimentos concetuais divergentes que conduzem a variadas concretizações práticas da avaliação, seria necessário encontrar pontos de convergência facilitados pela clarificação concetual quer do entendimento das modalidades de avaliação, quer do que se entende por comportamentos/attitudes/valores e de que forma são incorporados na avaliação (PACHECO, 2002; HARLEN, 2007; ROLDÃO; FERRO, 2015).

MODOS DE FUNCIONAMENTO E ORGANIZAÇÃO DA ESCOLA E DAS AULAS

O nível de decisão curricular sobre os modos de funcionamento e organização da escola, no âmbito do enquadramento das atividades de aprendizagem, é uma das partes mais desenvolvidas nos PCEs, sendo preponderantes as matérias sobre as estratégias para a constituição das turmas, os horários dos professores, os horários dos alunos e a distribuição do serviço docente. Constata-se que a maioria das medidas preconizadas consta dos normativos legais, poucas são as medidas avançadas pelas escolas.

As estratégias dirigidas à família e ao município são mais raras, possivelmente por constituírem uma parte significativa dos PEs, como tivemos oportunidade de comprovar a partir da sua consulta.

A gestão estratégica ao nível do espaço das aulas e dos tempos está praticamente ausente nos PCEs. O espaço da

aula não é concebido numa perspetiva pedagógica, mas especialmente no sentido logístico a julgar pelas medidas elencadas apenas em três PCEs: estabelecer a coabitação limitada de turmas no ginásio; garantir a atribuição de sala própria para disciplinas de cariz mais prático; e fixação de uma sala para cada turma.

Já a gestão dos tempos letivos apresenta-se muito uniforme em todas as escolas, sobretudo organizados em períodos de 45 minutos, sendo que raramente se selecionam diferentes tempos letivos entre ciclos ou entre disciplinas. Porém, nesta matéria a legislação portuguesa é bastante flexível. O Decreto-Lei n. 6/2001, que estabeleceu os princípios orientadores da organização e da gestão curricular do ensino básico, determinava que a carga horária semanal se organizava em múltiplos de 45 minutos, sobretudo, em períodos de 90 minutos (PORTUGAL, 2001a). Uma década mais tarde, o Decreto-Lei n. 139/2012 procedeu a uma revisão de normas para o desenho curricular, de forma a permitir maior autonomia às escolas na organização das atividades letivas, incluindo a não obrigatoriedade, no 2º e 3º ciclos, da organização dos horários de acordo com tempos letivos de 45 minutos ou seus múltiplos, desde que as escolas respeitassem as cargas horárias semanais previstas nas matrizes curriculares do ensino básico (PORTUGAL, 2012).

A uniformidade da seleção da unidade de tempo letivo nos PCEs encontra-se em consonância com um estudo recente do Conselho Nacional da Educação (CNE) em parceria com o Projeto *Curriculum Monitor* desenvolvido no Centro Interdisciplinar de Ciências Sociais da Universidade Nova de Lisboa (ALMEIDA et al., 2017). Este estudo mostra, a partir de uma amostra representativa de escolas do 5º e 9º anos do ensino básico de Portugal Continental, no ano letivo de 2015/2016, num total de 231 Unidades Orgânicas (UO), que a maioria dos seus horários apresenta uma carga horária semanal organizada em períodos de 45 minutos (72,7%), seguindo-se os tempos letivos de 50 minutos (25,5%). Como podemos verificar pela Tabela 1, apenas duas UO organizam os tempos por unidades de 60 minutos, uma por unidades de 70 minutos e outra por tempos diferenciados nos 5º e 9º anos e por disciplinas.

TABELA 1 - Unidade de tempo letivo nas Unidades Orgânicas e nas turmas de 5º e 9º anos de escolaridade. 2015/2016

UNIDADES ORGÂNICAS	UNIDADE DE TEMPO LETIVO										TOTAL
	45 MINUTOS		50 MINUTOS		60 MINUTOS		70 MINUTOS		DIFERENCIADO		
	N	%	N	%	N	%	N	%	N	%	
Sem contrato de autonomia	128	71,91	47	26,40	2	1,12	0	0,00	1	0,56	178
Com contrato de autonomia	40	75,47	12	22,64	0	0,00	1	1,89	0	0,00	53
Total	168	72,73	59	25,54	2	0,87	1	0,43	1	0,43	231

Fonte: CNE e Projeto *Curriculum Monitor*, 2017.

Em relação às justificações para a escolha da unidade de tempo letivo, apesar de nos PCEs se acumularem páginas com as matrizes curriculares com os tempos das disciplinas, apenas um PCE justifica sua opção pelos 45 minutos com base numa maior concentração das atividades letivas no período da manhã (seis tempos) e numa melhor gestão da afetação da carga horária por disciplina e área disciplinar.

Os princípios enquadradores das opções organizativas do trabalho dos professores, designadamente o trabalho colaborativo a começar pela articulação vertical, são mencionados pela quase totalidade dos PCEs, sendo a realização/planificação de atividades comuns, a troca de informações sobre processos dos alunos e a garantia da sequencialidade das aprendizagens, as práticas mais referenciadas, seguidas pela respetiva calendarização.

Já a articulação horizontal apresenta uma fraca presença nos PCEs, a acrescer à ausência completa de calendarização das poucas medidas avançadas, o que, ao contrário da articulação vertical, pode significar uma prática ainda pouco enraizada no quotidiano destas escolas. A articulação horizontal parece remeter mais para um discurso retórico induzido pela avaliação externa das escolas por não ser sustentado, na maioria dos casos, em medidas concretas e respetiva calendarização, mas sim por expressões que implicam a necessidade de “Reunir para a interdisciplinaridade” (ALMEIDA, 2017); ou, ainda, reenviar para o “efeito transformador” falacioso da “bondade dos princípios” (BARROSO, 2014), um equívoco recorrente nas práticas da escola e dos professores,

ao acreditar-se que determinados conceitos/princípios enunciados nos documentos reguladores (não operacionalizados) atribuem aos seus destinatários capacidade de ação.

As medidas de articulação da biblioteca escolar com o trabalho pedagógico têm alguma presença nos PCEs, possivelmente pela adesão das escolas ao Plano Nacional de Leitura lançado pelo governo em 2006.³ Em três PCEs, regista-se a existência de articulação entre a educação literária desenvolvida nas salas de aula e as iniciativas desenvolvidas pela biblioteca escolar, no âmbito do Plano Nacional de Leitura. Outras formas de articulação com a biblioteca escolar são sugeridas pela disponibilização de apoio pedagógico a professores e pela realização de atividades em articulação com as práticas letivas, como o apoio às matérias correntes nas salas de aulas fornecendo materiais de consulta atualizados aos alunos.

A articulação entre as áreas disciplinares e não disciplinares raramente se inscreve no registo dos PCEs e, tal como a articulação com a biblioteca escolar, não é referida a calendarização para eventuais reuniões.

A formação dos professores é outra dimensão pouco sublinhada nos PCEs: apenas duas escolas estabelecem as áreas de formação a investir, mas sem agendar calendarização para qualquer uma delas.

A AVALIAÇÃO DO RESULTADO DAS OPÇÕES DO PROJETO CURRICULAR

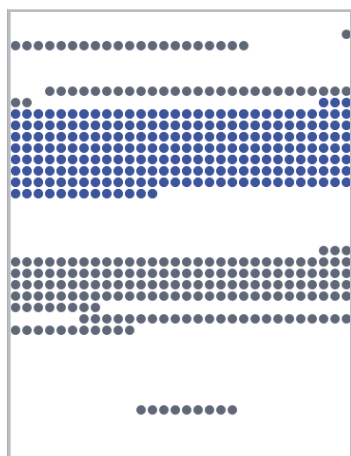
Em relação à avaliação da gestão curricular preconizada nos PCEs, esta parece pouco naturalizada na cultura destas escolas. Os campos de avaliação são mencionados por metade das escolas e caracterizam-se pela sua limitação, restringindo-se ao trabalho pedagógico desenvolvido nas turmas ou apenas às orientações nos documentos não seguidas pelos atores educativos. A maioria dos PCEs menciona a necessidade da sua avaliação periódica, mas raramente especifica a calendarização, indicadores ou os responsáveis pela prática avaliativa.

Com efeito, a avaliação dos PCEs configura uma parte fraca dos campos de decisão curricular que, por não serem concebidos essencialmente como um conjunto de ações estratégicas, dificilmente são passíveis de avaliação.

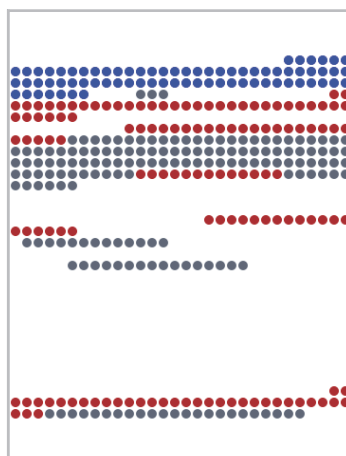
3 Programa lançado pelo Ministério da Educação, em 2006, com o objetivo central de elevar os níveis de literacia dos portugueses e colocar o país no nível do contexto europeu. Nas escolas do ensino básico e secundário materializa-se num conjunto de medidas e apoios financeiros para estimular nos alunos o prazer da leitura, intensificando o contacto com o livro na escola, seja nas salas de aula, seja nas respetivas bibliotecas. Consultar Alçada (2006).

Desde logo, o registo escrito dos PCEs deixa perceber que estas escolas ainda não se apropriaram do conceito de projeto. Este facto é, em parte, explicável pela lista de informações/definições nos PCEs, ora com disposições normativas para os docentes, ora com disposições sobre a organização da escola, como vimos no Quadro 2, oscilando entre o registo escrito do regulamento e do relatório, que ocupa uma parte substancial destes documentos, como podemos verificar pelos dois exemplos de Document Portrait⁴ (Figuras 1 e 2).

**Figura 1 - Document Portrait:
Projeto curricular n. 3**



**Figura 2 - Document Portrait:
Projeto curricular n. 5**



4 O Document Portrait é uma imagem total de cada documento em análise, que mostra tanto a sequência de categorias ou subcategorias que selecionamos para visualizar, como a percentagem que os segmentos de texto alocados naquelas ocupam na totalidade de cada documento analisado.

Code	
●	Provisions for teachers
●	Provisions on school organization
●	Curriculum matrices

Fonte: 12 Projetos Curriculares de Escola/agrupamento de uma rede de escolas portuguesas. Realizado em MAXQDA 12.3.1.

Em síntese, a análise avaliativa destes PCEs, intencionalizados na sua criação e prescrição como documentos de gestão estratégica, operacionalizadores supostamente do Projeto da Escola ou agrupamento no que se refere à dimensão curricular, revela escasso potencial estratégico, nomeadamente pelas seguintes características que a análise de conteúdo destes 12 PCEs permitiu destacar:

- caráter descritivo da organização sem retirar dela nem fundamentos nem implicações no plano pedagógico e curricular;

- proximidade a documento normativo pela inclusão de direitos, tempos atribuídos a tarefas e outras medidas de natureza regulamentar;
- invisibilidade quase total de opções da escola no plano curricular, limitadas à adoção de áreas transversais no campo da cidadania, à inclusão de algumas áreas complementares e ao reforço de tempos de trabalho para alunos com dificuldades nas áreas assumidas como estruturantes, como a Matemática e o Português;
- ausência maioritária de decisões relativas a modos de trabalho de professores ou a estratégias de ensino e aprendizagem preconizadas de forma estratégica;
- escassa visibilidade de orientações para o papel regulador da avaliação no desenvolvimento e melhoria das aprendizagens curriculares, expressando contudo de forma nominal as tipologias diagnóstica, formativa e sumativa, e sendo referidos genericamente os campos a abordar;
- insuficiência de elementos sobre avaliação do próprio projeto curricular, explicável em parte pelo facto de as ações e opções estratégicas serem pouco operacionalizadas, o que não possibilita sua análise avaliativa;
- difícil identificação de usos eficazes deste instrumento na melhoria do desempenho da escola e dos professores.

A natureza prescritiva da produção deste tipo de documentos, articulada com uma escassa tradição de autonomia na condução do ensino por parte das escolas, e ainda a tradição predominante de trabalho docente em lógica individual configuram explicações possíveis para esta escassa valia estratégica de instrumentos de gestão contextualizada do currículo que, todavia, ocupam grande quantidade de tempo e esforço na sua produção, sem correspondente evidência do seu impacto na transformação das rotinas da escola e na melhoria da sua eficácia curricular.

REFERÊNCIAS

- AINSCOW, M. (Ed.). *Effective schools for all*. London: David Fulton, 1991.
- ALÇADA, I. (Coord.). *Plano Nacional de Leitura: relatório síntese*. Lisboa: Ministério da Educação, 2006.
- ALMEIDA, S. de. Modos de apropriação das políticas educativas para uma gestão curricular contextualizada. In: CONGRESSO PORTUGUÊS DE SOCIOLOGIA, 9., 2017, Faro, PT. *Atas...* Faro, Portugal: Universidade do Algarve/Faculdade de Economia, 2017. p. 1281-1307.
- ALMEIDA, S. de et al. *Organização escolar: o tempo*. Lisboa: Conselho Nacional da Educação, 2017.
- ANANIADOU, K.; CLARO, M. *21st century skills and competences for new millennium learners*. Paris: OECD, 2009. (OECD Education Working Papers, n. 41).
- BARBIER, J.-M. *Élaboration de projets d' action et planification*. Paris: PUF, 1991.
- BARDIN, L. *Análise de conteúdo*. 4. ed. Lisboa: Edições 70, 2006. Conforme edição de 1977.
- BARROSO, J. Fazer da escola um projeto. In: CANÁRIO, R. (Org.). *Inovação e projeto educativo da escola*. Lisboa: Educa, 1992. p. 17-55.
- BARROSO, J. *Para o desenvolvimento de uma cultura de participação na escola*. Lisboa: Instituto de Inovação Educacional, 1995. (Cadernos de Organização e Gestão Escolar, n. 1).
- BARROSO, J. *Autonomia e gestão das escolas*. Lisboa: Ministério da Educação, 1997.
- BARROSO, J. Autonomia, contratos e direção das escolas. In: MACHADO, J.; ALVES, J. M. (Org.). *Melhorar a Escola-Educação para Todos: igualdade, diversidade e autonomia*. Porto: Universidade Católica Editora, 2014. p. 91-111.
- BARROSO, J. ; SJORSLEV S. *Administration and evaluation structures for primary and secondary schools in the twelve member states of the EC*. Brussels: Eurydice, 1990.
- BENAVOT, A.; BRASLAVSKY, C. (Ed.). *School knowledge in comparative and historical perspective*. Changing curricula in the primary and secondary education. Dordrecht, The Netherlands: Springer, 2007.
- BOURDIEU, P.; PASSERON, J. C. *Les héritiers: les étudiants et la culture*. Paris: Minuit, 1964.
- CACHAPUZ, F. et al. (Org.). *Saberes básicos de todos os cidadãos no século XXI*. Lisboa: Conselho Nacional de Educação, 2004. (Estudos e Relatórios).
- CARMEN, L. del; ZABALA, A. *Guia para la elaboracion, seguimiento y valoracion de proyectos curriculares de centro*. Madrid: Cide, 1991.

CLARK, C. et al. Theories of inclusion, theories of schools: deconstructing and reconstructing the “inclusive school”. *British Educational Research Journal*, v. 25, n. 2, p. 157-176, 1999.

COSTA, J.; RAMOS, M. Os professores e a (re)construção do currículo na escola: a construção de projetos curriculares de escola e de turma. In: COSTA, J. A. et al. *Gestão curricular: percursos de investigação*. Aveiro: Universidade de Aveiro, 2004. p. 79-97.

CUBAN, L. A fundamental puzzle of school reform. In: LIEBERMAN, A. (Ed.). *Schools as collaborative cultures: creating the future now*. New York: The Falmer, 1990. p. 71-77.

FINNISH NATIONAL BOARD OF EDUCATION. *National Curriculum*. Helsinki: FNBE, 2016.

GIMENO, J. *El curriculum: una reflexión sobre la práctica*. Madrid: Morata, 1989.

GODET, M. *Prospective et planification stratégique*. Paris: Economica, 1985.

GODET, M. Prospective et stratégie: approches intégrées. *Futuribles*, novembro 1989.

GONÇALVES, A. M.; MORGADO, J. C. Construção de um projeto curricular de escola: percepções e predisposições dos professores. 2013. Disponível em: <<http://hdl.handle.net/1822/23181>>. Acesso em: 18 maio 2017.

GOODSON, I. *The making of curriculum*. London: Falmer, 1988.

GOODSON, I. *Learning, curriculum and life politics*. The selected works of I. Goodson. Abingdon, Great Britain: Routledge, 2005.

HARLEN, W. Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, v. 33, n. 1, p. 15-28, 2007.

HATTIE, J. *Visible learning: a synthesis of 800 meta-analyses relating to achievement*. London: Routledge, 2009.

HUTMACHER, W. A escola em todos os seus estados: das políticas de sistemas às estratégias de estabelecimento. In: NÓVOA, A. (Coord.). *As organizações escolares em análise*. Lisboa: Educa, 1992. p. 45-76.

KUCKARTZ, U. *Qualitative text analysis. A guide to methods, practice and using software*. London: Sage, 2014.

LE BOTERF, G. *De la compétence. Essai sur un attracteur étrange*. Paris: Les Éditions d'Organization, 1994.

MORGADO, J. C. Projeto curricular e autonomia da escola: possibilidades e constrangimentos. *Revista Galego-Portuguesa de Psicología e Educación*, v. 8, n. 10, p. 335-344, 2003.

MORGADO, J. C. Projeto curricular e autonomia da escola: das intenções às práticas. *Revista Brasileira de Política e Administração da Educação*, v. 27, n. 3, p. 361-588, 2011.

MORGADO, J. M.; MARTINS, F. B. Projecto curricular: mudança de práticas ou oportunidade perdida? *Revista de Estudos Curriculares*, v. 6, n. 1, p. 3-19, 2008.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Environnement, école et pédagogie active*. Paris: OECD/CERI, 1991.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *La réforme des programmes scolaires: l'évaluation en question*. Paris: OECD/CERI, 1993.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *The curriculum redefined: schooling for the 21st century*. Paris: OECD/CERI, 1994.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Making the curriculum work*. Paris: OECD/CERI, 1998.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Innovating schools* (Schooling for tomorrow project). Paris: OECD/CERI, 1999.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Trends shaping education*. Paris: OECD, 2013.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *TALIS 2013 Results: an international perspective on teaching and learning*. Paris: OECD Documents, 2014.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Towards defining character/attitudes/values/behaviours. In: EDUCATION 2030 WORKSHOP, 16 November 2015. Paris: OECD, 2015. (EDU/EDPC/M(2015)5: Education Working Papers, n. 41). Draft summary.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Education 2030 International Curriculum Analysis: Twenty Years of Background Analysing Mathematics and Science Curricula. In: MEETING OF THE INFORMAL WORKING GROUP OF THE FUTURE OF EDUCATION AND SKILLS, 3rd, 3-4 May 2016, Paris, France. *Annals...* Paris: OECD, 2016.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Project "Future of education and skills: education 2030"*. Paris: OECD, 2017. (OECD Working papers).

PACHECO, J. A. *Projeto curricular integrado*. Lisboa: Ministério da Educação, 1998. (PEPT 2000, n. 18).

PACHECO, J. A. Critérios de avaliação na escola. In: ABRANTES, P.; ARAÚJO, F. (Org.). *Avaliação das aprendizagens: das concepções às práticas*. Lisboa: Ministério da Educação, DEB, 2002. p. 55-64.

PACHECO, J. A.; MORGADO, J. C. *Construção e avaliação do projeto curricular de escola*. Porto: Porto, 2002.

PACHECO, J. A.; PEREIRA, N. Estudos curriculares: das teorias aos projetos de escola. *Educação em Revista*, n. 45, p. 197-221, 2007.

PERRENOUD, P. *Construire des compétences dès l'école*. Paris: ESF, 1997.

- PERRENOUD, P. *Novas competências para ensinar*. São Paulo: Artmed, 2000.
- PORTUGAL. Ministério da Educação. Decreto-Lei n. 43, de 3 de fevereiro de 1989. *Diário da República*, Lisboa, n. 29/1989, Série I, 3 fev. 1989.
- PORTUGAL. Ministério da Educação. Departamento da Educação Básica. *Relatório do Projeto “Reflexão participada sobre os Currículos do Ensino Básico”*. Lisboa: ME-DEB, 1997.
- PORTUGAL. Ministério da Educação. Decreto-Lei n. 6, de 18 de janeiro de 2001. *Diário da República*, Lisboa, n. 15/2001, Série I-A, 18 jan. 2001a.
- PORTUGAL. Ministério da Educação. *Currículo nacional do ensino básico – competências essenciais*. Lisboa: Ministério da Educação, 2001b.
- PORTUGAL. Ministério da Educação. Decreto-Lei n. 139, de 5 de julho de 2012. *Diário da República*, Lisboa, n. 129/2012, Série I, 5 jul. 2012.
- RED JESUÍTAS EDUCACIÓ. *Projeto Horizonte 2020*. Barcelona, 2015.
- RODRIGUES, M.; SÁ-CHAVES, I. Gestão curricular e cultura de escola: a relação entre as dimensões curriculares instituída e instituinte. In: COSTA, J. et al. *Gestão curricular, percursos de investigação*. Aveiro: Universidade de Aveiro, 2004. p. 69-87.
- ROLDÃO, M. C. *O diretor de turma e a gestão curricular*. Lisboa: Instituto de Inovação Educacional, 1995. (Cadernos de Organização e Gestão Escolar, n. 2).
- ROLDÃO, M. C. Currículo: um processo de construção, gestão e formação reflexiva centrado na escola. In: ENCONTRO DE INVESTIGAÇÃO EM EDUCAÇÃO MATEMÁTICA, 6., 6, 7 e 8 de abril de 1997, Castelo de Vide. *Actas...* Portalegre: SPCE, 1998. p. 31-39.
- ROLDÃO, M. C. Currículo como projeto: o papel das escolas e dos professores. In: MARQUES, R.; ROLDÃO, M. C. (Org.). *Reorganização e gestão curricular no ensino básico*. Porto: Porto, 1999a. p. 11-21.
- ROLDÃO, M. C. *Gestão curricular*. Fundamentos e práticas. Lisboa: Ministério da Educação, Departamento de Educação Básica, 1999b.
- ROLDÃO, M. C. *Os professores e a gestão do currículo: perspectivas e práticas em análise*. Porto: Porto, 1999c.
- ROLDÃO, M. C. O currículo escolar: da uniformidade à contextualização – campos e níveis de decisão curricular. *Revista de Educação*, v. 9, n. 9, p. 81-92, 2000.
- ROLDÃO, M. C. A escola como instância de decisão curricular. In: ALARCÃO, I. (Org.). *Escola reflexiva e supervisão: uma escola em desenvolvimento e aprendizagem*. Porto: Porto, 2001. p. 67-77. (CIDInE, n. 14).
- ROLDÃO, M. C. *Avaliação de competências e gestão do currículo: as questões dos professores*. Lisboa: Presença, 2003a.
- ROLDÃO, M. C. *Diferenciação curricular revisitada: conceito, discurso e práxis*. Porto: Porto, 2003b.

ROLDÃO, M. C. *Formação e práticas de gestão curricular: crenças e equívocos*. Porto: Asa, 2005a.

ROLDÃO, M. C. Para um currículo do pensar e do agir: as competências enquanto referencial de ensino e aprendizagem. *En Direct de l'APPF*, suplemento, p. 9-20, fev. 2005b.

ROLDÃO, M. C. A função curricular da escola e o papel dos professores: políticas, discurso e práticas de contextualização e diferenciação curricular. In: COLÓQUIO LUSO-BRASILEIRO DAS QUESTÕES CURRICULARES, 4., Florianópolis, Brasil, setembro de 2008. *Anais...* Florianópolis: Universidade Federal de Santa Catarina, 2008.

ROLDÃO, M. C. Currículo, formação e trabalho docente. In: SANTOS, L. L. C. P.; FAVACHO, A. M. P. (Org.). *Políticas e práticas curriculares: desafios contemporâneos*. Curitiba, Brasil: CRV, 2012. p. 65-79.

ROLDÃO, M. C. Currículo, didáticas e formação de professores – a triangulação esquecida? In: OLIVEIRA, M. R. (Org.). *Professor: formação, saberes e problemas*. Porto: Porto, 2014. p. 91-103.

ROLDÃO, M. C. Currículo enquanto conhecimento necessário: eixos de uma discussão. In: CONSELHO NACIONAL DA EDUCAÇÃO. *Pareceres e Conferências*. Lisboa: CBE, 2016.

ROLDÃO, M. C.; FERRO, N. O que é avaliar? Reconstrução de práticas e concepções de avaliação. *Estudos em Avaliação Educacional*, v. 26, n. 63, p. 570-594, set./dez. 2015.

SACRISTÁN, J. G. *El curriculum: una reflexión sobre la práctica*. 4. ed. Madrid: Morata, 1994.

SKILBECK, M. *Curriculum reform: an overview of trends*. Paris: OECD, 1990.

SKILBECK, M. The core curriculum. In: ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *The curriculum redefined: schooling for the 21st century*. Paris: OECD, 1994. p. 95-100.

TARDIF, J. Le transfert des compétences analysé à travers la formation de professionnels. In: MEIRIEU, P. et al. (Ed.). *Le concept de transfert de connaissances en formation initiale et en formation continue*. Lyon: CRDP, 1996. p. 55-72.

TYACK, D.; CUBAN, L. *Tinkering toward utopia. A century of public school reform*. Cambridge: Harvard University, 1995.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. *Curriculum in the twenty-first century: challenges, tensions and open questions*. Paris: Unesco Education Research and Foresight, 2014. (ERF Working Papers, n. 9). Disponível em: <<http://unesdoc.unesco.org/images/0022/002294/229458e.pdf>>. Acesso em: 10 maio 2017.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. *The curriculum in debates and in educational reforms to 2030: for a curriculum agenda of the twenty-first century*. Geneva, 2015. (IBE Working Papers on Curriculum Issues, n. 15).

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. *Repensar a educação: rumo a um bem comum mundial?* Brasília, DF: Unesco, 2016.

YOUNG, M. *The curriculum of the future. From the new sociology of education to a critical theory of learning.* London: Falmer, 1998.

ZABALZA, M. Do currículo ao projeto da escola. In: CANÁRIO, R. (Org.). *Inovação e projeto educativo de escola.* Lisboa: Educa, 1992. p. 87-107.

MARIA DO CÉU ROLDÃO

Investigadora integrada do Centro de Estudos em Desenvolvimento Humano da Universidade Católica Portuguesa (CEDH/UCP), Porto, Portugal
mrceuroldao@gmail.com

SÍLVIA DE ALMEIDA

Investigadora integrada do Centro Interdisciplinar de Ciências Sociais da Universidade Nova de Lisboa (CICS.NOVA-UNL). Professora auxiliar convidada da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (FCSH-UNL), Lisboa, Portugal
silvia.almeida@fcs.unl.pt

ANEXO 1 – Sistema de categorias

CAMPOS DE DECISÃO CURRICULAR	CATEGORIAS
1. Filosofia distintiva e oferta educativa	Especificidade face à oferta educativa
	Especificidade face aos métodos de ensino
2. Opções e prioridades	Aspetos do <i>core curriculum</i> a investir
	Referências às orientações/programas/metas curriculares
	Medidas de promoção do sucesso escolar
3. Organização das aprendizagens	Seleção de um perfil de competências
	Justificação da seleção da oferta complementar
	Justificação da seleção da Oferta da Escola
	Justificação da seleção dos projetos
	Justificação da seleção dos clubes/oficinas
4. Métodos e estratégias de ensino e avaliação	Métodos pedagógicos
	Técnicas pedagógicas
	Recursos pedagógicos
	Princípios orientadores da avaliação
	Finalidades da avaliação
	Modalidades de avaliação das aprendizagens
	Campos da avaliação
5. Modos de funcionamento e organização da escola e das aulas	Estratégias da constituição de turmas
	Estratégias da constituição dos horários dos professores
	Estratégias da constituição dos horários dos alunos
	Estratégias da distribuição do serviço docente
	Estratégias relativas à família
	Estratégias relativas ao município
	Estratégias sobre a unidade de tempo letivo
	Estratégias relativas aos espaços
	Articulação das áreas disciplinares e não disciplinares
	Articulação horizontal
	Articulação vertical
	Articulação com a biblioteca escolar
	Formação dos docentes
6. Avaliação do resultado de cada uma das opções do projeto curricular	Campos de avaliação do trabalho curricular

Fonte: 12 Projetos Curriculares de Escola/agrupamento de uma rede de escolas portuguesas.

ANEXO 2 - Variáveis de documento

VARIÁVEIS DE DOCUMENTO
Reflexão crítica sobre o diagnóstico do PE
Operacionalização das prioridades do PE
Número de páginas com matrizes curriculares
Indicação da Oferta Educativa
Designação da Oferta Complementar no 1º ciclo
Designação da Oferta Complementar no 2º ciclo
Designação da Oferta Complementar no 3º ciclo
Designação da Oferta Complementar no ensino secundário
Designação da Oferta de Escola no 7º ano de escolaridade
Designação da Oferta de Escola no 8º ano de escolaridade
Número de projetos
Número de clubes/oficinas
Conceção de diferenciação curricular
Articulação horizontal: calendarização de reuniões
Articulação vertical: calendarização de reuniões
Articulação das áreas disciplinares e não disciplinares: calendarização de reuniões
Articulação com a biblioteca escolar: calendarização de reuniões
Indicadores de avaliação da gestão curricular
Calendarização da avaliação da gestão curricular
Intervenientes no processo de avaliação da gestão curricular

Fonte: 12 Projetos Curriculares de Escola/agrupamento de uma rede de escolas portuguesas.

Recebido em: 08 JUNHO 2017

Aprovado para publicação em: 26 JULHO 2017

IMPACTO DA BONIFICAÇÃO EDUCACIONAL EM PERNAMBUCO

**CLAYTON SIRILO DO VALLE FURTADO
TUFI MACHADO SOARES**

RESUMO

O artigo busca avaliar o impacto do Bônus de Desempenho Educacional, que constitui um incentivo pago aos funcionários das escolas da rede estadual de Pernambuco mediante o cumprimento de metas educacionais preestabelecidas. O objetivo principal é verificar se essa política de bonificação propiciou um melhor desempenho nas proficiências das disciplinas de Língua Portuguesa e Matemática entre 2008 e 2012. A metodologia para abordar essa questão consiste no uso de escores de propensão para produzir pareamentos, em que cada escola da rede estadual (que participa da política de bonificação) é comparada a uma escola similar da rede municipal (que não dispõe de tal política). Afere-se, como principal resultado, que as escolas estaduais obtiveram, no 9º ano do ensino fundamental, um ganho de cinco pontos em relação àquelas da rede municipal.

PALAVRAS-CHAVE ACCOUNTABILITY • AVALIAÇÃO DE POLÍTICAS EDUCACIONAIS • BÔNUS DE DESEMPENHO EDUCACIONAL • MÉTODO DAS DIFERENÇAS EM DIFERENÇAS.

IMPACTO DE LA BONIFICACIÓN EDUCACIONAL EN PERNAMBUCO

RESUMEN

El artículo tiene el propósito de evaluar el impacto del Bono de Desempeño Educacional, un incentivo que se les paga a los empleados de las escuelas de la red estadual de Pernambuco mediante el cumplimiento de metas educacionales preestablecidas. El principal objetivo es el de verificar si dicha política de bonificación propició un mejor desempeño en las asignaturas de Lengua Portuguesa y Matemáticas entre 2008 y 2012. La metodología para abordar este tema consiste en el uso de notas que producen cruces de datos, en que cada escuela de la red estadual (que participa de la política de bonificación) es comparada a una escuela similar de la red municipal (que no dispone de tal política). Como principal resultado, se verificó que las escuelas estaduais obtuvieron, en el 9º año de la educación básica, un aumento de cinco puntos en relación a las de la red municipal.

PALABRAS CLAVE ACCOUNTABILITY • EVALUACIÓN DE POLÍTICAS EDUCACIONALES • BONO DE DESEMPEÑO EDUCACIONAL • MÉTODO DE LAS DIFERENCIAS EN DIFERENCIAS.

IMPACT OF EDUCATIONAL BONUS IN PERNAMBUCO

ABSTRACT

The present article seeks to assess the impact of the Educational Performance Bonus, an incentive paid to employees of the state school network in Pernambuco, when they accomplish pre-established educational goals. The main objective of this paper is to verify whether this bonus policy led to better performance in Portuguese and Mathematics between 2008 and 2012. The methodology used to address this issue uses propensity scores to produce pairings, in which each school in the state network (participating in the bonus policy) is compared to a similar municipal school (which does not have such a policy). The main result shows that state schools have obtained, in the ninth grade, a gain of five points compared to those schools in the municipal network.

KEYWORDS ACCOUNTABILITY • ASSESSMENT OF EDUCATIONAL POLICIES • EDUCATIONAL PERFORMANCE BONUS • METHOD OF DIFFERENCES-IN-DIFFERENCES.

INTRODUÇÃO

Esse artigo analisa uma política de pagamento de incentivos salariais (bonificação), versão *high-stakes* das políticas de responsabilização em educação. Em se tratando de políticas *high-stakes* em educação, a modalidade por meio de pagamento de incentivos salariais é predominantemente adotada no Brasil, ao contrário de intervenções punitivas como as que ocorrem nos Estados Unidos, nas quais professores podem ser dispensados e escolas fechadas. De maneira geral, a concessão do pagamento do bônus às escolas considera o desempenho dos alunos a partir do cumprimento de metas acordadas previamente.

As políticas de responsabilização em educação vêm, aos poucos, se espalhando na gestão da educação brasileira. Em 2011, de acordo com Brooke (2011), pelo menos seis estados brasileiros (Amazonas, Ceará, Pernambuco, Rio de Janeiro, Minas Gerais e São Paulo) já apresentavam experiências de responsabilização e políticas de incentivos salariais e outros dois (Espírito Santo e Distrito Federal) estudavam a implementação desses modelos, havendo uma tendência de

aumento desses números com base na tônica do governo federal, que pode ser representada, dentre outros elementos, pela avaliação censitária da Prova Brasil e o Índice de Desenvolvimento da Educação Básica (Ideb).

Contudo, evidencia-se a importância de conhecer as consequências dessas políticas, pois, conforme aponta a literatura, os resultados têm sido bem variados; um resumo de diversas pesquisas pode ser encontrado em Pontual (2008).

Em se tratando de políticas de responsabilização em educação e suas consequências, o *No Child Left Behind* (UNITED STATES OF AMERICA, 2002) caracterizou-se por ser um dos maiores movimentos nos Estados Unidos e que, muito provavelmente, influenciou na modelagem dessa intervenção em outros países. A avaliação de seu impacto foi analisada no trabalho de Guisbond *et al.* (2012), que não identificaram ganhos significativos na adoção desse programa. Contudo, Hout e Elliott (2011) estabelecem alguns critérios para a validação e seleção dos estudos sobre responsabilização educacional. Tais elementos nem sempre são adotados pelas pesquisas, tornando assim suas conclusões menos robustas.

Essa modalidade de intervenção sob o enfoque da responsabilização educacional é bastante polêmica, até mesmo por conta da diversidade de resultados apontados anteriormente, e engendrou inúmeras críticas, das quais, as principais encontram-se nos trabalhos de Ravitch (2011) e Freitas (2012). Entretanto, a origem das mesmas varia desde argumentos mais ideológicos até os mais empíricos. No sentido de organizar o debate, Brooke (2013) empreendeu um enfoque que permite uma melhor compreensão dessas críticas, elucidando o caminho para os debates futuros.

Uma preocupação que se destaca refere-se ao impacto da responsabilização no estreitamento curricular, ou seja, as escolas e professores dariam maior ou quase exclusiva atenção aos conteúdos que seriam avaliados pelos testes empregados na política de responsabilização. Levantamento de pesquisas (AU, 2007) que abordaram o impacto das políticas de *accountability* no currículo aponta em diferentes direções, pois há trabalhos que indicam o estreitamento e outros a expansão do conteúdo curricular. Assim, o impacto no

currículo depende do desenho da política, do contexto no qual é aplicado e até mesmo da dimensão da matriz curricular antes e após a aplicação das políticas de *accountability*. Um estudo realizado nos estados norte-americanos pelo Conselho Nacional de Pesquisa (*National Research Council*) e financiado pelo *Hunt Institute* mostra que

[...] na ausência de padrões claros e concisos e sem um currículo a eles alinhado, os professores se apoiam nos testes estaduais para obterem diretrizes acerca do que devem ensinar. (THE HUNT INSTITUTE, 2012, p. 178)

Dessa forma, se os efeitos esperados dos programas são variados, o mesmo acontece em relação aos não desejados.

As pesquisas sobre responsabilização educacional vêm crescendo nos últimos anos. O trabalho de Corvalán (2006) reúne estudos relevantes sobre essa temática na América Latina. Já no Brasil, Brooke *et al.* (2011) abordam a utilização das políticas de incentivos salariais em alguns estados, fornecendo um bom suporte para conhecer um pouco mais sobre essa modalidade de intervenção no país.

Na literatura brasileira, existem alguns estudos sobre o desenho e o impacto das políticas de pagamentos de incentivos salariais em educação. Quanto aos diferentes desenhos, Ferraz (2009) abordou o pagamento de incentivos salariais de São Paulo e acrescentou, em sua pesquisa, o Bônus de Desempenho Educacional de Pernambuco (BDE), permitindo a análise do impacto dos diferentes desenhos de bonificação entre os dois estados. Em relação ao impacto, o estudo de Oshiro e Scorzafave (2011), realizado no estado de São Paulo, encontrou um ganho de 6,4 pontos para Matemática e 3,7 pontos para Língua Portuguesa na escala Saeb (Sistema de Avaliação da Educação Básica).

Diante da gradual implantação das políticas de bonificação, cabe uma questão: a utilização das políticas de bonificação em educação, no Brasil, tem produzido efeitos positivos no desempenho dos alunos? Por mais instigante que esse questionamento possa parecer, seria impossível esgotar esse tema tão abrangente no âmbito nacional, considerando o tamanho do Brasil e a diversidade de desenhos de bonificação utilizados.

A alternativa mais viável é o enfoque das análises em um sistema de bonificação estadual. Desse modo, a opção do presente estudo recaiu sobre o estado de Pernambuco. Alguns fatores foram decisivos para tal escolha, entre eles a presença de um sistema próprio de avaliação de larga escala desde 2000, a implementação da política do Bônus de Desempenho Educacional em 2008, além do cálculo de metas com base em um indicador sintético, o Índice de Desenvolvimento da Educação Básica de Pernambuco (Idepe), que conjuga desempenho nos testes com fluxo escolar. Alinhado a esses fatores, existe o fato de o estado ter apresentado um Ideb bem inferior às médias nacionais em 2007. Enquanto o Ideb para o total do Brasil, em 2007, era de 4,2 para os anos iniciais do ensino fundamental (EF), 3,8 para os anos finais e 3,5 para o ensino médio, em Pernambuco, esses valores correspondiam a, respectivamente, 3,6, 2,8 e 3,0 (BRASIL, 2007). Os valores baixos do Ideb, aliados à implantação dessa política de bonificação, esboçam um cenário interessante e de potencial amplitude para melhorias.

Dessa forma, o presente trabalho busca identificar o efeito produzido pelo BDE na rede estadual de Pernambuco entre 2008 e 2012, procurando verificar se foram conquistados melhores desempenhos nas proficiências de Língua Portuguesa e Matemática.

Antes de adentrar nas análises, é relevante salientar que essa intervenção foi baseada na remuneração variável por desempenho. Assim sendo, em 2008, por meio da Lei n. 13.486, de 01 de julho de 2008, instituiu-se o Bônus de Desempenho Educacional. Em seu Art. 1º, já se salienta a vinculação entre bonificação e desempenho:

[...] o Bônus de Desempenho Educacional – BDE, correspondente a uma premiação por resultados, destinado aos servidores lotados e em exercício nas unidades escolares da Rede Pública Estadual de Ensino. (PERNAMBUCO, 2008a, p. 1)

O bônus é pago para todos os funcionários das escolas estaduais que atingem o mínimo de 50% da meta preestabelecida, sendo que o seu valor é semelhante ao percentual de

cumprimento das metas; dessa forma, se a escola atingir 55% da meta, o valor do bônus será igual a esse percentual. Além disso, os funcionários devem estar alocados na escola por no mínimo seis meses no ano do cálculo da bonificação.

Pode-se dizer que esse processo de responsabilização por meio da bonificação consistiria em uma

[...] tentativa de melhorar os resultados das escolas mediante a criação de consequências para a escola ou para professores, sejam elas materiais ou simbólicas, de acordo com o desempenho dos alunos medido por procedimentos avaliativos estaduais ou municipais. (BROOKE, 2008, p. 94)

No tocante às percepções dos professores e diretores, pelo que pode ser percebido até o momento, os mesmos não participaram da construção desse processo em Pernambuco. Entretanto, o resultado de outra pesquisa indica que, nesse estado, não houve resistência por parte dos professores (SEGATTO, 2011). Ao que parece, a política de bonificação ficou confundida com uma melhoria salarial, o que geraria menos resistência daqueles que concorrem ao seu recebimento e, provavelmente, não contam com salários tão atrativos. Nesse caso, pode haver uma mobilização nas escolas para buscar o cumprimento das metas estabelecidas, o que poderia influenciar em mudanças diversas, tais como melhoria no processo de ensino-aprendizagem, gestão e relações intraescolares.

Em 2012, 56% das escolas estaduais de Pernambuco atingiram as metas e fizeram jus ao recebimento do bônus. Os pagamentos variaram entre R\$ 726 e R\$ 3.873, correspondendo a uma média de R\$ 2.976,24 por funcionário das escolas receptoras desse ano (PERNAMBUCO, 2013). Esse valor médio pode ser considerado expressivo se levarmos em conta que o teto da carreira do magistério (classe IV, faixa d, com doutorado), em 2012, era de R\$ 3.000,46. Dessa forma, há chances de o valor do bônus ser suficiente, de modo a fomentar algum tipo de motivação e mobilização nas escolas estaduais de Pernambuco.

Estudos realizados pelo Banco Mundial (FERRAZ; BRUNS, 2010; BRUNS; FERRAZ, 2011; BRUNS; FILMER; PATRINOS, 2011; BRUNS; EVANS; LUQUE, 2012) indicaram algum efeito do BDE

em Pernambuco, sobretudo no 9º ano do ensino fundamental em 2009, devendo-se dar continuidade nessas análises para verificar se os ganhos continuaram alguns anos após esses estudos. Foram verificadas algumas melhorias nas práticas pedagógicas, como, por exemplo, o aproveitamento do tempo dentro da sala de aula pelas escolas que receberam a bonificação. Ressaltam-se, ainda, um melhor aproveitamento do tempo no processo de ensino-aprendizagem e uma redução no tempo do professor fora de sala de aula, bem como no tempo consumido para organização e gerenciamento da sala.

No tocante à pressão exercida para o cumprimento das metas, salienta-se que se trata de um ingrediente inerente a essa modalidade de política. Contudo, se tal pressão sobre as escolas e seus membros não surtirem os efeitos esperados, a política se finda, uma vez que se constitui em um elemento de “desgaste” desnecessário. Obviamente, os resultados das análises de impacto também poderiam servir para subsidiar modificações no desenho dessas políticas, levando a uma espécie de aprimoramento para que os objetivos sejam atingidos com a menor quantidade possível de efeitos não desejados.

Entretanto, verificar o efeito de qualquer política não é algo simples, tendo em vista que a unidade de análise – no caso deste trabalho, a escola – pode estar submetida a mais de uma intervenção de políticas e programas, tanto do governo quanto de instituições privadas e ONGs. Por outro lado, além da diversidade de intervenções políticas que poderiam atuar nas escolas, tem-se todo um contexto socioeconômico e geográfico que pode influenciar nas comparações entre escolas. Assim, para cumprir o proposto nesse trabalho, requer-se o emprego de uma metodologia capaz de neutralizar ou, ao menos, minimizar o efeito de outras variáveis que possam influenciar no desempenho dos alunos.

METODOLOGIA

Em um estudo de impacto de uma política pública seria mais eficaz a realização de estudos experimentais, nos quais seriam escolhidos, aleatoriamente, um grupo de intervenção

(que concorreria pela bonificação salarial) e um grupo controle (que não seria submetido à política de bonificação). Essa metodologia permitiria o controle da influência de outras variáveis sob os indicadores utilizados para mensurar o impacto da adoção da política. Entretanto, é difícil a utilização desse procedimento para uma política educacional, pois, dentre as muitas questões que podem ser colocadas, está a de como excluir, *a priori*, um grupo de participar de tal política. Essa impossibilidade reforça a necessidade de projetos-piloto, já que

[...] o benefício de tal pesquisa seria inestimável, pois as evidências geradas permitiriam que a política finalmente adotada tivesse a maior probabilidade possível de alcançar seu objetivo de melhorar a aprendizagem dos alunos. (PONTUAL, 2008, p. 36)

No entanto, é raro no Brasil se pensar em avaliar determinada política antes de implantá-la em grande escala. Por esse e outros motivos, os estudos experimentais são inviáveis na avaliação de políticas educacionais e estudos observacionais.

O presente trabalho toma como fonte de análise as bases com as proficiências em Matemática e Língua Portuguesa do Sistema de Avaliação Educacional de Pernambuco (Saepe) e da Prova Brasil. Embora a ideia inicial para a análise fosse utilizar apenas as bases do Saepe, que possui aplicações anuais desde 2008, as justificativas para a utilização das análises em duas bases de dados, de origens distintas, foram se consolidando no decorrer do trabalho. A primeira justificativa é que o Saepe apresentou avaliações anuais entre 2008 e 2012, mas não teve uma avaliação em 2007. Considerando-se que o BDE teve início em 2008, seria interessante analisar o período de transição da sua implantação. Uma segunda justificativa é que o cálculo da bonificação não utiliza os resultados da Prova Brasil, o que seria, a princípio, mais adequado metodologicamente, como apontado por Hout e Elliott (2011, p. 5-6):

[para] evitar ter os resultados determinados pela inflação de escore que ocorre nos testes *high-stakes* ligados aos incentivos, os pesquisadores devem utilizar testes *low-stakes*

que não imitem os testes de *high-stakes* para avaliar como os incentivos baseados no teste afetam os resultados.

Por último, como ambas as avaliações encontram-se na mesma escala de proficiência do Saeb, torna-se possível realizar comparações, constituindo, assim, uma forma complementar de análise que permite também uma maior robustez nas conclusões.

A utilização da escola como unidade de análise deve-se ao fato de a premiação ser feita pelo desempenho da escola e não pelo desempenho dos alunos individualmente. Logo, as informações dos mesmos, tal como a proficiência, serão agregadas por escola.

Entretanto, o desempenho do aluno é determinado por inúmeros fatores (SOARES, 2007; SOARES, 2010) e os resultados da escola dependem, além das características socioeconômicas de seus alunos, de uma complexa rede de relações entre diferentes variáveis. A maior dificuldade reside em isolar o efeito das demais variáveis para que seja possível observar apenas a melhoria do desempenho proveniente das políticas de responsabilização.

Para resolver esses problemas, pelo menos em parte, utilizam-se a metodologia de escore de propensão (*propensity score*) e o subsequente pareamento das unidades escolares pelo pareamento dos escores de propensão (*matching score*). Essa técnica vem sendo aplicada no estudo de políticas de pagamentos de incentivos salariais a professores. No Brasil têm-se, por exemplo, as pesquisas realizadas sobre os efeitos do sistema de bonificação existente no estado de São Paulo (OSHIRO; SCORZAFAVE, 2011) e em Pernambuco (BRUNS; FERRAZ, 2011). Tais aplicações sugerem que essa metodologia se demonstra como uma boa alternativa para a observação do efeito do BDE em Pernambuco.

Conforme mencionado, métodos experimentais são difíceis de serem empregados em avaliações de políticas públicas educacionais, mas pode-se tentar um método observacional que se aproxime do método experimental no sentido de, tendo o grupo de tratamento (as escolas que receberam o bônus) já definido, construir um grupo de controle que funcione como um bom contrafactual do grupo tratado.

Esse grupo pode ser definido com o uso de procedimentos que minimizem o viés de seleção, por exemplo, por meio do controle de variáveis sabidamente associadas ao desempenho dos alunos ao final de um período escolar, tais como medidas da condição econômica e as proficiências prévias.

Embora o método proposto no trabalho não seja tão robusto quanto o experimental, no que tange à possível influência de outras intervenções, devemos esclarecer que o BDE era a principal política educacional atuante no estado de Pernambuco, no período analisado nesse artigo, até mesmo pela sua abrangência, visto que todas as escolas da rede estadual poderiam concorrer ao recebimento do bônus. A outra intervenção que temos conhecimento são algumas escolas da rede estadual que funcionariam em tempo integral; para evitar um possível viés, elas não entraram nos pareamentos. Além disso, como nosso grupo controle foi obtido dentro do próprio estado (rede municipal), procuramos reduzir o “risco” de o grupo controle estar submetido a outras ofertas de políticas e programas de outros estados, embora garantir-se disso seja praticamente impossível num estudo observacional e, talvez, nem mesmo em um experimental na área educacional.

O método de pareamento (*matching*) que usa os escores de propensão (*propensity score*) constitui uma alternativa frequentemente sugerida na literatura para esse fim e será usado na produção das análises acerca dos resultados do BDE. No pareamento, escolhe-se uma escola no grupo de controle (no caso, da rede municipal) com características semelhantes a cada escola no grupo de tratamento (rede estadual). Essa escola do grupo de controle funciona como o par da escola do grupo de tratamento, emulando-se, portanto, o resultado que ela teria se não participasse do programa, conforme observado por Pinto (2012, p. 85):

De outra forma, as hipóteses do pareamento postulam que, ao comparar dois indivíduos, um no grupo de controle e outro no grupo de tratamento, com as mesmas características observáveis, o único fator que diferencia os resultados destes indivíduos é a participação ou não no programa.

Ainda de acordo com Pinto (2012, p. 85), existem, pelo menos, dois pressupostos importantes para os resultados das análises de impacto baseadas na técnica do pareamento:

- ignorabilidade (não confundimento): as características observáveis utilizadas devem ser suficientes para controlar o resultado potencial na ausência de tratamento;
- sobreposição: as mesmas variáveis devem representar as características importantes das escolas nos indivíduos presentes dos grupos tratados e não tratados.

O termo *escore de propensão* apareceu pela primeira vez em 1983, em um texto de Rosenbaum e Rubin (1983) que descrevia a estimação de efeitos causais de um estudo de dados observacionais. Essa metodologia consegue reduzir as muitas variáveis que representam as características observáveis importantes a uma única dimensão (GUO; FRASER, 2010).

Uma das formas de estimar o *escore de propensão* é a utilização do modelo *logit* ou *probit* para estimar a probabilidade de participação em determinado programa ou política. Dessa maneira,

[...] o *escore de propensão* é uma pontuação de equilíbrio que representa um vetor de covariáveis. Neste contexto, um par tratado e controle compartilharia um *escore de propensão* semelhante, logo, os elementos do par poderiam ser comparáveis, embora possam diferir em valores de covariáveis específicas. (GUO; FRASER, 2010, p. 130)

O pareamento busca, então, aproximar os participantes dos dois grupos, em termos de *escore de propensão* estimado, ou seja, realiza o pareamento tentando reduzir a perda de participantes pela dificuldade de encontrar um par. Em sua forma mais simples, o pareamento utiliza alguma medida de distância entre tratados e grupo controle em relação a um conjunto de variáveis observáveis (vetor X).

A técnica de pareamento utilizada nesse trabalho é o *nearest neighbor matching* (vizinho mais próximo), por meio da qual os pares seriam formados como um caso de tratamento e outro de controle e como um critério para formação dos pares. Essa técnica emprega a menor diferença entre os *escores*

de propensão entre todas as possibilidades de pares, sendo que cada par formado remove o caso de tratamento e controle da base inicial em que estão os elementos sem pares.

Para evitar escolhas ruins, pode-se usar o *caliper matching*, que é um método de correção do pareamento que determina a distância máxima aceitável entre os escores de propensão. Outra técnica mais recente é o *optimal matching* (pareamento otimizado), que leva em consideração a distância total de todos os pares, podendo refazer pares já formados para otimizar o conjunto total dos pares (GUO; FRASER, 2010).

Após o pareamento, são realizadas as análises nos pares, com algum tipo de regressão ajustada, em que se possam observar as diferenças nos escores das covariáveis entre o grupo de tratamento e o de controle. Uma comparação entre os dois grupos, antes do pareamento, deverá indicar diferenças nas médias das variáveis observáveis. No entanto, após o pareamento, as médias deverão ficar muito próximas entre ambos.

É importante salientar que após o pareamento nenhum grupo, seja ele de controle ou intervenção, possui um desempenho inicial desfavorável. A técnica do escore de propensão e do pareamento engendra um “ponto de partida” comum entre intervenção e controle, sendo comparadas escolas com características semelhantes entre aquelas consideradas no pareamento. Assim, se o incremento da variação acontecer de forma semelhante entre controle e intervenção, isso é um indício de que a política não estaria induzindo a um esforço maior pelo grupo intervenção.

Uma vez formados os grupos de tratamento e controle, por alguma técnica de pareamento, pode-se utilizar o método das diferenças em diferenças (DD) para testar o impacto do tratamento. Em um primeiro momento, calcula-se a diferença das médias, em um intervalo de tempo, para o grupo controle e para o grupo tratamento. Em um segundo momento, faz-se a diferença dessas diferenças. Esse resultado seria o que se pode atribuir ao efeito do programa. Observe um exemplo fictício na Tabela 1.

TABELA 1 - Exemplo fictício do método das diferenças em diferenças considerando a proficiência média de duas redes públicas de ensino

GRUPO	ANO DA AVALIAÇÃO		DIFERENÇA (2012-2008)	DD
	2008	2012		
Escolas da Rede Estadual (Tratamento)	100	180	80	30
Escolas da Rede Municipal (Controle)	100	150	50	

Fonte: Elaboração dos autores.

Nesse exemplo, o impacto do programa seria de 30 pontos pelo método das diferenças em diferenças. Como vantagem, essa técnica lidaria bem com o viés de seleção associado às características não observáveis, principalmente as que não variam muito no decorrer do tempo. Algumas condições devem ser obedecidas ao se utilizar o método DD. De acordo com Foguel (2012), a composição dos grupos não deve se alterar de maneira significativa entre os períodos analisados. Além disso, é necessário garantir que os grupos não sejam afetados de maneira heterogênea por mudanças externas.

PAREAMENTO COM AS BASES DO SAEPE

As etapas escolares avaliadas, e que serão analisadas, correspondem ao 5º e 9º ano do ensino fundamental. Embora o Saepe também avalie o 3º ano do ensino fundamental, essa etapa não entra no cálculo de metas para o recebimento do BDE e, portanto, tais dados não serão aqui analisados. Por outro lado, o 3º ano do ensino médio não apresenta casos suficientes para a formação do grupo controle (rede municipal), logo, permanecerá fora das análises. Essa escolha apoia-se na composição da rede escolar pública de Pernambuco, conforme observado na Tabela 2, que apresenta a distribuição de alunos nos anos avaliados por rede de ensino.

Foram realizados dois pareamentos, um para o 5º ano e outro para o 9º ano do ensino fundamental. Antes de ser feita a regressão para o cálculo do escore de propensão, a base de dados foi organizada com os resultados dos alunos de ambas as redes nas avaliações do Saepe.

TABELA 2 – Distribuição absoluta e relativa dos alunos por série e rede de ensino em Pernambuco (2008-2012)

ANO	5º ANO EF		9º ANO EF		3º ANO EM	
	ESTADUAL	MUNICIPAL	ESTADUAL	MUNICIPAL	ESTADUAL	MUNICIPAL
	N	N	N	N	N	N
	%	%	%	%	%	%
2008	45.440	216.394	98.544	71.418	113.416	7.754
	17,4	82,6	58,0	42,0	93,6	6,4
2009	42.266	220.074	102.922	74.850	109.478	3.466
	16,1	83,9	57,9	42,1	96,9	3,1
2010	59.142	428.896	211.668	147.560	231.620	4.636
	12,1	87,9	58,9	41,1	98,0	2,0
2011	57.924	417.982	211.044	154.526	242.108	3.416
	12,2	87,8	57,7	42,3	98,6	1,4
2012	41.154	410.922	208.040	157.498	247.560	2.068
	9,1	90,9	56,9	43,1	99,2	0,8

Fonte: Elaboração dos autores a partir de microdados do Saepe (PERNAMBUCO, 2012).

As bases continuam, preliminarmente, escolas localizadas tanto na zona urbana quanto na rural. Contudo, as escolas rurais encontravam-se sobremaneira na rede municipal, não permitindo realizar o pareamento adequadamente nesse grupo, o que levou a descartá-las das análises. Além disso, foram também observados os seguintes critérios para incluir uma escola na análise:

- ter ao menos dez alunos em cada ano avaliado (5º e 9º anos);
- possuir proficiência em Língua Portuguesa e Matemática em todas as edições do Saepe entre 2008 e 2012;
- não mudar de rede de ensino entre 2008 e 2012.

A análise por meio da regressão logística que gerou os escores de propensão contou com as seguintes variáveis:

- *BDE* = variável dependente no modelo, admitindo-se o valor “0” para o grupo controle (escolas municipais) e 1 para intervenção (escolas estaduais);
- *PRF_MT_2008* = proficiência de Matemática em 2008;
- *PRF_LP_2008* = proficiência de Língua Portuguesa em 2008;
- *Nº_ALUNOS_ESC* = número de alunos na escola; o dado é relevante, pois o resultado dos alunos tende a ser maior em escolas menores (SOARES, 2004);

- *IDHM* = Índice de Desenvolvimento Humano Municipal,¹ que é um controle importante das diferenças entre os municípios;
- *ISE* = Índice Socioeconômico² dos alunos, que contribui para o controle de diferenças de origem socioeconômica; essa medida, quando agregada por escola, constitui um importante controle na observação das escolas (ALVES; SOARES, 2009).

1 Envolve a transformação das três dimensões por ele contempladas (longevidade, educação e renda) em índices que variam entre 0 (pior) e 1 (melhor) (ORGANIZAÇÃO DAS NAÇÕES UNIDAS, 2016).

2 O cálculo do ISE foi realizado no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd-UFJF), conforme Anexo I.

Após o cálculo do escore de propensão foi realizado o pareamento das escolas que possuíam o 5º e o 9º ano do ensino fundamental. O método para formar os pares foi o *vizinho mais próximo* (GUO; FRASER, 2010). Essa técnica emprega a menor diferença (mínimos) entre os escores de propensão entre todas as possibilidades de pares entre as unidades do grupo controle e de intervenção. Cada unidade pareada terá um único par. Na formação dos pares, foi observada, ainda, uma distância máxima aceitável entre os escores de propensão (*caliper matching*). Dessa forma, pares com diferenças muito altas foram retirados, evitando, assim, o pareamento de escolas que sejam pouco semelhantes segundo o conjunto das variáveis utilizadas. As Tabelas 3 e 4 apresentam as médias das variáveis, antes e após a realização do processo de pareamento das escolas. Como se pode observar, as diferenças existentes entre os grupos de controle e intervenção antes do processo de pareamento foram reduzidas de forma significativa. Essa observação é válida tanto para o 5º ano, com 137 pares formados (73,3% do total da rede estadual submetida ao pareamento), quanto para o 9º ano do ensino fundamental, com 304 pares formados (49,4% do total da rede estadual submetida ao pareamento).³

3 Conforme a Tabela 2, o fato de a rede municipal de Pernambuco no 9º ano ser menor do que a rede estadual limitou a escolha e a formação dos pares da rede estadual; o mesmo não ocorreu com o 5º ano, cuja oferta da rede municipal é maior, facilitando assim a formação dos pares. Essa diferença também será observada no pareamento com as bases da Prova Brasil.

TABELA 3 – Valores médios das variáveis utilizadas no cálculo dos escores de propensão das escolas públicas com 5º ano do EF, com base no Saepe, antes e depois do pareamento

VARIÁVEIS UTILIZADAS NO CÁLCULO DOS ESCORES DE PROPENÇÃO	ANTES DO PAREAMENTO		APÓS O PAREAMENTO	
	CONTROLE	INTERVENÇÃO	CONTROLE	INTERVENÇÃO
IDHM	0,657	0,696	0,683	0,684
Número de alunos por escola	57,7	50,6	49,9	52,7
ISE	0,086	0,269	0,192	0,202
Proficiência em Língua Portuguesa 2008	163,3	169,8	166,1	166,3
Proficiência em Matemática 2008	166,9	173,1	169,1	169,5

Fonte: Elaboração dos autores a partir dos microdados do Saepe (PERNAMBUCO, 2012).

TABELA 4 – Valores médios das variáveis utilizadas no cálculo dos escores de propensão das escolas públicas com 9º ano do EF, com base no Saepe, antes e depois do pareamento

VARIÁVEIS UTILIZADAS NO CÁLCULO DOS ESCORES DE PROPENÇÃO	ANTES DO PAREAMENTO		APÓS O PAREAMENTO	
	CONTROLE	INTERVENÇÃO	CONTROLE	INTERVENÇÃO
IDHM	0,635	0,672	0,648	0,649
Número de alunos por escola	60,8	67,0	63,1	60,7
ISE	-0,069	0,205	0,027	0,029
Proficiência em Língua Portuguesa 2008	209,3	215,9	211,1	211,1
Proficiência em Matemática 2008	220,8	225,7	221,9	222,0

Fonte: Elaboração dos autores a partir dos microdados do Saepe (PERNAMBUCO, 2012).

PAREAMENTO COM A BASE DA PROVA BRASIL

Conforme relatado anteriormente, as bases do Saepe não contemplaram o ano de 2007, impossibilitando, assim, as análises do ano que antecede a implantação do BDE, que ocorreu em 2008. Dessa forma, foram utilizadas as bases da Prova Brasil de 2007, 2009 e 2011. Como a base da edição de 2007 apresentava apenas escolas urbanas,⁴ esse mesmo recorte foi considerado para 2009 e 2011, pois, ainda que existam escolas rurais em 2009 e 2011, não seria possível acompanhar o desempenho dessas instituições partindo de 2007.

Foram adotados os seguintes critérios para a manutenção das escolas nas análises:

- ter ao menos dez alunos em cada ano avaliado (5º e 9º anos);

⁴ As escolas rurais passaram a ser avaliadas na edição da Prova Brasil de 2009.

- possuir proficiência de Língua Portuguesa e Matemática nas edições de 2007, 2009 e 2011 da Prova Brasil;
- não mudar de rede de ensino entre 2007 e 2011.

Os procedimentos que seguem são muito semelhantes aos realizados nas bases do Saepe e, portanto, serão descritos de maneira sucinta. Na regressão logística para a geração do escore de propensão, foram utilizadas as seguintes variáveis:

- *BDE* = variável dependente no modelo, admitindo-se o valor “0” para o grupo controle (escolas municipais) e 1 para intervenção (escolas estaduais);
- *PRF_MT_2007* = proficiência em Matemática em 2007;
- *PRF_LP_2007* = proficiência em Língua Portuguesa em 2007;
- *Nº_ALUNOS_ESC* = número de alunos na escola;
- *IDHM* = Índice de Desenvolvimento Humano Municipal.

Nesse pareamento não foi empregado o ISE, pois, em 2007, não existia esse indicador nas bases da Prova Brasil. Portanto, a estimação do escore de propensão foi obtida com uma variável a menos em relação ao cálculo do escore de propensão realizado no Saepe.

Após o cálculo do escore de propensão, foi feito o pareamento das escolas. O método para formar os pares foi o mesmo utilizado no pareamento do Saepe. As Tabelas 5 e 6 mostram as médias das variáveis antes e após a realização do processo de pareamento das escolas. As diferenças existentes antes do processo de pareamento foram reduzidas de maneira perceptível. Após o pareamento, estavam formados 238 pares no 5º ano (98,3% do total da rede estadual submetida ao pareamento) e 314 no 9º ano do ensino fundamental (54,7% do total da rede estadual submetida ao pareamento).

TABELA 5 – Valores médios das variáveis utilizadas no cálculo dos escores de propensão das escolas públicas com 5º ano do EF, com base na Prova Brasil, antes e depois do pareamento

VARIÁVEIS UTILIZADAS NO CÁLCULO DOS ESCORES DE PROPENÇÃO	ANTES DO PAREAMENTO		APÓS O PAREAMENTO	
	CONTROLE	INTERVENÇÃO	CONTROLE	INTERVENÇÃO
IDHM	0,654	0,681	0,682	0,679
Número de alunos por escola	53,3	46,4	47,4	46,6
Proficiência em Língua Portuguesa 2007	156,2	161,6	161,3	161,3
Proficiência em Matemática 2007	172,7	176,5	176,5	176,3

Fonte: Elaboração dos autores a partir dos microdados da Prova Brasil (BRASIL, 2011).

TABELA 6 – Valores médios das variáveis utilizadas no cálculo dos escores de propensão das escolas públicas com 9º ano do EF, com base na Prova Brasil, antes e depois do pareamento

VARIÁVEIS UTILIZADAS NO CÁLCULO DOS ESCORES DE PROPENÇÃO	ANTES DO PAREAMENTO		APÓS O PAREAMENTO	
	CONTROLE	INTERVENÇÃO	CONTROLE	INTERVENÇÃO
IDHM	0,631	0,668	0,637	0,638
Número de alunos por escola	60,8	63,5	60,9	58,6
Proficiência em Língua Portuguesa 2007	208,4	213,0	209,6	209,4
Proficiência em Matemática 2007	219,8	222,3	220,1	220,0

Fonte: Elaboração dos autores a partir dos microdados da Prova Brasil (BRASIL, 2011).

RESULTADOS

Após a formação dos pares de escolas foi aplicado o método das diferenças em diferenças, nas análises tanto do Saepe quanto da Prova Brasil, para verificar a efetividade do BDE. Dessa forma, caso o BDE tenha tido algum impacto no desempenho dos alunos, esse “ganho” seria superior e favorável ao grupo de intervenção (rede estadual) nas séries históricas analisadas.

RESULTADOS PARA O 5º ANO DO ENSINO FUNDAMENTAL

No tocante aos pareamentos realizados no 5º ano do ensino fundamental, conforme podemos observar na Tabela 7 (base do Saepe), há uma pequena diferença das médias favorável ao grupo de intervenção. Contudo, o teste *t* apontou diferenças significativas ($p < 0,05$) apenas nos intervalos

entre 2008 e 2011 em Matemática, que apresentaram um ganho médio em torno de 4 pontos do grupo de tratamento *vis-à-vis* o grupo controle. Em Língua Portuguesa, o intervalo entre 2008 e 2012 apresentou um ganho médio em torno de 3 pontos. No entanto, são resultados muito tímidos se consideramos esse intervalo de quatro anos.

TABELA 7 - Resultados da análise de diferenças em diferenças para as escolas públicas urbanas com 5º ano do EF, com base no Saepe (137 pares)

INTERVALO	CONTROLE		INTERVENÇÃO		DIFERENÇAS DAS MÉDIAS	I.C. 95%	P-VALOR
	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS			
MATEMÁTICA							
2008-2009	8,49	12,46	10,16	13,97	1,67	[-1,24 ; 4,57]	0,2594
2008-2010	7,73	14,91	9,79	19,74	2,06	[-1,88 ; 6,00]	0,3032
2008-2011	19,65	14,56	23,76	16,72	4,10	[0,54 ; 7,66]	0,0243*
2008-2012	21,56	16,83	24,87	15,97	3,31	[-0,46 ; 7,08]	0,0850
LÍNGUA PORTUGUESA							
2008-2009	1,67	10,22	2,11	10,59	0,43	[-1,97 ; 2,84]	0,7226
2008-2010	2,66	11,45	4,74	16,11	2,08	[-1,10 ; 5,25]	0,1980
2008-2011	8,63	12,10	11,27	13,46	2,64	[-0,31 ; 5,58]	0,0789
2008-2012	9,13	14,36	12,36	14,25	3,23	[0,10 ; 6,35]	0,0430*

Fonte: Elaboração dos autores a partir dos microdados do Saepe (PERNAMBUCO, 2012).

* Valor de p significativo até 0,05.

Com base nos resultados das bases do Saepe para o 5º ano, encontrados no limite, as mesmas análises foram feitas com a utilização das bases da Prova Brasil. Isso poderia trazer alguma informação, tendo em vista que seria possível captar os resultados do ano de 2007, antecedente à implantação do BDE. Entretanto, mesmo considerando a transição do ano de 2007 para 2009, os resultados (conforme demonstrado na Tabela 8) ficaram muito próximos dos encontrados com a base do Saepe, sendo favorável ao grupo de intervenção, com uma diferença das médias em torno de 5 pontos para Matemática no período de 2007 a 2011. Resultados mais tímidos,

com valores em torno de 3 pontos, foram observados em Língua Portuguesa no mesmo intervalo de tempo.

TABELA 8 - Resultados da análise de diferenças em diferenças para as escolas públicas urbanas com 5º ano do EF, com base na Prova Brasil (238 pares)

INTERVALO	CONTROLE		INTERVENÇÃO		DIFERENÇAS DAS MÉDIAS	I.C. 95%	P-VALOR
	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS			
MATEMÁTICA							
2007-2009	6,91	14,72	8,80	14,75	1,89	[-0,68 ; 4,47]	0,1489
2007-2011	9,79	14,89	14,63	17,29	4,84	[2,24 ; 7,44]	0,0003*
LÍNGUA PORTUGUESA							
2007-2009	4,58	12,88	4,75	12,56	0,17	[-2,05 ; 2,39]	0,8813
2007-2011	8,45	13,72	11,42	14,96	2,97	[0,61 ; 5,32]	0,0138*

Fonte: Elaboração dos autores a partir dos microdados da Prova Brasil (BRASIL, 2011).

* Valor de p significativo até 0,05.

Portanto, os resultados dos pareamentos para o 5º ano do ensino fundamental foram positivos apenas se consideramos as diferenças entre o início e o fim das séries históricas, tanto pelo Saepe quanto pela Prova Brasil. É importante destacar que, na transição entre o período que não existia o BDE, ou seja, de 2007 para 2009, as diferenças não foram estatisticamente significativas. Outro ponto relevante é que o uso da base de dados do Saepe, que foi empregada nos cálculos do BDE, ao que parece, não provocou uma inflação de escores quando comparada à da Prova Brasil. Tal defesa justifica-se pela proximidade dos resultados obtidos entre as duas bases (Saepe e Prova Brasil). Esse achado torna indiferente a ideia de que seria melhor utilizar testes *low-stakes* para evitar a inflação de escore no estudo de impacto de uma política de responsabilização (HOUT; ELLIOTT, 2011). Pelo menos, no caso do 5º ano, os resultados aproximados indicam que não fez muita diferença utilizar uma base originada de um teste *low-stakes* ou de um teste *high-stakes*.

RESULTADOS PARA O 9º ANO DO ENSINO FUNDAMENTAL

Passando para as análises do 9º ano do ensino fundamental, temos também dois pareamentos: o primeiro com a base do Saepe e o segundo com a base da Prova Brasil.

A Tabela 9 apresenta os resultados realizados com a base do Saepe, podendo-se observar que as diferenças das médias foram favoráveis, em torno de 5 pontos, ao grupo de intervenção em todos os intervalos de tempo analisados. Além disso, os valores do teste *t* revelaram-se bastante significativos em toda a série histórica.

Todavia, no 9º ano, os ganhos gerados no intervalo inicial (2008-2009) praticamente se mantiveram. Isso também ocorre em períodos de maior amplitude, como no de 2008 a 2012, revelando que o BDE não conseguiu, no decorrer dos anos, elevar os ganhos para o grupo de intervenção e permanecendo, praticamente, no mesmo patamar do início do programa.

TABELA 9 – Resultados da análise de diferenças em diferenças para as escolas públicas urbanas com 9º ano do EF, com base no Saepe (304 pares)

INTERVALO	CONTROLE		INTERVENÇÃO		DIFERENÇAS DAS MÉDIAS	I.C. 95%	P-VALOR
	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS			
MATEMÁTICA							
2008-2009	0,71	11,26	5,56	12,99	4,85	[2,88 ; 6,83]	< 0,0001
2008-2010	0,13	13,35	5,85	14,91	5,73	[3,53 ; 7,92]	< 0,0001
2008-2011	5,52	14,92	10,51	13,74	4,99	[2,81 ; 7,17]	< 0,0001
2008-2012	6,97	14,64	11,26	14,52	4,29	[1,95 ; 6,63]	< 0,0001
LÍNGUA PORTUGUESA							
2008-2009	3,27	11,04	7,95	11,28	4,68	[2,88 ; 6,48]	< 0,0001
2008-2010	4,40	12,25	10,14	12,20	5,74	[3,82 ; 7,67]	< 0,0001
2008-2011	5,81	14,82	11,14	12,80	5,32	[3,18 ; 7,47]	< 0,0001
2008-2012	10,80	14,73	15,97	13,51	5,17	[2,94 ; 7,40]	< 0,0001

Fonte: Elaboração dos autores a partir dos microdados do Saepe (PERNAMBUCO, 2012).

* Valor de *p* significativo até 0,05.

Também foi realizado o pareamento do 9º ano com as bases da Prova Brasil. De maneira análoga às análises do 5º ano, tal procedimento tem, como já foi dito, a função de comparar os resultados das análises, de modo a garantir resultados mais fidedignos e evitar uma possível inflação de resultados no Saepe.

Dessa forma, a Tabela 10 traz os resultados do pareamento do 9º ano na base da Prova Brasil. As diferenças das médias nos períodos avaliados ficaram em torno de 5 pontos; a exceção foi Matemática que, entre 2007 e 2009, apresentou uma diferença de 4 pontos. Resumidamente, os resultados encontrados, nessa análise, foram muito parecidos com os do Saepe, sendo que, igualmente aqui, não fez diferença utilizar uma avaliação com desenho *low-stakes* (Prova Brasil) ou *high-stakes* (Saepe) e não ocorreu uma inflação de escores na base *high-stakes*.

O efeito inicial também foi mantido em toda a série histórica, sendo que apenas a disciplina de Matemática apresentou uma discreta variação, de 1 ponto nas diferenças das médias entre o intervalo de 2007 a 2009 (4,05) e o de 2007 a 2011 (5,19). Ou seja, aqui também ocorreu uma estabilização dos ganhos gerados pelo programa de bonificação. A hipótese é que o estímulo provocado tenha alcançado um ponto de saturação, de maneira que os indivíduos submetidos à influência dessa política, uma vez que tenham empreendido esforços após a implantação do BDE (2008), não tiveram uma motivação extra para elevação dos esforços nos anos seguintes.

TABELA 10 – Resultados da análise de diferenças em diferenças para as escolas públicas urbanas com 9º ano do EF, com base na Prova Brasil (314 pares)

INTERVALO	CONTROLE		INTERVENÇÃO		DIFERENÇAS DAS MÉDIAS	I.C. 95%	P-VALOR
	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS	MÉDIA DAS DIFERENÇAS	D.P. DAS DIFERENÇAS			
MATEMÁTICA							
2007-2009	3,01	14,03	7,07	11,80	4,05	[2,10 ; 6,00]	< 0,0001
2007-2011	4,76	14,78	9,95	13,33	5,19	[3,00 ; 7,38]	< 0,0001
LÍNGUA PORTUGUESA							
2007-2009	10,29	15,75	15,69	12,65	5,40	[3,21 ; 7,58]	< 0,0001
2007-2011	8,90	0,84	14,61	12,98	5,71	[3,68 ; 7,73]	< 0,0001

Fonte: Elaboração dos autores a partir dos microdados da Prova Brasil (BRASIL, 2011).

* Valor de p significativo até 0,05.

CONSIDERAÇÕES FINAIS

Respondendo à questão inicial deste trabalho, pode-se dizer que, com base nas análises dos pareamentos realizados tanto pela base do Saepe quanto pela Prova Brasil, os resultados foram semelhantes e positivos; além disso, corroboram na mesma direção de estudos anteriores (FERRAZ; BRUNS, 2010; BRUNS; FERRAZ, 2011) e, portanto, existem boas chances de a política de bonificação ter sido responsável pela variação positiva, ou ainda pela maior parte dos ganhos de desempenho da rede estadual no período analisado.

Os resultados positivos ocorreram de maneira mais tímida no 5º ano, com teste de significância validando apenas os intervalos com distâncias maiores de tempo (no caso do Saepe, entre 2008 e 2012, e no da Prova Brasil, entre 2007 e 2011). Isso, provavelmente, indica que a bonificação produziu um efeito menor nos resultados do 5º ano do EF.

As análises dos pareamentos do 9º ano do ensino fundamental indicaram ganhos favoráveis ao grupo de intervenção, variando em torno de 5 pontos, e com teste de significância validado em todos os intervalos analisados. Assim, por meio do pareamento, pode-se dizer que, desde o início do programa, em 2008, a bonificação demonstrou ganhos favoráveis ao grupo de intervenção.

Entretanto, esses ganhos não foram sendo incrementados no decorrer da série histórica; ao que parece, a melhoria do desempenho induzida pelo bônus, de maneira inicial, não se apresenta constante ao longo do tempo; uma vez alcançado um determinado patamar de desempenho, o desafio da escola seria manter-se nele. Dessa forma, cabe aos gestores da bonificação avaliar se o retorno de no máximo 5 pontos justificaria a continuação dessa política de bonificação no estado de Pernambuco.

Portanto, uma alternativa seria engendrar mecanismos que motivem mais os professores das escolas que não atingiram o bônus e, paralelamente, que o estado forneça suporte adequado para as escolas em maior dificuldade, seja melhorando sua infraestrutura ou, ainda, investindo na capacitação dos professores, algo já previsto no documento do Programa de Modernização da Gestão (PERNAMBUCO, 2008b).

REFERÊNCIAS

ALVES, Maria Teresa G.; SOARES, José Francisco. Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. *Opinião Pública*, Campinas, SP, v. 15, n. 1, p. 1-30, jun./2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-62762009000100001&lng=pt&nrm=iso>. Acesso em: 1 jul. 2013.

AU, Wayne. High-stakes testing and curricular control: a qualitative meta synthesis. *Educational Researcher*, Thousand Oaks, CA, v. 36, n. 5, p. 258-267, Aug. 2007. Disponível em: <<http://edr.sagepub.com/content/36/5/258>>. Acesso em: 6 jul. 2013.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *SAEB – 2005 primeiros resultados: médias de desempenho do SAEB/2005 em perspectiva comparada*. Brasília, DF: Inep, 2007. Disponível em: <http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/resultados/SAEB1995_2005.pdf>. Acesso em: 16 ago. 2013.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados Prova Brasil*. Brasília, DF: Inep, 2011. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 16 ago. 2013.

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

BROOKE, Nigel. As novas políticas de incentivo salarial para professores: uma avaliação. In: ENCONTRO DE PESQUISA EM EDUCAÇÃO DA REGIÃO SUDESTE, 10., 2011, Rio de Janeiro. *Políticas públicas, movimentos sociais: desafios à pós-graduação em educação em suas múltiplas dimensões*. Rio de Janeiro: Anped, 2011. Disponível em: <<http://www.fe.ufrj.br/anpedinha2011/ebook3.pdf>>. Acesso em: 8 ago. 2013.

BROOKE, Nigel. Políticas estaduais de responsabilização: buscando o diálogo. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Ciclo de debates: vinte e cinco anos de avaliação de sistemas educacionais no Brasil*. Florianópolis: Insular, 2013. v. 2.

BROOKE, Nigel et al. *A avaliação externa como instrumento da gestão educacional nos estados*. Belo Horizonte: Game/FAE/UFMG, 2011.

BRUNS, Barbara; EVANS, David; LUQUE, Javier. *Achieving world-class education in Brazil: the next agenda*. Directions in development: human development. Washington, D.C.: The World Bank, 2012.

BRUNS, Barbara; FERRAZ, Claudio. *Paying teachers to perform: the effects of bonus pay in Pernambuco, Brazil*. Washington, D.C.: The World Bank, December 2011.

BRUNS, Barbara; FILMER, Deon; PATRINOS, Harry Anthony. *Making schools work: new evidence on accountability reforms*. Human development perspectives. Washington D.C.: The World Bank, 2011.

CORVALÁN, Javier. Accountability educacional: rendición de cuentas más responsabilización por los procesos y resultados de la educación. In: CORVALÁN, Javier; MCMEEKIN, Robert W. (Ed.). *Accountability educacional: posibilidades y desafíos para América Latina a partir de la experiencia internacional*. Santiago: Preal/Cide, 2006.

FERRAZ, Cláudio. Sistemas educacionais baseados em desempenho, metas de qualidade e a remuneração de professores: os casos de Pernambuco e São Paulo. In: VELOSO, Fernando (Org.). *Educação básica no Brasil: construindo o país do futuro*. Rio de Janeiro: Elsevier, 2009.

FERRAZ, Claudio; BRUNS, Barbara. Incentivando a qualidade: avaliando os efeitos de um bônus por desempenho na educação. In: WORKSHOP AVALIAÇÃO DE IMPACTO-SIEF, 1., 26-30 de abril de 2010, Rio de Janeiro. *Curso...* Rio de Janeiro: Banco Mundial, 2010. Disponível em: <http://siteresources.worldbank.org/EXTTHDOFFICE/Resources/5485726-1256762343506/6518748-1273071334136/7041933-1273164045906/17.b.Pagamento_por_desempenho_Ferraz_PORT.pdf>. Acesso em: 3 jul. 2014.

FOGUEL, Miguel N. Modelos de resultados potenciais. In: MENEZES FILHO, Naércio. (Org.). *Avaliação econômica de projetos sociais*. São Paulo: Fundação Itaú Social; Dinâmica, 2012.

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. 353-377, abr./jun. 2012. Disponível em: <<http://www.cedes.unicamp.br>>. Acesso em: 12 ago. 2013.

GUISBOND, Lisa et al. A década de progresso educativo perdida sob a NCLB: que lições tirar desse fracasso político? *Educação & Sociedade*, Campinas, SP, v. 33, n. 119, p. 353-377, abr./jun. 2012. Disponível em: <<http://www.cedes.unicamp.br>>. Acesso em: 12 ago. 2013.

GUO, Shenyang Y.; FRASER, Mark W. *Propensity score analysis: statistical methods and applications*. California: Sage, 2010.

HOUT, Michael; ELLIOTT, Stuart W. *Incentives and test-based accountability in education*. Washington, D.C.: National Academy, 2011.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Programa das Nações Unidas para o Desenvolvimento. *O que é IDHM?* 2016. Disponível em: <http://www.pnud.org.br/IDH/IDHM.aspx?indiceAccordion=0&li=li_IDHM>. Acesso em: 20 jul. 2014.

OSHIRO, Cláudia Hiromi; SCORZAFAVE, Luiz Guilherme. Efeito do pagamento de bônus aos professores sobre a proficiência escolar no Estado de São Paulo. In: ENCONTRO NACIONAL DE ECONOMIA, 39., 2011, Foz do Iguaçu. *Anais...* Foz do Iguaçu: Anpec, 2011.

PERNAMBUCO. Lei n. 13.486, de 01 de julho de 2008. Institui o Bônus de Desempenho Educacional – BDE, no âmbito do Estado de Pernambuco, e dá outras providências. *Diário Oficial*, Recife, ano LXXXIV, n. 1717, 09 de setembro de 2008a.

PERNAMBUCO. Secretaria de Educação. *Programa de Modernização da Gestão Pública*. 2008b. Disponível em: <<http://www.educacao.pe.gov.br/portal/?pag=1&men=69>>. Acesso em: 11 maio 2011.

PERNAMBUCO. Secretaria de Educação. *Boletim Contextual de Avaliação da Educação – SAEPE 2008: fatores contextuais associados ao desempenho escolar*. Juiz de Fora: UFJF/CAEd, 2008c. Disponível em: <http://www.saepe.caedufjf.net/wp-content/uploads/2016/05/BOLETIM_CONTEXTUAL_SAEPE_2008.pdf>. Acesso em: 19 ago. 2013.

PERNAMBUCO. Secretaria de Educação. *Microdados SAEPE*. 2012.

PERNAMBUCO. Assessoria de Comunicação do Governo do Estado. *Profissionais da educação recebem bônus por desempenho em Pernambuco*. Recife: SEE, 2013. Disponível em: <<http://www.educacao.pe.gov.br/portal/?pag=1&cat=37&art=1351>>. Acesso em: 4 ago. 2014.

PINTO, Cristine C. X. Pareamento. In: MENEZES FILHO, Naércio (Org.). *Avaliação econômica de projetos sociais*. São Paulo: Fundação Itaú Social; Dinâmica, 2012.

PONTUAL, Teresa C. *Remuneração por mérito, desafio para a educação*. São Paulo: Fundação Lemann, 2008. Disponível em: <[http://www.fundacaolemann.org.br/arquivos/uploads/arquivos/Remunera%C3%A7%C3%A3o%20por%20m%C3%A9rito%20\(paper\)%20\(Tereza%20Pontual\).pdf](http://www.fundacaolemann.org.br/arquivos/uploads/arquivos/Remunera%C3%A7%C3%A3o%20por%20m%C3%A9rito%20(paper)%20(Tereza%20Pontual).pdf)>. Acesso em: 19 ago. 2013.

RAVITCH, Diane. *Vida e morte do grande sistema escolar americano: como os testes padronizados e o modelo de mercado ameaçam a educação*. Porto Alegre: Sulina, 2011.

ROSENBAUM, Paul R.; RUBIN, Donald B. The central role of the propensity score in observational studies for causal. *Biometrika*, Oxford, v. 70, n. 1, p. 41-55, Apr. 1983.

SEGATTO, Catarina Ianni. *Como ideias se transformam em reformas: um estudo comparativo das mudanças educacionais orientadas pelo desempenho nos estados brasileiros*. 2011. 134 f. Dissertação (Mestrado em Administração Pública e Governo) – Escola de Administração de Empresas, Fundação Getúlio Vargas, São Paulo, 2011.

SOARES, José F. O efeito da escola no desempenho cognitivo de seus alunos. *REICE – Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, Madrid, v. 2, n. 2, p. 83-104, jul./dic. 2004. Disponível em: <<http://www.ice.deusto.es/rinace/reice/vol2n2/Soares.pdf>>. Acesso em: 5 ago. 2015.

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

SOARES, Tufi M. Utilização da teoria da resposta ao item na produção de indicadores socioeconômicos. *Pesquisa Operacional*, Rio de Janeiro, v. 25, n. 1, p. 83-112, jan./abr. 2005.

SOARES, Tufi M. 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.

THE HUNT INSTITUTE. Cumprindo a promessa da reforma baseada em padrões. In: BROOKE, Nigel (Org.). *Marcos históricos na reforma da educação*. Belo Horizonte: Fino Traço, 2012.

UNITED STATES OF AMERICA. Department of Education. Office of elementary and secondary education. *No Child Left Behind: a desktop reference*. Washington, D.C.: Department of Education, 2002.

CLAYTON SIRILO DO VALLE FURTADO

Analista de Medidas Educacionais no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF), Juiz de Fora, Minas Gerais, Brasil
clayton@caed.ufff.br

TUFI MACHADO SOARES

Professor do Departamento de Estatística e do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora (UFJF). Coordenador da Unidade de Pesquisa do Centro de Políticas Públicas e Avaliação da Educação (CAEd) da UFJF, Juiz de Fora, Minas Gerais, Brasil
tufi@caed.ufff.br

ANEXO I - CÁLCULO DO ISE

O cálculo do Índice Socioeconômico dos alunos foi realizado no *software* Parscale utilizando a metodologia da TRI (Teoria da Resposta ao Item); existe literatura específica de suporte para cálculos semelhantes em Soares (2005). No caso de Pernambuco, a estimação dos valores do ISE tomou as variáveis dos questionários dos alunos, a maior parte de posse de bens e acesso aos serviços básicos, mas também foram incluídas questões acerca da escolaridade dos pais, compreendendo um total de 15 questões descritas no quadro a seguir.

QUADRO 1 - Relação dos itens do questionário dos alunos do Saepe utilizados no cálculo do ISE

QUESTÃO	OPÇÕES
06. Até que série sua mãe ou a responsável por você estudou?	1 - Nunca estudou.
	2 - 1ª a 4ª série do ensino fundamental (1º Grau)
	3 - 5ª a 8ª série do ensino fundamental (1º Grau)
	4 - 1ª a 3ª série do ensino médio (2º Grau)
	5 - Até a faculdade (ensino superior)
07. Até que série seu pai ou o responsável por você estudou?	1 - Nunca estudou
	2 - 1ª a 4ª série do ensino fundamental (1º Grau)
	3 - 5ª a 8ª série do ensino fundamental (1º Grau)
	4 - 1ª a 3ª série do ensino médio (2º Grau)
	5 - Até a faculdade (ensino superior)
08. A rua em que você mora tem calçamento?	1 - Sim; 2- Não
10. Sua casa tem energia elétrica?	1 - Sim; 2- Não
11. Sua casa possui água na torneira?	1 - Sim; 2- Não
12. Sua família recebe bolsa família?	1 - Sim; 2- Não
13. Banheiro?	1 - Nenhum; 2- Um; 3- Dois ou mais
14. Rádio?	1 - Nenhum; 2- Um; 3- Dois ou mais
15. Geladeira?	1 - Nenhum; 2- Um; 3- Dois ou mais
16. Televisão em cores?	1 - Nenhum; 2- Um; 3- Dois ou mais
17. Máquina de lavar roupa?	1 - Nenhum; 2- Um; 3- Dois ou mais
18. DVD?	1 - Nenhum; 2- Um; 3- Dois ou mais
19. Automóvel?	1 - Nenhum; 2- Um; 3- Dois ou mais
20. Além dos livros didáticos, quantos livros há em sua casa? (Não conte jornais, revistas ou gibis)	1 - 1 a 20 livros
	2 - 21 a 100 livros
	3 - Mais de 100 livros
	4 - Nenhum
21. Em sua casa tem computador?	1 - Sim, com acesso à Internet
	2 - Sim, sem acesso à Internet
	3 - Não

Fonte: Elaboração dos autores a partir do Boletim Contextual do Saepe (PERNAMBUCO, 2008c, p. 26).

Recebido em: 08 JUNHO 2016

Aprovado para publicação em: 27 OUTUBRO 2016

A (DES)IGUALDADE DE CONHECIMENTO NO CICLO DE ALFABETIZAÇÃO*

NAIRA DA COSTA MUylaert LIMA

ALICIA MARIA CATALANO DE BONAMINO

LUIS ANTONIO FAJARDO PONTES

RESUMO

Este trabalho busca contribuir para o debate acerca do tema das desigualdades educacionais, por meio da investigação da (re)produção da desigualdade de conhecimento. Com base nos dados da avaliação longitudinal da alfabetização do Espírito Santo, das edições de 2012, 2013 e 2014, o estudo analisou a variação do percentual de acerto entre turmas de alto e baixo desempenho nos testes de Leitura do 1º e 3º ano do ensino fundamental. O principal resultado aponta que a aquisição das habilidades é mais heterogênea no 1º ano, o que permite supor que, ao longo do ciclo de alfabetização, os alunos podem ter passado por processos de reorganização intraescolar que tornaram seus resultados mais homogêneos, ou seja, menos desiguais. A despeito disso, há determinadas habilidades que não estão consolidadas até o fim do 3º ano, o que evidencia a ineficácia das políticas atuais para fazer com que os alunos adquiram habilidades imprescindíveis ao processo de alfabetização.

* Os autores agradecem o financiamento da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), Observatório da Educação, Edital n. 038/2010/Capes/Inep.

PALAVRAS-CHAVE ALFABETIZAÇÃO • AVALIAÇÃO DA EDUCAÇÃO • CONHECIMENTOS • DESIGUALDADES EDUCACIONAIS.

LA (DES)IGUALDAD DE CONOCIMIENTO EN EL CICLO DE ALFABETIZACIÓN

RESUMEN

Este trabajo intenta contribuir con el debate acerca del tema de las desigualdades educacionales, por medio de la investigación de la (re)producción de la desigualdad de conocimiento. En base a los datos de la evaluación longitudinal de la alfabetización del estado de Espírito Santo en las ediciones de 2012, 2013 y 2014, el estudio analizó la variación del porcentual de acierto entre grupos de alto y bajo desempeño en las pruebas de Lectura del 1er y 3er año de la educación básica. El principal resultado señala que la adquisición de las habilidades es más heterogénea en el 1er año, lo que permite suponer que, a lo largo del ciclo de alfabetización, los alumnos pueden haber pasado por procesos de reorganización intraescolar que hicieron que sus resultados fueran más homogéneos, es decir, menos desiguales. A pesar de ello, hay determinadas habilidades que no se consolidan hasta fines del 3er año, lo que pone de manifiesto la ineficacia de las actuales políticas para que los alumnos aprendan habilidades imprescindibles al proceso de alfabetización.

PALABRAS CLAVE ALFABETIZACIÓN • EVALUACIÓN DE LA EDUCACIÓN • CONOCIMIENTOS • DESIGUALDADES EDUCACIONALES.

THE (IN)EQUALITY OF KNOWLEDGE IN THE LITERACY CYCLE

ABSTRACT

The present study seeks to contribute to the debate about educational inequalities, using the study of the (re)production of the inequality of knowledge. Based on data from the longitudinal evaluation of literacy in Espírito Santo, from the 2012, 2013 and 2014 publications, this study has analyzed the percentage of variation between the high and low performance groups in the Reading Tests of the 1st and 3rd years of high school. The main results point out that the acquisition of skills is more heterogenous in the 1st year, which supports the hypothesis that, during the literacy cycle, students may have gone through intrascholastic reorganization processes which make their results more homogeneous. That is, less unequal. Despite this, there are certain skills that are not consolidated by the end of the 3rd year, which provides evidence of the ineffectiveness of the current policies for students to learn essential skills throughout the literacy process.

KEYWORDS LITERACY • EDUCATIONAL EVALUATION • KNOWLEDGE • EDUCATIONAL INEQUALITIES.

INTRODUÇÃO

A discussão da justiça na educação pode ser abordada à luz de três perspectivas: igualdade de oportunidades, de tratamento e de conhecimento (CRAHAY, 2002). Há tempo, as duas primeiras perspectivas são debatidas no campo da Sociologia da Educação, que relaciona desigualdades educacionais e desigualdades sociais. Mais recentemente, o conceito de igualdade de conhecimento também passou a ser incorporado às discussões acerca das desigualdades.

Crahay (2002), ao refletir sobre a justiça no âmbito da educação básica, mostra os limites das duas noções mais difundidas no campo educacional: a igualdade de oportunidade e a igualdade de tratamento. A primeira se sustenta na teoria dos dons naturais e, portanto, na garantia de condições iguais de desenvolvimento das pessoas, independentemente do nascimento, da cor e do sexo. A segunda noção, ao propor tratar os desiguais como iguais, estaria contribuindo para a reprodução das desigualdades sociais.

Para o autor (CRAHAY, 2002), embora essas noções apontem para questões pertinentes, elas são insuficientes para

garantir a materialização de uma escola justa, ou seja, aquilo que, no Brasil, vem sendo legalmente definido como o direito público subjetivo a uma educação básica de qualidade, conforme a Constituição Federal de 1988 (Art. 208, §1º).

Além disso, Crahay (2002) introduz uma nova noção, a de igualdade de conhecimentos adquiridos, como uma possibilidade para fazer avançar não apenas a discussão das relações entre desigualdades sociais e educacionais, mas também a busca pela plena igualdade de conhecimentos de base, de forma a garantir uma escola eficaz e justa.

A principal característica da igualdade de conhecimento diz respeito às formas de organização escolar que devem ter como foco o aprendizado dos conhecimentos de base por parte de todos os estudantes (CRAHAY, 2002). Nessa perspectiva, o ensino deve ser organizado e planejado em função de objetivos a serem atingidos por todos. A compreensão de que o conhecimento é um direito de todos e que cada indivíduo possui ritmos e formas diferentes de aprendizagem é o principal aspecto da *Pedagogia do Domínio dos Conhecimentos* (PDC), proposta por Crahay (2002). Dela decorre a necessidade de se dedicar um tempo suplementar de aprendizagem aos alunos que ainda não conseguiram consolidar os conhecimentos de base, ou seja, aqueles considerados universais e que não podem deixar de ser ensinados. Segundo Ribeiro (2014, p. 1.106), para Crahay (2002), políticas educacionais e práticas escolares baseadas na igualdade dos conhecimentos adquiridos fariam jus ao princípio de justiça corretiva no âmbito da escola e incluiriam

[...] iniciativas de discriminação positiva, tais como reforço escolar para crianças com dificuldade de aprendizagem, aulas com diversidade de práticas pedagógicas para crianças com diferentes níveis de aprendizagem, inexistência de reprovação associada a baixos desempenhos e baixa desigualdade escolar e professores mais experientes trabalhando com crianças com maiores dificuldades naquele momento. Salas heterogêneas (constituídas por alunos com maiores facilidades e dificuldades no momento), estratégias de ensino coletivas e não pautadas pela individualização da aprendizagem e foco nos objetivos de

aprendizagem também se associam a essa perspectiva, tendo em vista evidências de correlação entre elas e maiores níveis de justiça na escola.

Assim, para ser justa, a escola precisa lançar mão de ações pedagógicas que garantam o aprendizado e a máxima igualdade de conhecimento entre os alunos, de forma a tornar a variação dos resultados escolares a menor possível. Para isso, a escola deveria nortear suas ações ancorando-se nos princípios da igualdade de conhecimento (CRAHAY, 2002).

O presente estudo investiga como se apresenta e como evolui a igualdade de conhecimentos adquiridos nos primeiros anos escolares, relativos ao ciclo de alfabetização. Para tanto, faz uso de dados das edições de 2012, 2013 e 2014¹ do Programa de Avaliação da Educação Básica do Espírito Santo voltado para a alfabetização – Paebes-Alfa – e analisa o percentual médio de acerto dos itens difíceis de Leitura e sua variação nos testes aplicados aos alunos do 1º e 3º ano do ensino fundamental (EF), em turmas de alto e baixo desempenho.

Implementado em 2008, o Paebes-Alfa aplica testes de Leitura, Escrita e Matemática aos alunos dos três primeiros anos do ensino fundamental. Trata-se de uma avaliação com desenho longitudinal, que permite verificar a evolução do aprendizado dos mesmos alunos ao longo dos três anos do ciclo de alfabetização.

A avaliação dos anos iniciais do ensino fundamental ganhou destaque nas avaliações nacionais em larga escala, tanto no âmbito federal quanto no estadual, em consequência do nível de aprendizado insatisfatório da maioria dos alunos do 5º ano do ensino fundamental, que está relacionado à aprendizagem insuficiente dos alunos nos anos anteriores a essa etapa escolar.

Assim, a partir da compreensão de que os resultados insatisfatórios dos alunos são decorrentes de uma defasagem acumulada de conhecimentos e habilidades desde o início da alfabetização, surgiram não apenas as avaliações em larga escala para os anos iniciais do ensino fundamental, mas também outras iniciativas como a ampliação da duração desse nível de escolarização para nove anos. Esta implicou o

¹ A análise do percentual dos itens não levou em consideração os dados do 2º ano. Esses dados foram considerados apenas para selecionar os alunos, uma vez que o estudo observou somente o percentual de acerto dos itens dos alunos participantes de todas as ondas da avaliação do Paebes-Alfa.

ingresso dos alunos na escola aos seis anos de idade e a ênfase na alfabetização ao longo dos três primeiros anos escolares (Lei n. 11.274/2006).

Além disso, foi criado o Pacto Nacional pela Alfabetização na Idade Certa (Pnaic), um compromisso formal assumido pelos municípios, estados, distrito federal e governo federal para a garantia da alfabetização de todas as crianças até os oito anos de idade, quando estas deveriam concluir o 3º ano do ensino fundamental.

ANÁLISE DOS DADOS

Os estudos estatísticos sobre a desigualdade educacional no Brasil, segundo Alves e Franco (2008, p. 493), têm sido realizados com o objetivo de investigar as diferenças de desempenho dos alunos do ponto de vista interno às escolas. Por exemplo, Franco *et al.* (2002, p. 21), analisando dados do Sistema Brasileiro de Avaliação da Educação Básica (Saeb), verificaram que, mesmo após a introdução do controle dos resultados educacionais pelo índice socioeconômico (ISE) médio das escolas, ainda persiste uma associação significativa entre o índice socioeconômico individual dos alunos e seus respectivos resultados de proficiência, muito embora tal associação também varie entre as escolas. Esse último ponto é relevante, porque permite constatar que as escolas em geral podem variar não apenas quanto a seus resultados médios de proficiência, mas também quanto à maior ou menor equanimidade com que ensinam seus alunos.

Nessa mesma linha, Soares e Andrade (2006, p. 126), em um estudo sobre escolas na capital de Minas Gerais, atestaram que “o sistema de educação básica de Belo Horizonte só consegue produzir qualidade na presença de alta iniquidade”. Isso porque as escolas com os maiores resultados médios nas avaliações utilizadas no estudo foram também aquelas que, em geral, apresentaram as maiores variações internas de resultados. As variações de desempenho, além de estarem associadas a diferenças socioeconômicas dos alunos, também podem estar relacionadas a desigualdades raciais, conforme demonstraram Soares e Alves (2003, p. 159) que,

em estudo com dados do Saeb, verificaram a existência de um grande hiato entre alunos brancos e negros e, em menor grau, entre alunos brancos e pardos em relação ao desempenho escolar. Constataram, ainda, que os fatores produtores de eficácia do ensino não se distribuem de forma equânime, favorecendo principalmente o desempenho escolar dos alunos brancos e contribuindo para acirrar, e não para reduzir, a diferença entre os grupos raciais.

No presente estudo, é possível observar alguns pontos de interseção com as referências supracitadas. Pretende-se analisar eventuais desigualdades de conhecimento dos alunos dentro das escolas, com base nos resultados das avaliações em larga escala, levando em conta a influência de determinadas características dos alunos, tais como condição socioeconômica, cor/raça e sexo.

Tendo em vista o propósito de investigar a desigualdade de conhecimento num contexto escolar público brasileiro, optou-se por trabalhar com os resultados dos testes de Leitura de uma avaliação em larga escala – o Paebes-Alfa –, adotando-se, para isso, uma abordagem estatística que permite verificar se o conhecimento adquirido pelos estudantes é satisfatório e se sua distribuição é mais ou menos igualitária. O Quadro 1 apresenta o número total de estudantes matriculados no sistema educacional do Espírito Santo em cada um dos três primeiros anos do ensino fundamental, seguido do número total de alunos que realizaram os testes cognitivos e do número de alunos considerados nesta análise.

QUADRO 1 – Número de alunos matriculados no ensino fundamental no estado do Espírito Santo, número de participantes do Paebes-Alfa e dos alunos considerados para a análise dos dados

ANO DA AVALIAÇÃO	ANO DO EF	NÚMERO DE ALUNOS		
		MATRICULADOS*	PARTICIPANTES PAEBES-ALFA**	CONSIDERADOS PARA A ANÁLISE**
2012	1º	54.271	34.778	7.781
2013	2º	55.872	36.316	
2014	3º	63.068	39.277	

Fonte: (*) *site* QEDU. Endereço eletrônico: <<http://www.qedu.org.br/>>;
 (**) Dados do Paebes-Alfa.

Por tratar-se de uma pesquisa que utiliza dados de uma avaliação longitudinal com o objetivo de investigar a desigualdade de conhecimento, alguns critérios de seleção dos alunos e das escolas foram estabelecidos. O primeiro critério diz respeito à seleção de estudantes das escolas públicas, excluindo-se, assim, aqueles pertencentes às escolas particulares participantes. O segundo critério considerou escolas com duas turmas ou mais, excluindo-se, portanto, escolas com menos de duas turmas. Por fim, para preservar a natureza longitudinal dos dados, foram considerados os estudantes que possuíam informações de proficiência nas quatro ondas de avaliação, a saber: avaliação de entrada no 1º ano, em 2012; avaliação de saída no 1º ano, em 2012; avaliação de saída no 2º ano, em 2013; e avaliação de saída no 3º ano, em 2014. Assim, após a filtragem dos casos de acordo com os critérios adotados, o número total considerado para a análise foi de 7.781 estudantes, distribuídos em turmas de alto e baixo desempenho, de escolas públicas urbanas e rurais, seguindo procedimentos que serão apresentados mais à frente.

O estudo também optou por trabalhar com uma modelagem linear hierárquica, pois os resultados educacionais, em regra, revestem-se de uma natureza pela qual os alunos se agrupam em turmas, e estas em escolas, de modo que turmas e escolas podem variar significativamente entre si quanto a diversas variáveis de interesse, como o índice socioeconômico médio de seus integrantes, variáveis específicas dos professores e as atitudes médias dos alunos em relação ao ensino ou à aprendizagem (RAUDENBUSH; BRYK, 2002).

No presente caso, analisaram-se os percentuais de acerto dos itens nos testes, bem como sua variação. Dessa maneira, se o item tiver um alto percentual de acerto, isso indica que a habilidade de Leitura avaliada pelo item foi adquirida e consolidada pela maioria dos alunos. Por outro lado, se o item tiver um baixo percentual de acerto, isso indica que poucos alunos adquiriram a habilidade avaliada. Se a variação desse percentual entre as turmas for alta, significa que, no interior da escola, a aquisição da habilidade foi desigual. Se a variação for pequena, o aprendizado da habilidade na escola foi mais homogêneo ou mais igualitário.

Quando se analisam os percentuais de acerto dos itens de uma avaliação educacional, um problema que se coloca é a instabilidade das médias, uma vez que elas dependem diretamente da proficiência dos alunos (HOGAN, 2006). Dessa forma, há uma ampla gama de situações possíveis, estendendo-se entre dois limites: um superior, no qual os itens apresentam médias de acerto próximas a 100%, revelando a existência de um estado de consolidação da aprendizagem em aproximadamente todos os estudantes avaliados, e um inferior, no qual os itens possuem médias de acerto nulas ou quase nulas (especialmente quando corrigidos para o acerto casual). Nessa última situação, os resultados sugerem que a habilidade mensurada no item pode ter sido tratada pelo professor de forma inadequada ou, ainda, pode nem ter sido trabalhada com os estudantes em sala de aula.

No presente estudo, as médias de desempenho em cada item foram estimadas a partir dos dados das respostas dos alunos, variáveis que tinham uma codificação binária, onde o valor nulo corresponde ao erro e o valor unitário, ao acerto. Além disso, e como já observado, levou-se em conta, na análise, a natureza hierárquica dos dados, visto que, nesse caso, desejava-se estudar o desempenho dos alunos agrupados em turmas e escolas. Outro ponto de interesse foi incorporar, na estimação dos percentuais de acerto, os efeitos de algumas variáveis do nível do aluno que, conforme amiúde demonstrado na literatura, encontram-se associadas de forma significativa ao desempenho escolar, tais como o nível ou índice socioeconômico do aluno ou de sua família (XAVIER; ALVES, 2015), sua cor/raça (SOARES; ALVES, 2003) e gênero (CARVALHO, 2003).

O Quadro 2 apresenta as estatísticas descritivas referentes aos 7.781 alunos que compuseram a presente amostra e que foram avaliados inicialmente no 1º ano do ensino fundamental na rede pública do Espírito Santo, em 2012.

QUADRO 2 - Estatísticas descritivas da amostra de alunos do 1º ano EF²

	ÍNDICE SOCIO-ECONÔMICO (ISE)	RAÇA (BRANCO)	SEXO (FEMININO)
Média	0,19	0,34	0,53
Desvio-padrão	0,93	0,47	0,50

Fonte: Dados do Paebes-Alfa (elaboração própria).

Por esse quadro, observa-se que o índice socioeconômico, na escala para ele criada, possui valores próximos de zero (0,19) e de uma unidade (0,93), respectivamente para a média e o desvio-padrão. Quanto à cor da pele e o sexo, expressos em proporção, constata-se que cerca de um terço (34%) dos alunos é branco e que pouco mais da metade (53%) é do sexo feminino.

Com base nessas variáveis, obteve-se a formulação de um modelo linear hierárquico de três níveis, no qual o primeiro deles é o do aluno, o segundo é o da turma e o terceiro, o da escola, conforme se vê nas seguintes equações:

Nível 1 (aluno):

$$\text{LOG}(AC)_{ijk} = \pi_{0jk} + \pi_{1jk} (\text{ISE}) + \pi_{2jk} (\text{BRANCO}) + \pi_{3jk} (\text{FEM}) + e_{ij}$$

onde as variáveis são:

LOG(AC): logaritmo natural da chance de acerto P que um aluno i , estudando na turma j da escola k , possui em relação a um determinado item.

ISE: índice socioeconômico do aluno,³ centralizado na grande média, do mesmo modo que as duas variáveis explicativas seguintes.

BRANCO: cor do aluno (0 = não branco, 1 = branco).

FEM: gênero do aluno (0 = masculino, 1 = feminino).

E os coeficientes são:

π_{0jk} : percentual médio do logaritmo da chance de acerto dos alunos da turma j da escola k quanto a determinado item, controlando-se pelo seu respectivo índice socioeconômico, cor e gênero.

π_{1jk} , π_{2jk} , π_{3jk} : correspondem respectivamente aos coeficientes de regressão para o índice socioeconômico, a cor e o gênero dos alunos da turma j da escola k .

e_{ij} : erro estocástico da medida do aluno.

2 A forma de obtenção do índice socioeconômico é discutida na nota 3, a seguir. Quanto à raça (branca) e ao sexo (feminino) dos alunos, trata-se de variáveis binárias, valendo respectivamente 1 para brancos ou meninas, e 0 para não brancos ou meninos. A média dessas variáveis binárias é igual a P , dada pela expressão $(\sum X)/N$, onde X corresponde ao valor individual da variável e N , ao total de casos. Por sua vez, o desvio-padrão DP é obtido pela expressão $DP = \sqrt{P*(1 - P)/N}$.

3 O índice socioeconômico foi calculado com base na Teoria da Resposta ao Item (TRI), mediante a utilização do *software* Bilog MG, segundo Mislevy e Bock (1990), e foi construído com base nos itens do questionário contextual aplicado aos pais/responsáveis dos alunos dos 1º (2012), 2º (2013) e 3º (2014) anos do ensino fundamental. Os itens que compõem o índice são: escolaridade dos pais, situação de trabalho, quantidade de livros (não escolares) em casa, frequência com que realiza leituras e acesso a bens e serviços (rua asfaltada, energia elétrica, água na torneira, coleta de lixo, benefício do Bolsa Família, empregada doméstica/faxineira, banheiro em casa, rádio, TV em cores, geladeira, máquina de lavar roupa, aparelho de DVD, automóvel (carro/moto), computador, dicionário de Língua Portuguesa/outras línguas). Por sua vez, essas medidas foram obtidas pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd) na Universidade Federal de Juiz de Fora (UFJF), em parceria com a Secretaria de Educação do Espírito Santo. Informações mais detalhadas sobre o ISE podem ser obtidas nos questionários contextuais publicados por essa Secretaria, como em Espírito Santo (2013).

Já no nível 2 (da turma), as equações para as turmas de uma escola k ficam do seguinte modo:

$$\pi_{nj^k} = \beta_{n0k} + r_{nj^k} \quad \text{onde } n \in \{0, 1, 2, 3\} \text{ e:}$$

β_{n0k} : média da turma j da escola k , correspondente ao percentual médio de acerto esperado de um determinado item.

r_{nj^k} : erro estocástico da média da turma.

E, no nível 3 (escola), tem-se:

$$\beta_{nk} = \gamma_{n0} + u_{nk} \quad \text{onde } n \in \{0, 1, 2, 3\} \text{ e:}$$

γ_{n0} : média da escola k , correspondente ao percentual médio de acerto esperado de um determinado item.

u_{nk} : erro estocástico da média da escola.

Percebe-se portanto que, no nível do aluno, há três variáveis explicativas – índice socioeconômico, cor da pele e gênero –, ao passo que, no nível da turma e da escola, o modelo é incondicional. E, em todos os três níveis, há um termo residual, ou estocástico, que indica a diferença entre o seu valor observado e o seu respectivo valor previsto. Dessa forma, valores positivos desses resíduos indicam que o sujeito em análise possui um desempenho observado superior ao previsto, e valores negativos indicam o contrário. Portanto, somando algebricamente os resíduos dos alunos no nível das turmas às quais estes pertencem (ponderando-se pelo seu número de estudantes), é possível diferenciar as turmas dentro de uma mesma escola quanto ao seu desempenho médio não explicado pelo modelo. E algo análogo é obtido quando se agregam as médias das turmas (ponderando-se pelos seus respectivos tamanhos) para o nível das suas escolas.

O Quadro 3 resume algumas estatísticas relevantes desse modelo:

QUADRO 3 - Coeficientes de fidedignidade e percentuais de variância para as chances de acerto dos itens, por nível de análise no modelo linear hierárquico adotado

	COEFICIENTE DE FIDEDIGNIDADE		% DA VARIÂNCIA		
	TURMA	ESCOLA	ALUNO	TURMA	ESCOLA
Média	0,413	0,317	37,5	36,9	25,7
Desvio-padrão	0,181	0,100	14,6	15,0	9,1

Fonte: Dados do Paebes-Alfa (elaboração própria).

É possível observar, pelo quadro, que os coeficientes de fidedignidade no nível da turma e no da escola – cujas médias se situam respectivamente na casa de 0,4 e 0,3 – são elevados o suficiente para se concluir que, em geral, as médias de desempenho, tanto das turmas dentro de uma mesma escola como das escolas entre si, variam consideravelmente, o que justifica a adoção de tal modelo de três níveis. Essa observação é também corroborada, como visto no quadro, pela análise dos percentuais de variância associados às turmas (que corresponde a mais de um terço da variância total dos resultados observados) e às escolas (correspondente a cerca de um quarto dessa variância total).

Após a utilização desse modelo, que fornece o logaritmo da chance de acerto, é possível chegar ao percentual previsto deste último, p , a partir da equação do nível do aluno já mencionada (e onde as variáveis de controle são todas nulas), da seguinte forma:

$$AC = p/(1 - p) = \text{EXP}[\pi_{0jk}]$$

Ou, resolvendo-se para a probabilidade p de acerto:

$$p = AC/(1 + AC) = \text{EXP}[\pi_{0jk}]/(1 + \text{EXP}[\pi_{0jk}])$$

Dessa maneira, foi possível estimar os percentuais de acerto de cada item do Paebes-Alfa, controlados pelas três variáveis já citadas no nível do aluno. Porém, antes de analisar esses percentuais, optou-se por levar em conta também o acerto casual dos itens. No teste de Leitura, todos os itens eram de múltipla escolha e continham quatro opções de resposta. Com essa configuração, a probabilidade de acerto é significativa, mesmo que o respondente não tenha os conhecimentos necessários para acertar o item. A fim de atenuar esse problema na interpretação das médias dos percentuais de acerto, utilizou-se um procedimento para a correção do acerto casual, dado pela seguinte fórmula:

$$Perc_corrigido = Perc_real - (100 - Perc_real)/(N - 1)$$

onde:

Perc_corrigido: percentual esperado corrigido do item para o acerto casual.

Perc_real: percentual esperado *bruto* de acerto do item.

N: número de opções da múltipla escolha para o item específico (no Paebes-Alfa, as opções de resposta de cada item de Leitura são sempre quatro).

No total, foram analisadas 349 escolas e 1.168 turmas. Nessa análise, os resultados individuais dos alunos foram agregados por suas respectivas turmas e, numa segunda etapa, também por suas escolas. Desse modo, as medidas mostradas nos gráficos 1 e 2, a seguir, correspondem às médias e aos desvios padrão dos percentuais médios e previstos de acerto na totalidade dos itens apresentados na prova, e agregados inicialmente para as diferentes turmas dentro de uma mesma escola, ponderando-se pelas respectivas quantidades de alunos avaliados.

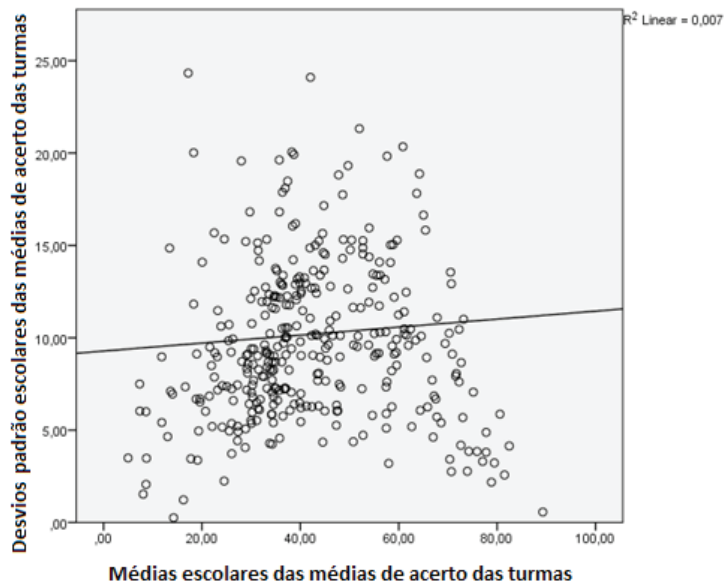
Tal abordagem permitiu analisar a associação entre as médias dos percentuais de acerto dos itens que cada escola respondeu e os seus respectivos desvios padrão, como se observa nos gráficos.

Por exemplo, conforme mostra o Gráfico 1, no 1º ano, em 2012, a associação dessas duas medidas é muito fraca, correspondendo a uma correlação de Pearson igual a 0,084.⁴ Embora esse valor seja muito próximo de zero, a inclinação da reta parece indicar uma sutil relação entre as médias e os desvios padrão do número de acertos dos itens. Ou seja, escolas cujas turmas obtêm, em média, maiores resultados tendem a ser aquelas em que se observa uma maior variabilidade entre as turmas. Já no 3º ano, em 2014, o Gráfico 2 mostra que a correlação entre essas duas medidas é de -0,53, o que implica o contrário do que se observou na (embora tênue) associação do 1º ano. Tem-se, agora, um novo padrão pelo qual, quanto maiores as médias dos percentuais de acerto nas turmas das escolas, maior também a homogeneidade de seus respectivos resultados, ou seja, menores os seus desvios padrão.

Isso possivelmente significa que, ao longo do ciclo de alfabetização, as escolas que mais agregaram valor quanto à variável de proficiência aqui utilizada (as chances de acerto em Língua Portuguesa) podem ter passado por alguns processos de reorganização que tornaram seus respectivos resultados de proficiência mais homogêneos ou equitativos. E tal associação, inexistente no 1º ano, passou a ser observada no 3º ano.

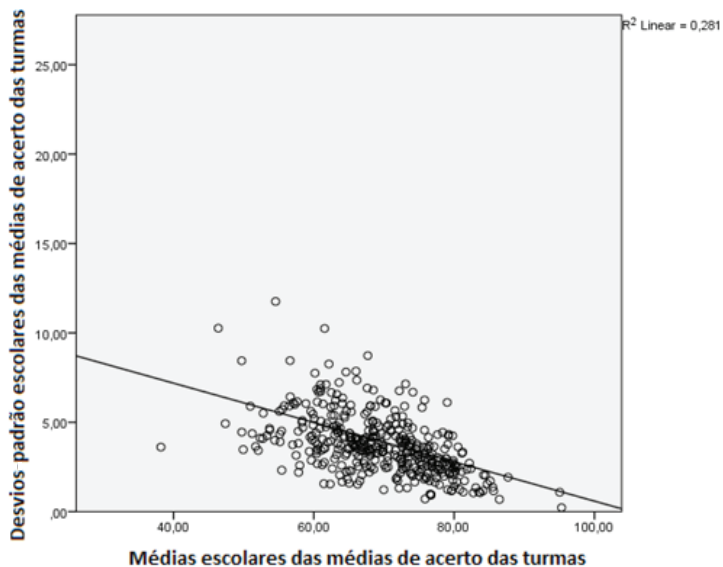
⁴ A correlação de Pearson é calculada por meio da raiz quadrada do R^2 .

GRÁFICO 1 - Associação entre a média (eixo horizontal) e o desvio-padrão (vertical) dos percentuais de acerto dos itens entre as turmas dentro de uma mesma escola, no 1º ano do ensino fundamental



Fonte: Dados do Paebes-Alfa (elaboração própria).

GRÁFICO 2 - Associação entre a média (eixo horizontal) e o desvio-padrão (vertical) dos percentuais de acerto dos itens entre as turmas dentro de uma mesma escola, no 3º ano do ensino fundamental



Fonte: Dados do Paebes-Alfa (elaboração própria).

Ao mesmo tempo, ao levar em conta especificamente nesta análise o grau de complexidade das habilidades de Leitura avaliadas no 3º ano, também se observa que as médias dos percentuais de acerto das turmas dentro das escolas tendem a ser relativamente maiores, ao passo que seus respectivos desvios padrão tendem a ser menores do que no 1º ano. Em outras palavras, poucos alunos adquirem aprendizagens complexas, e essa aquisição se dá de forma desigual no interior da escola.

A partir desses resultados mais gerais, buscou-se averiguar como as habilidades mais complexas se encontram associadas às características sociais clássicas – nível socioeconômico, cor/raça e sexo dos alunos. Com esse propósito, primeiramente se procedeu a uma classificação dos itens baseada na norma, fornecida pelo valor de seu respectivo parâmetro de dificuldade B, da TRI, expresso na escala de proficiência do teste em questão.

De acordo com Hambleton *et al.* (1991), por meio da TRI, é possível, para cada item do teste, construir uma função que relaciona a proficiência estimada de um indivíduo e a chance de ele acertar um dado item do teste, segundo uma escala logarítmica. Dessa forma, cada item possui a sua curva característica, por esse motivo chamada de Curva Característica do Item (CCI).

Na CCI, as curvas dos itens apresentam, *grosso modo*, a forma de um “S” alongado, estendendo-se lateralmente segundo duas assíntotas horizontais: uma inferior e igual ou maior que zero na extremidade esquerda do gráfico (e inferior das proficiências), e outra superior e igual a um na extremidade direita do gráfico (e superior das proficiências), indicando que, nesse último caso, a possibilidade de ocorrência de um acerto do item se aproxima de 100% dos casos. Ainda em relação às CCIs, quanto mais elas se deslocam para a direita (ou seja, na direção dos maiores níveis de proficiência), mais difíceis se consideram os seus respectivos itens. E o parâmetro B de dificuldade se relaciona, precisamente, a esse deslocamento lateral da curva no gráfico do item, associado à sua respectiva dificuldade. Também cabe mencionar que os parâmetros B, assim como várias outras medidas da

TRI, após serem obtidos originalmente numa escala padronizada, em geral, sofrem transformações lineares que não alteram a essência de sua informação, mas, ao mesmo tempo, têm a vantagem de expressá-los em escalas mais adequadas para a sua compreensão ou interpretação, em especial para um público tecnicamente não especializado na obtenção de tais medidas.

Desse modo, a escala da medida de dificuldade dos itens foi dividida em quatro partes (quartis), de forma a classificar os itens com grau de dificuldade alto (4º quartil), médio (3º e 2º quartis) e baixo (1º quartil). Os itens considerados difíceis no 1º ano apresentaram o parâmetro B superior a 529,7 na escala de proficiência, e os itens considerados difíceis no 3º ano obtiveram parâmetro B superior a 715,1.

Para os itens difíceis, tanto do 1º quanto do 3º ano, os resultados mostraram que, na maioria dos casos, não houve efeito estatisticamente significativo das variáveis sociais sobre o percentual de acerto dos itens.⁵ E, nos casos em que esses efeitos foram significativos, eles atuaram no mesmo sentido observado pelas referências já mencionadas, ou seja, associando-se positivamente ao aumento do índice socioeconômico e à ocorrência de alunos brancos e do sexo feminino.

Dessa forma, e em face de tais resultados, complementou-se a análise anterior utilizando o percentual de acerto dos itens difíceis e sua variação segundo o nível de desempenho das turmas, considerando as habilidades avaliadas por esses itens. Implementou-se o controle do efeito da composição social da escola, e basearam-se as comparações do percentual de acerto de itens difíceis nos testes de Leitura de acordo com o parâmetro B, de modo a aumentar as chances de as diferenças entre os dois grupos de turmas refletirem o currículo efetivamente ensinado e suas semelhanças e diferenças de ênfase, quer se trate de um ou de outro grupo de turmas.

Por um lado, considerou-se na análise que o alto percentual de acerto de um item mostra que a habilidade avaliada foi dominada por praticamente todos os estudantes, e que esse resultado evidenciaria a eficácia do trabalho da escola em garantir o aprendizado daquele conteúdo. Por outro

⁵ Entre os itens analisados e num nível de significância de 95% ou mais, não houve resultados significativos em 21 itens quanto ao índice socioeconômico dos alunos, em 29 itens quanto à cor da pele e em 30 itens quanto ao gênero dos alunos.

lado, foi considerado, também, que um item com baixo percentual de acerto indicaria que poucos alunos adquiriram a habilidade avaliada e que, portanto, a escola possivelmente teria trabalhado de forma inadequada aquele conteúdo. Por sua vez, em termos da categoria *igualdade de conhecimento*, que baliza este estudo, itens com alta variação do percentual de acerto entre as turmas evidenciam que a habilidade avaliada foi trabalhada de forma diferenciada entre os dois agrupamentos, inclusive entre os itens que apresentam alta variação do percentual de acerto, mas para os quais a média desse percentual é baixa para as turmas de baixo e de alto desempenho.

Além disso, o estudo identificou as diferenças na evolução geral de aprendizagem entre os dois grupos de turmas e a composição dessas diferenças em termos de conteúdo curricular de Leitura.

Nesse processo de análise, levou-se em conta a matriz de referência em Língua Portuguesa do Paebes-Alfa,⁶ que é composta por cinco tópicos. Em cada tópico, são apresentadas as competências, as habilidades, os níveis de complexidade das habilidades e os anos em que cada habilidade é avaliada. Embora a matriz do Paebes-Alfa, como toda matriz de referência, represente apenas um recorte do currículo, relacionado ao que pode ser aferido por meio de testes padronizados tipicamente utilizados na avaliação em larga escala, ela apresenta importantes elementos associados a habilidades e competências a serem desenvolvidas na etapa da alfabetização. Isso permite a análise do percentual de acerto dos itens em cada ano e, portanto, pode fornecer pistas acerca da ênfase dada a diferentes dimensões do conteúdo curricular no processo de ensino e de aprendizagem relativo ao ciclo de alfabetização.

1º ANO DO ENSINO FUNDAMENTAL

A partir deste ponto, serão apresentados os percentuais de acerto dos itens difíceis e sua variação segundo o nível de desempenho das turmas, além das habilidades avaliadas.

O procedimento adotado para observar a *igualdade de conhecimento* (CRAHAY, 2002) entre as turmas foi o de dividir

⁶ Disponível em: <<http://www.paebesalfa2onda.caedufff.net/avaliacao-educacional/matriz-de-referencia/>>. Acesso em: 10 maio 2016.

o desempenho obtido por elas em alto, médio e baixo. Para realizar essa classificação, os dados de proficiência foram divididos em tercis, de forma que o tercil mais baixo se refere às turmas de baixa proficiência, o tercil do meio, às turmas de média proficiência e o tercil mais alto às turmas de alta proficiência. Neste estudo, foram analisados os dados das turmas de alta e baixa proficiência, de modo a observar se as habilidades complexas são adquiridas de forma homogênea ou heterogênea entre os dois agrupamentos de turmas.

O Quadro 4 apresenta o percentual de acerto, a diferença do percentual de acerto entre as turmas de alto e baixo desempenho e a habilidade avaliada pelo item. As linhas do quadro estão ordenadas de forma decrescente, segundo a coluna que indica a diferença do percentual de acerto entre os dois agrupamentos.

QUADRO 4 – Percentual de acerto dos itens difíceis e sua variação segundo o nível de desempenho do grupo – 1º ano do ensino fundamental

ITEM	HABILIDADE	NÍVEL DE DESEMPENHO DO GRUPO		DIFERENÇA ENTRE OS GRUPOS
		BAIXO	ALTO	
P42	Identificar o número de sílabas de uma palavra (consciência silábica).	18,1	64,6	46,5
P72	Reconhecer a finalidade de gêneros diversos.	3,1	44,0	40,9
P66	Distinguir, como leitor, diferentes tipos de letras.	2,6	41,8	39,2
P61	Ler frases e pequenos textos, localizando informações explícitas contidas neles.	3,4	42,6	39,2
P67	Identificar o número de sílabas de uma palavra (consciência silábica).	13,8	48,7	34,9
P63	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.	23,2	57,3	34,1
P68	Ler frases e pequenos textos, localizando informações explícitas contidas neles.	0,0	33,2	33,2
P80	Identificar assunto de textos.	25,0	57,7	32,7
P64	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	8,5	40,9	32,3
P75	Identificar assunto de textos.	1,1	32,8	31,6
P34	Identificar elementos que constroem a narrativa.	30,2	60,9	30,7
P58	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	31,8	62,2	30,4
P71	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.	0,9	26,7	25,9
P57	Identificar gêneros textuais diversos.	31,4	56,3	24,9
P69	Reconhecer a finalidade de gêneros diversos.	2,1	26,3	24,2
P73	Identificar elementos que constroem a narrativa.	1,2	25,1	23,9
P62	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.	4,0	27,0	23,1
P14	Identificar sílabas e sons (consciência silábica e consciência fonêmica).	75,7	81,9	6,2

Fonte: Dados do Paebes-Alfa (elaboração própria).

Em geral, pode-se dizer que os itens avaliados no 1º ano foram difíceis para as turmas de baixo desempenho e médios para as de alto desempenho. Isso pode ser percebido nas diferenças dos valores do percentual de acerto entre esses dois grupos, que variam, na maioria dos casos, entre 47% e 25%. Somente no caso do item P14, que avalia a habilidade relacionada à consciência silábica e fonêmica, a diferença dos percentuais de acerto dos dois grupos de turmas decresce consideravelmente, para ficar na casa dos 6%.

Uma ressalva importante sobre essas diferenças é que os itens mostrados no Quadro 4 fizeram parte da avaliação dos alunos no início do 1º ano. Isso significa que os resultados apresentados ainda não conseguem refletir o trabalho das escolas e dos professores. O mais provável é que esses resultados sejam produto do *background* familiar e/ou da frequência prévia desses alunos à educação infantil. De qualquer forma, o percentual de acerto dos itens no 1º ano foi incluído nesta pesquisa por ser considerado útil para o estudo tanto da aprendizagem agregada pela escola como da evolução da variação dessas aprendizagens das turmas entre o início do 1º e o fim do 3º ano.

Como pode ser observado ainda no Quadro 4, nos dois itens que avaliam a habilidade de ler frases e pequenos textos localizando informações explícitas (P68 e P61), o percentual de acerto entre as turmas com baixo desempenho é nulo ou não ultrapassa 3,4%, revelando que os alunos não possuem a habilidade avaliada. Nas turmas de alto desempenho, os percentuais de acerto desses dois itens foram, respectivamente, 33,2% e 42,6%, o que indica que cerca de um terço ou um pouco mais dos alunos dessas turmas já apresentava essa habilidade ao iniciar o 1º ano escolar. Também pode ser observado, entre as turmas de baixo desempenho, que nos itens P75 (identificar assunto de textos), P71, P62 (reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética), P73 (identificar elementos que constroem a narrativa) e P69 (reconhecer a finalidade de gêneros diversos), o percentual de acerto mal se aproxima de 3%, indicando que os alunos dessas turmas não possuem as habilidades avaliadas. Já entre os alunos das turmas de alto desempenho, o

percentual de acerto desses mesmos itens oscila entre 25% e 32%. No entanto, é na habilidade de identificar o número de sílabas de uma palavra, avaliada pelos itens P42 e P67, que a variação do percentual de acerto entre as turmas está entre as mais altas (46,5% e 34,9%, respectivamente), confirmando a desigualdade de conhecimento presente no âmbito intraescolar no 1º ano.

Esses itens estão ancorados no padrão de desempenho avançado e a correlação bisserial item-teste⁷ é de 0,465 (item P42) e 0,446 (P67), o que indica que eles discriminam bem os alunos que dominam a habilidade daqueles que ainda não a dominam, confirmam o bom ajuste do item ao teste e informam que os alunos que mais acertaram no teste também acertaram esses itens. Ou seja, são itens que avaliam bem a habilidade que se propõem a avaliar.

Em síntese, as habilidades de reconhecer a finalidade de gêneros diversos; distinguir, como leitor, diferentes tipos de letras; ler frases e pequenos textos, localizando informações explícitas; identificar o número de sílabas de uma palavra (consciência silábica); e reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética se configuram como aquelas que apresentam maior desigualdade entre as turmas do 1º ano do ensino fundamental.

O item P64 (localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados), utilizado na avaliação de entrada no 1º ano (2012), foi o único divulgado pelo Paebes-Alfa, o que viabilizou sua interpretação pedagógica. O item está ancorado no padrão de desempenho avançado e a correlação bisserial é alta, de 0,524, indicando que esse item discrimina bem os alunos que dominam a habilidade daqueles que não a dominam e que o ajuste do item ao teste é bom. Além disso, indica que os alunos que se saíram melhor no teste também acertaram esse item. A Figura 1 apresenta o item.

7 A correlação bisserial é considerada boa a partir do valor 0,30. Itens que possuem uma correlação bisserial abaixo de 0,30 convencionalmente são aqueles que não discriminam bem os alunos que dominam a habilidade avaliada daqueles que não a dominam.

FIGURA 1 - Item P64

Questão ES.04.2014

O ELEFANTINHO

ONDE VAIS, ELEFANTINHO,
CORRENDO PELO CAMINHO
ASSIM TÃO DESCONSOLADO?
ANDAS PERDIDO, BICHINHO?
ESPETASTE O PÉ NO ESPINHO?
QUE SENTES, POBRE COITADO?
– ESTOU COM MEDO DANADO.
ENCONTREI UM PASSARINHO!

Vinicius de Moraes. A Arca de Noé. Editora José Olympio.

O ELEFANTINHO ESTAVA COM MEDO PORQUE

ANDAVA DESCONSOLADO.

ENCONTROU UM PASSARINHO.

ESPETOU O PÉ NO ESPINHO.

ESTAVA PERDIDO.

ES.04.2014
H15- Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.

Fonte: Revista Pedagógica - Paebes-Alfa (ESPÍRITO SANTO, 2012).

Conforme pode ser observado, esse item exige a identificação de uma informação explícita sobre o encontro do protagonista com outro personagem, em um poema infantil, gênero textual comumente abordado nos anos iniciais do ensino fundamental. Alunos que têm experiência na escola com a leitura de poemas reconhecem a estrutura organizacional desse gênero e fatos ficcionais que transcorrem com os personagens. Para realizar a tarefa proposta com êxito, o estudante deve inferir a causa de um evento que, no caso, é o medo do passarinho. A habilidade avaliada diz respeito à evolução dos eventos, que leva em conta a relação de causa e consequência, ou seja, qual fato dá origem a outro fato. No caso, é o medo do passarinho que faz com que o elefantinho tenha um passo apressado, um andar perdido e um ar desconsolado. O conectivo, marca linguística de uma relação de causa, não está sinalizado no poema, o que exige do aluno a habilidade de fazer inferências. Entretanto, o conectivo *porque* está presente no comando, o que torna mais clara a relação de causa e consequência.

O item apresenta baixo índice de acerto no grupo de baixo desempenho, de 8,5%, sugerindo que a grande maioria dos participantes do teste que se encontra nessas turmas não domina a habilidade avaliada. Entre as turmas com maior desempenho, esse percentual foi de 40,9%. Esse resultado indica que, provavelmente, no grupo das turmas de alto desempenho, apesar da diferença de 32,3% com relação às de baixo desempenho, a maioria dos alunos também não domina a habilidade avaliada em sua completude.

3º ANO DO ENSINO FUNDAMENTAL

O Quadro 5 apresenta os mesmos dados para o 3º ano do ensino fundamental. Nesse ano escolar, os resultados dos alunos nos testes já conseguem refletir o trabalho das escolas e dos professores ao longo do período destinado à alfabetização.

QUADRO 5 – Percentual de acerto dos itens difíceis e sua variação, segundo o nível de desempenho do grupo - 3º ano do ensino fundamental

ITEM	HABILIDADE	NÍVEL DE DESEMPENHO DO GRUPO		DIFERENÇA ENTRE OS GRUPOS
		BAIXO	ALTO	
P64	Identificar o número de sílabas de uma palavra (consciência silábica).	11,7	26,1	14,4
P72	Estabelecer relações de continuidade temática, a partir da recuperação de elementos da cadeia referencial do texto.	47,5	61,6	14,2
P51	Identificar assunto de textos.	38,3	51,0	12,7
P52	Reconhecer a finalidade de gêneros diversos.	45,6	57,9	12,3
P79	Inferir informações implícitas em textos.	33,8	45,9	12,1
P62	Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.	56,4	67,6	11,2
P70	Identificar efeitos de humor em textos diversos.	44,6	53,9	9,3
P65	Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.	9,1	18,2	9,2
P32	Identificar gêneros textuais diversos.	66,2	74,9	8,7
P48	Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.	52,2	60,8	8,6
P29	Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.	69,7	78,2	8,5
P63	Reconhecer a finalidade de gêneros diversos.	23,4	31,7	8,3
P25	Identificar assunto de textos.	72,8	80,6	7,8
P45	Estabelecer relações lógico-discursivas presentes no texto.	63,4	70,7	7,3
P57	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	56,4	63,3	6,9
P71	Identificar gêneros textuais diversos.	81,2	87,5	6,2
P78	Identificar efeitos de humor em textos diversos.	49,0	54,8	5,8
P30	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	59,5	65,1	5,6
P60	Inferir informações implícitas em textos	54,9	60,4	5,5
P66	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	76,0	81,4	5,4
P47	Estabelecer relações lógico-discursivas presentes no texto.	83,4	88,3	4,9
P77	Reconhecer a finalidade de gêneros diversos.	76,3	80,1	3,7
P56	Estabelecer relações lógico-discursivas presentes no texto.	68,5	71,9	3,4
P36	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	38,5	41,4	2,9

Fonte: Dados do Paebes-Alfa (elaboração própria).

A observação do Quadro 5 permite constatar que a variação do percentual de acerto entre as turmas de baixo e alto desempenho é bem menor que no 1º ano, oscilando entre uma diferença máxima de 14,4% e uma diferença mínima de acerto de 2,9%. Isso significa que, no início do 1º ano do ensino fundamental, havia maior desigualdade nas habilidades de Leitura entre as turmas. Outra observação importante é que os percentuais de acerto nos dois grupos de turmas aumentam em relação aos resultados do 1º ano, ultrapassando,

na maioria das vezes, o percentual de 60% e aproximando-se, em alguns deles, da casa de 80%. Por exemplo, nos itens P25 (identificar assuntos de textos), P47 (estabelecer relações lógico-discursivas presentes no texto), P66 (localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados), P71 (identificar gêneros textuais diversos) e P77 (reconhecer a finalidade de gêneros diversos), o percentual de acerto está acima de 70% para os dois grupos de desempenho, o que indica que as habilidades avaliadas por esses itens foram adquiridas pela maioria dos alunos, independentemente da turma frequentada.

É possível observar, ainda, que o item P36, que avalia a habilidade de localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados, é o que tem a menor variação do percentual de acerto (2,94%), e que esse percentual não é muito alto nas turmas de baixo e alto desempenho: 38,5% e 41,4%, respectivamente.

Já o item P64, que avalia a habilidade de identificar o número de sílabas de uma palavra, tem a maior variação do percentual de acerto, e esse valor é muito baixo para os dois grupos (11,7% nas turmas de baixo desempenho e 26,1% nas de alto). Isso revela que poucos alunos possuem essa habilidade, mesmo sendo pequena a variação entre os dois grupos. Esse item apresentou uma correlação bisserial de 0,474, indicando que ele está bem ajustado ao teste.

O item P65, utilizado na avaliação de saída no 3º ano (2014), foi o único item difícil divulgado pelo Paebes-Alfa, o que possibilita sua interpretação pedagógica. O item está ancorado no padrão de desempenho avançado e possui uma correlação bisserial alta, de 0,513.

FIGURA 2 – Item P65

(P040052C2) **Leia o texto abaixo.**

De férias e ligado na natureza

Quando a gente está de férias, não tem de acordar cedo, se preparar para as provas nem fazer lição de casa. Mas há uma tarefa que não podemos nos esquecer de jeito nenhum: cuidar bem da natureza.

NA PRAIA

- Não jogue embalagens no mar. Além de fazer a maior sujeira, elas podem matar animais.

NA ESTRADA

- Leve um saquinho para juntar o lixo. E não jogue nada pela janela do carro.
- Fale com seus pais para evitar viajar nas horas mais quentes do dia. Se o calor não estiver muito forte, peça a seus pais que desliguem o ar-condicionado e sigam viagem com as janelas abertas. O ar-condicionado faz o carro gastar mais combustível, o que causa mais poluição.

Disponível em: <http://planetasustentavel.abril.com.br/planetinha/natureza/conteudo_planetinha_479218.shtml>. Acesso em: 12 abr. 2011. Fragmento.(P040052C2_SUP)

Esse texto foi escrito para

- ambientalistas.
- estudantes.
- pais.
- professores.

Fonte: Revista Pedagógica - Paebes-Alfa (ESPÍRITO SANTO, 2012).

Esse item, conforme pode ser observado na Figura 2, relaciona-se com a habilidade de identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto. O item exige do participante inferir, a partir de pistas linguísticas (“lição de casa”, “está de férias”, “preparar para as provas”), a quem o texto se dirige uma vez que os possíveis interlocutores não são mostrados. Tal item também avalia o uso social do texto, pois há uma finalidade e público específicos, o que faz com que os alunos precisem desenvolver outras estratégias de Leitura, para além das marcas linguísticas, a fim de conseguirem responder a questão com êxito. Esse item teve um percentual de acerto de 9,1% nas turmas de baixo desempenho e de 18,2% nas turmas de alto desempenho, com variação entre esses dois grupos de 9,2%. As baixas taxas de acerto desse item em ambos os grupos permitem aventar a hipótese de que essa habilidade tenha sido trabalhada de modo insuficiente ao longo dos três primeiros anos do ensino fundamental pelas escolas investigadas nesta pesquisa.

CONCLUSÕES

De maneira geral, é possível dizer que, no início do 1º ano do ensino fundamental, o desempenho é mais heterogêneo entre as turmas observadas na presente amostra, mas essa desigualdade vai diminuindo de forma que, ao final do 3º ano, os resultados escolares são mais homogêneos. Essa tendência também é observada quando se analisam, especificamente, os itens difíceis do Paebes-Alfa. Nesse caso, quanto maior a dificuldade do item, maior a variação do acerto entre as turmas de alto e baixo desempenho nesses dois anos escolares. Apesar disso, essa variação entre os dois grupos de turmas é bem menor no 3º ano, quando comparada à do 1º ano.

Mesmo a variação do desempenho sendo menor entre as turmas no 3º ano, algumas habilidades não foram plenamente adquiridas por nenhum dos dois grupos até o final do ciclo de alfabetização, visto que o percentual de acerto nas turmas de baixo e alto desempenho é inferior a 50%. O Quadro 6 apresenta as habilidades com baixo percentual de acerto nos dois agrupamentos de turmas e nos anos escolares analisados.

QUADRO 6 – Habilidades que apresentaram baixo percentual de acerto nos grupos de baixo e alto desempenho nos anos avaliados

HABILIDADES	AValiação
Identificar o número de sílabas de uma palavra (consciência silábica).	1º e 3º ano
Reconhecer a finalidade de gêneros diversos.	
Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.	
Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.	3º ano
Inferir informações implícitas em textos.	
Distinguir, como leitor, diferentes tipos de letras.	1º ano
Ler frases e pequenos textos, localizando informações explícitas.	
Identificar assuntos de textos.	
Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.	
Identificar elementos que constroem a narrativa.	

Fonte: Dados do Paebes-Alfa (elaboração própria).

Trata-se claramente de habilidades básicas de Leitura, cujos percentuais de acerto evidenciam que praticamente a maioria dos alunos começa e termina o ciclo de alfabetização sem dominá-las por completo. Isso mostra a incapacidade de

a escola trabalhar, de forma eficaz, determinados aspectos curriculares, ao mesmo tempo que nivela por baixo o desempenho da maioria de seus alunos.

Para percorrer o objetivo de promoção de justiça e eficácia, a escola precisaria pautar suas ações nos princípios da igualdade de conhecimento, o que pressupõe que as formas de organização escolar deveriam priorizar o aprendizado dos conhecimentos de base para todos os estudantes.

Se essa perspectiva indica que as ações escolares e as práticas pedagógicas precisam ser planejadas em função de objetivos a serem atingidos por todos, as análises realizadas neste estudo com dados do Paebes-Alfa permitem depreender que os princípios da igualdade de conhecimento não se concretizam. Nesse caso, os alunos das turmas de alto e baixo desempenho não consolidam parte dos conhecimentos de base referentes à etapa da alfabetização, o que pode comprometer a aquisição de outros conhecimentos e o desenvolvimento de aprendizagens futuras da maioria dos alunos.

REFERÊNCIAS

ALVES, M. T. G.; FRANCO, C. A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. In: BROOKE, N.; SOARES, J. F. (Org.). *Pesquisa em eficácia escolar no Brasil: origem e trajetórias*. Belo Horizonte: Editora UFMG, 2008. p. 483-502.

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

BRASIL. Lei n. 11.274, de 6 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. *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 7 fev. 2006. Disponível em: <www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Lei/L11274.htm>. Acesso em: maio 2016.

CARVALHO, M. P. Sucesso e fracasso escolar: uma questão de gênero. *Educação e Pesquisa*, São Paulo, v. 29, n. 1, p. 185-193, jun. 2003.

CRAHAY, M. *Poderá a escola ser justa e eficaz?* Da igualdade de oportunidades à igualdade de conhecimentos. Tradução de Vasco Farinha. Lisboa: Instituto Piaget, 2002.

ESPÍRITO SANTO. Secretaria de Educação. *Revista Pedagógica do Paebes-Alfa 2012*, Juiz de Fora, 2012. Disponível em: <http://www.paebes.caedufff.net/wp-content/uploads/2013/06/PAEBES_Alfa_RPEDAGOGICA_LP_MAT_2EF_2012.pdf>. Acesso em: 30 nov. 2016.

ESPÍRITO SANTO. Secretaria de Educação. *Paebes – 2013*, Juiz de Fora, v. 2, jan./dez. 2013. Disponível em: <<http://www.paebes.caedufff.net/wp-content/uploads/2014/07/PAEBES-RG-RE-WEB.pdf>>. Acesso em: 30 nov. 2016.

ESPÍRITO SANTO. Secretaria de Educação. *Paebes Alfa – 2014*, Juiz de Fora, v. 1, jan./dez. 2014. Disponível em: <http://www.paebesalfa2onda.caedufff.net/wp-content/uploads/2015/04/PAEBES-ALFA-2014-RP-LP_MT-3EF_WEB.pdf>. Acesso em: 30 nov. 2016.

FRANCO, C. et al. *Qualidade e equidade na educação fundamental brasileira*, 2002. Disponível em: <<http://www.econ.puc-rio.br/pdf/td455.pdf>>. Acesso em: 10 ago. 2016.

HAMBLETON, R. et al. *Fundamentals of item response theory*. Newbury Park: Sage Publications, 1991.

HOGAN, T. *Introdução à prática de testes psicológicos*. Rio de Janeiro: LTC, 2006.

MISLEVY, D.; BOCK, R. *Bilog: Item analysis and test scoring with binary models*. Chicago: Scientific Software, 1990.

RAUDENBUSH, S. W.; BRYK, A. S. *Hierarchical Linear Models*. 2nd. ed. Thousand Oaks: Sage, 2002.

RIBEIRO, V. Que princípios de justiça para a Educação Básica? *Cadernos de Pesquisa*, São Paulo, v. 44, n. 154, p. 1.094-1.109. out./dez. 2014.

SOARES, J. F.; ALVES, M. T. G. Desigualdades raciais no sistema brasileiro de educação básica. *Educação e Pesquisa*, São Paulo, v. 29, n. 1, p. 147-165, jan./jun. 2003.

SOARES, J. F.; ANDRADE, R. Nível sócio-econômico, qualidade e equidade das escolas de Belo Horizonte. *Ensaio: Avaliação de Políticas Públicas Educacionais*, Rio de Janeiro, v. 14, n. 50, p. 107-126, jan./mar. 2006.

XAVIER, F. P.; ALVES, M. T. G. A composição social importa para os efeitos das escolas no ensino fundamental? *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 61, p. 216-243, jan./abr. 2015.

NAIRA DA COSTA MUYLEAERT LIMA

Pós-doutoranda do Departamento de Educação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Rio de Janeiro, Brasil
naira.muylaert@hotmail.com

ALICIA MARIA CATALANO DE BONAMINO

Professora do Departamento de Educação da Pontifícia
Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de
Janeiro, Rio de Janeiro, Brasil

alicia@puc-rio.br

LUIS ANTONIO FAJARDO PONTES

Analista de dados sênior do Centro de Políticas Públicas e
Avaliação da Educação da Universidade Federal de Juiz de
Fora (CAEd/UFJF), Juiz de Fora, Minas Gerais, Brasil

luis@caed.ufjf.br

Recebido em: 30 JUNHO 2016

Aprovado para publicação em: 14 DEZEMBRO 2016

AVALIAÇÃO DA EDUCAÇÃO INFANTIL NA AUSTRÁLIA: CONTRIBUIÇÕES PARA O BRASIL*

CLÁUDIA OLIVEIRA PIMENTA

RESUMO

Este artigo tem o propósito de apresentar resultados de investigação cujo objetivo foi identificar eventuais contribuições da experiência de avaliação da educação infantil australiana para a análise de iniciativas da mesma natureza, em curso no Brasil. Tem como base análise documental e informações coletadas in loco, quando da realização de estágio de pesquisa no exterior, na Graduate School of Education da Universidade de Melbourne, Austrália, em 2016. Os resultados do estudo evidenciam que o desenho avaliativo australiano reflete a preocupação com dimensões da qualidade consideradas fundamentais pela legislação e documentos norteadores da educação infantil no Brasil, ainda que os contextos social e educacional de ambos os países sejam bem diferentes. Ademais, indicam a importância da articulação e colaboração entre entes federados para a implantação de políticas voltadas para a primeira infância.

* Este artigo é fruto de estágio de pesquisa no exterior financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), Processo n. 2016/09612-7. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade da autora e não necessariamente refletem a visão da Fapesp.

PALAVRAS-CHAVE AVALIAÇÃO DA EDUCAÇÃO • EDUCAÇÃO INFANTIL • QUALIDADE DA EDUCAÇÃO • AUSTRÁLIA.

EVALUACIÓN DE LA EDUCACIÓN INFANTIL EN AUSTRALIA: CONTRIBUCIONES PARA BRASIL

RESUMEN

Este artículo tiene el propósito de presentar resultados de una investigación que tuvo el objetivo de identificar eventuales contribuciones de la experiencia de evaluación de la educación infantil australiana para analizar iniciativas de la misma naturaleza en curso en Brasil. Su base es el análisis documental e informaciones recogidas in loco, cuando se realizó la práctica de investigación en el exterior, en la Graduate School of Education de la Universidad de Melbourne, Australia, en el 2016. Los resultados del estudio evidencian que el diseño evaluativo australiano refleja la preocupación con dimensiones de la calidad consideradas como fundamentales por la legislación y documentos orientadores de la educación infantil en Brasil, aunque los contextos social y educacional de ambos países sean bastante diferentes. Además, indican la importancia de la articulación y colaboración entre entes federados para la implantación de políticas destinadas a la primera infancia.

PALABRAS CLAVE EVALUACIÓN DE LA EDUCACIÓN • EDUCACIÓN INFANTIL • CALIDAD DE LA EDUCACIÓN • AUSTRALIA.

EVALUATION OF EARLY CHILDHOOD EDUCATION IN AUSTRALIA: CONTRIBUTIONS FOR BRAZIL

ABSTRACT

This article aims to present research results intended to identify possible contributions from the Australian experience in evaluation early childhood education, in order to analyze similar initiatives existent in Brazil. It is based on documentary analysis and information collected in loco, when we conducted research internship abroad, at the Graduate School of Education of the University of Melbourne, Australia, in 2016. The results of the study show that the Australian evaluation initiative reflects the concern with dimensions of quality which are considered fundamental, by the legislation and documents guiding children's education in Brazil, even though the social and educational contexts are very different in both countries. Furthermore, they indicate the importance of articulation and federative collaboration between federal, state and municipal governments for the deployment of policies focused on early childhood.

KEYWORDS EDUCATION ASSESSMENT • EARLY CHILDHOOD EDUCATION • QUALITY OF EDUCATION • AUSTRALIA.

INTRODUÇÃO

Desde o final da última década, proposições de avaliação da educação infantil, ensejadas pelo arcabouço legal instituído a partir da Constituição Federal de 1988, vêm sendo formuladas, tanto por governos quanto por organizações da sociedade civil (SOUSA, 2014; PIMENTA, 2017).¹

Na esfera federal, o Ministério da Educação (MEC) publicou, em 2009, o material *Indicadores da qualidade na educação infantil* e, em 2012, o documento *Educação infantil: subsídios para a construção de uma sistemática de avaliação*.² Esses documentos, em conjunto com a legislação e outras publicações elaboradas ou apoiadas pelo MEC, concernentes à etapa, constituíram-se como referências para a proposição da Avaliação Nacional da Educação Infantil (Anei).³ Além das propostas do MEC, a Secretaria de Assuntos Estratégicos da Presidência da República (SAE) divulgou, em 2011, o instrumento denominado *Ages & Stages Questionnaires – Third Edition (ASQ-3)* e apoiou sua aplicação na rede municipal de educação do Rio de Janeiro.⁴

Iniciativas avaliativas da etapa também têm sido implementadas por governos municipais, seja por meio da

¹ Não trataremos, neste artigo, de proposições da sociedade civil.

² Ambos os documentos estão disponíveis no portal do MEC: <http://portal.mec.gov.br/index.php?option=com_content&view=article&id=12579%3Aeducacao-infantil&Itemid=859>. Acesso em: 30 nov. 2016.

³ A ANEI foi estabelecida pelo Ministério da Educação por meio da Portaria n. 369, de 5 de maio de 2016, que criava o Sistema Nacional de Avaliação da Educação Básica (Sinaeb), tendo sido revogada, no mesmo ano, pela Portaria n. 981, de 25 de agosto de 2016.

⁴ Disponível em: <<http://www.sae.gov.br/site/?p=6141>>. Acesso em: 20 mar. 2012.

contratação de sistemas apostilados que oferecem, entre outros elementos, instrumentos de avaliação (ADRIÃO et al., 2009; NASCIMENTO, 2012), seja adotando e/ou adaptando instrumentos produzidos em âmbito federal ou, ainda, criando procedimentos próprios de avaliação da educação infantil (BAUER et al., 2016).

As proposições gestadas no âmbito do MEC sugerem uma avaliação com foco nas condições de oferta da educação infantil. A proposta da SAE, por sua vez, promove uma abordagem que recai sobre o desenvolvimento/aprendizagem das crianças. No caso de iniciativas municipais implementadas, encontramos ambas as abordagens. Os diferentes desenhos avaliativos propostos para a etapa parecem indicar um cenário de disputas sobre quais noções de qualidade devem ser assumidas pela avaliação da educação infantil em nosso país.

Esse cenário levou-nos a investigar o funcionamento do sistema de credenciamento e avaliação da educação infantil australiana, durante estágio de pesquisa no exterior, entre os meses de agosto e outubro de 2016,⁵ com o objetivo de evidenciar potencialidades dessa experiência para contribuir com as análises de iniciativas em curso no Brasil. Em artigo anterior, tratamos de diferentes proposições de avaliação da educação infantil desenvolvidas em outros países, pois concordamos que tais experiências

[...] podem iluminar a proposição de caminhos próprios, que respondam às necessidades e aos desafios do país, marcado por diversidades e profundas desigualdades. (SOUSA; PIMENTA, 2016, p. 380)

Nossa escolha por estudar o caso australiano se pautou no fato de esse país ter implementado a avaliação da educação infantil como integrante de uma política nacional voltada para a primeira infância, assumida por todos os entes federados, por meio de articulação federativa promovida pela legislação e considerada fundamental para o desenvolvimento das políticas concernentes às crianças e, por conseguinte, para o estabelecimento da avaliação (FENECH; GIUGNI; BOWN, 2012; WESTON; TAYLER, 2016). A diversidade do contexto educativo australiano, que atende crianças

⁵ O estágio de pesquisa no exterior foi realizado na Graduate School of Education da Universidade de Melbourne, sob a orientação da Professora Doutora Jane Page.

nascidas em seu território, mas também crianças de famílias estrangeiras que vivem na Austrália, além da população aborígine, foi outro fator que contribuiu para essa escolha, dado que o Brasil também se configura como um país culturalmente diverso.

As características da educação infantil australiana são bem diferentes das do Brasil, quanto à provisão, obrigatoriedade, gratuidade, tipos de serviços oferecidos, formação de professores, entre outras. Entretanto, em ambos os países, é crescente a preocupação com a qualidade da oferta dessa etapa educacional.

No Brasil, documentos contruídos desde a década de 1990 vêm apontando padrões e critérios de qualidade que devem ser levados em conta pelos estabelecimentos educacionais e pelas diversas instâncias governamentais. Na Austrália, esse processo está institucionalizado, com a existência de padrões de qualidade nacionais utilizados para realizar o credenciamento, monitoramento e avaliação dos serviços.

O estudo aqui relatado, de natureza qualitativa, foi realizado por meio de levantamento e posterior análise de documentos governamentais e de produção acadêmica pertinentes a essa etapa educacional e à sua avaliação, bem como de entrevistas com pesquisadores da Universidade de Melbourne, diretoras de instituições de educação infantil e com um membro da Australian Children's Education and Care Quality Authority (ACECQA – Autoridade Australiana para Qualidade da Educação e Cuidados Infantis), órgão governamental responsável pela implementação das políticas voltadas para a primeira infância e pela avaliação da etapa na Austrália.

A seguir, tratamos da implementação do *National Quality Framework* (NQF – Estrutura Nacional de Qualidade) (2011b) e do *National Quality Standard* – NQS (Padrão Nacional de Qualidade) e apresentamos as características de ambos. Por fim, traçamos algumas considerações sobre possíveis contribuições dessa experiência para o Brasil.

A IMPLEMENTAÇÃO DO NATIONAL QUALITY STANDARD NA AUSTRÁLIA: HISTÓRICO E CARACTERIZAÇÃO

Até 2007, segundo Fenech, Giugni e Bown (2012), o investimento público da Austrália na primeira infância era muito pequeno, com políticas fragmentadas, formação dos professores e condições de trabalho inadequadas. Os processos de regulação de licenciamento e o sistema de credenciamento dos serviços de educação e cuidados eram considerados ineficientes e incapazes de induzir padrões e práticas de alta qualidade.

Para responder a essas críticas, bem como às recomendações da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) relacionadas à qualificação dos profissionais, proporção entre número de professores e quantidade de crianças/tamanho dos grupos – entendidos como indicadores de qualidade do processo –, a Austrália iniciou, em 2009, um conjunto de reformas para elevar os padrões e práticas de qualidade na Educação Infantil e Cuidados, em inglês, Early Childhood Education and Care (Ecec).

Segundo Weston e Tayler (2016), houve uma articulação entre todos os entes federados australianos, que concordaram em desenvolver e implementar propostas e abordagens consistentes para a Ecec, em nível nacional. Para as autoras, as “alavancas” das reformas foram a legislação, a regulamentação, o financiamento e a regulação, dimensões nas quais todos os níveis de governo estão envolvidos.

A Austrália é uma monarquia parlamentarista, formada por seis estados,⁶ dois territórios e um governo nacional,⁷ denominado Commonwealth. Os entes federados têm autonomia entre si e independência para elaborar suas legislações, não necessitando se reportar à legislação nacional. Dada a limitação de poderes da Commonwealth, estabelecida pela Constituição australiana, “cada estado mantém a sua própria soberania e o poder de fazer leis (sobre assuntos não controlados pela Commonwealth)”⁸ (WESTON; TAYLER, 2016, p. 30, tradução nossa). Um desses assuntos é a educação. Entretanto, a Commonwealth retém a maior parte das receitas tributárias e, portanto, apresenta mais capacidade de investimento que os estados (LEVY, 2011; WESTON;

6 Estados australianos: New South Wales (NSW); Victoria (Vic); Queensland (Qld); South Australia (SA); Western Australia (WA); Tasmania (Tas).

7 Territórios australianos: Northern Territory (NT) e Australian Capital Territory (ACT).

8 Trecho original: “each state retains its own sovereignty and power to make laws (over matters not controlled by Commonwealth)”.

TAYLER, 2016), o que implica induções de políticas por parte do governo central.

Weston e Tayler (2016, p. 30, tradução nossa) defendem a importância da promoção de debates sobre “os papéis e responsabilidades” de cada ente federado, bem como a colaboração por meio

[...] de fóruns e acordos intergovernamentais que buscam redistribuir o financiamento na Commonwealth com medidas de *accountability* adequadas para a obtenção de resultados.⁹

⁹ Trecho original: “intergovernmental forums and agreements that seek to re-distribute Commonwealth funding with appropriate accountability measures for the achievement of outcomes”.

Em entrevista concedida a nós em setembro de 2016, a Professora Collette Tayler (então vice-presidente da ACECQA) assinalou que a articulação entre os entes federados australianos – no âmbito da legislação – é o que demarca as diferenças entre as políticas existentes para a primeira infância antes e depois da reforma educacional da Austrália. Segundo ela, a nova legislação, constituída a partir de 2010, com a contribuição da Commonwealth, dos estados e territórios, e por eles assumida, lançou as bases legais para a tessitura das políticas educacionais e de cuidados, vigentes desde então.

No fim de 2010, a *Education and care services national law Act 2010* (Lei n. 69, de 12 de outubro de 2010) foi aprovada no Parlamento Vitoriano, atribuindo “*status legal para o acordo e para as exigências legais estabelecidas*”¹⁰ (FENECH; GIUGNI; BOWN, 2012, p. 6, tradução nossa). Tais exigências dizem respeito a:

Aprovação de provisão e serviço, certificação de supervisores, sistema de avaliação e classificação, políticas e procedimentos, saúde e segurança das crianças, ambiente físico dos centros, pessoal, informações e registros, e o estabelecimento, bem como os papéis a serem desempenhados pela ACECQA.¹¹ (FENECH; GIUGNI; BOWN, 2012, p. 6, tradução nossa)

¹⁰ Trecho original: “which gave legal status to the agreement and established legislative requirements”.

¹¹ Trecho original: “provider and service approvals, certified supervisors, the assessment and rating system, policies and procedures, children’s health and safety, a centre’s physical environment, staffing, information and records, and the establishing and roles of ACECQA”.

A referida lei enfatiza os objetivos do *National Quality Framework* (NQF) para a qualidade dos serviços oferecidos, quais sejam:

a) garantir a segurança, a saúde e o bem-estar das crianças que frequentam os serviços de educação e cuidados;

- b) melhorar os resultados educacionais e de desenvolvimento para as crianças que frequentam os serviços de educação e cuidados;
- c) promover a melhoria contínua na prestação de serviços de educação e cuidados de qualidade;
- d) estabelecer um sistema de integração nacional e responsabilidade partilhada entre jurisdições participantes e da Commonwealth na administração do quadro;
- e) melhorar o conhecimento público e o acesso à informação sobre a qualidade dos serviços de educação e cuidados;
- f) reduzir a carga regulamentar e administrativa para os serviços de educação e cuidados, permitindo que informações sejam compartilhadas entre as jurisdições participantes e a Commonwealth.¹² (VICTORIA, 2010, p. 12, tradução nossa)

O NQF compreende os seguintes elementos: o *National Quality Standard* (NQS); o desenvolvimento do quadro de aprendizagem para os primeiros anos – *Early years learning framework* (EYLF); a racionalização dos sistemas de licenciamento; e a acreditação e implementação de um sistema de classificação de qualidade. A coordenação do NQF é realizada pela ACECQA,¹³ já mencionada neste texto.

O NQS (aplicado desde 2012), em substituição ao antigo sistema de credenciamento e avaliação dos serviços de Ecec, criado em 1993, mesmo com flexibilizações relativas aos contextos locais, agora é o mesmo em toda a Austrália e integra os processos de regulação das políticas voltadas para a primeira infância.

O NATIONAL QUALITY FRAMEWORK

Em 2009, sob a liderança do Council of Australian Governments (Coag),¹⁴ foi construído, de forma colaborativa, o “primeiro quadro curricular nacional da Austrália”¹⁵ (WESTON; TAYLER, 2016, p. 33, tradução nossa), denominado *Belonging, being & becoming: the early years learning framework for Australia* (EYLF). De acordo com documento publicado pela Commonwealth of Australia (COUNCIL OF AUSTRALIAN

12 Trecho original: “a) to ensure the safety, health and wellbeing of children attending education and care services; b) to improve the educational and developmental outcomes for children attending education and care services; c) to promote continuous improvement in the provision of quality education and care services; d) to establish a system of national integration and shared responsibility between participating jurisdictions and the Commonwealth in the administration of the national education and care services quality framework; e) to improve public knowledge, and access to information, about the quality of education and care services; f) to reduce the regulatory and administrative burden for education and care services by enabling information to be shared between participating jurisdictions and the Commonwealth”.

13 A ACECQA substituiu o antigo National Childcare Quality Council (NCAC - Conselho Nacional de Credenciamento de Creches).

14 O Coag foi criado com a responsabilidade de: “aumentar a cooperação entre os governos no interesse nacional; prosseguir as reformas que visam a alcançar uma economia nacional integrada e eficiente e um mercado nacional único; continuar a reforma estrutural do governo e a revisão das relações entre os governos; considerar outras questões intergovernamentais ou de todo o governo” [Increase co-operation among governments in the national interest; Pursue reforms that aim to achieve an integrated, efficient national economy and single national market; Continue the structural reform of government and review of relationships among governments; and Consider other intergovernmental or whole-of-government issues] (GRIFFITH, 2009, p. 8, tradução nossa). O órgão também foi responsável pela articulação das políticas voltadas para a primeira infância na Austrália.

15 Trecho original: “Australia’s first national curriculum framework”.

GOVERNMENTS, 2009, p. 5, tradução nossa), o currículo está baseado em evidências internacionais que classificam a primeira infância como sendo “vital na aprendizagem e desenvolvimento das crianças”, tendo sido construído pelo “setor de educação infantil, acadêmicos da primeira infância e os governos da Austrália, estados e territórios”.¹⁶ Afirma que a estrutura curricular

16 Trecho original: “The Framework draws on conclusive international evidence that early childhood is a vital period in children’s learning and development. It has been developed with considerable input from the early childhood sector, early childhood academics and the Australian and State and Territory Governments”.

17 Trecho original: “The Framework forms the foundation for ensuring that children in all early childhood education and care settings experience quality teaching and learning”.

18 Trecho original: “It has a specific emphasis on play-based learning and recognises the importance of communication and language (including early literacy and numeracy) and social and emotional development. The Framework has been designed for use by early childhood educators working in partnership with families, children’s first and most influential educators”.

19 Trecho original: “This document may complement, supplement or replace individual State and Territory frameworks”.

20 Trecho original: “On 5 December 2008, State, Territory and Commonwealth Ministers of Education meeting as the Ministerial Council on Education, Employment, Training and Youth Affairs, released the *Melbourne declaration on educational goals for young Australians*”.

[...] constitui a base para garantir que as crianças, em todos os contextos de educação e cuidados na primeira infância, tenham uma experiência de qualidade de ensino e aprendizagem.¹⁷

E acrescenta:

Ela [a estrutura curricular] tem uma ênfase específica na aprendizagem baseada no jogo e reconhece a importância da comunicação e da linguagem (incluindo alfabetização e matemática) e do desenvolvimento social e emocional. A estrutura foi concebida para ser utilizada por educadores de primeira infância que trabalham em parceria com as famílias das crianças e educadores mais influentes.¹⁸ (COUNCIL OF AUSTRALIAN GOVERNMENTS, 2009, p. 5, tradução nossa)

Consubstanciado nas indicações previstas na Convenção das Nações Unidas sobre os Direitos da Criança, o documento estabelece que o EYLF “pode complementar, suplementar ou substituir as estruturas [curriculares] individuais do estado e território”, e que a relação deve ser “determinada por cada jurisdição”¹⁹ (COUNCIL OF AUSTRALIAN GOVERNMENTS, 2009, p. 5, tradução nossa). Informa que o quadro curricular suporta, de maneira mais ampla, a Meta 2 da *Melbourne declaration on education goals for young Australians* (Declaração de Melbourne sobre as metas de ensino para os jovens australianos), lançada em 5 de dezembro de 2008, pelos “Ministros da Educação dos estados, territórios e da Commonwealth, reunidos como Conselho Ministerial de Educação, Emprego, Formação e Juventude”²⁰ (COUNCIL OF AUSTRALIAN GOVERNMENTS, 2009, p. 5, tradução nossa).

Preocupados com os rumos econômicos mundiais e com a necessidade de tornar a Austrália um país “competitivo” e

influyente no mercado internacional, os governos locais decidiram que a educação deveria se tornar a alavanca principal no preparo das novas gerações para o desafio de enfrentar o mundo globalizado. A declaração abrange duas grandes metas, quais sejam:

- 1) a escolarização australiana deve promover a equidade e a excelência;
- 2) todos os jovens australianos devem se tornar:
 - alunos bem-sucedidos;
 - indivíduos confiantes e criativos;
 - cidadãos ativos e bem informados.²¹ (MINISTERIAL COUNCIL ON EDUCATION, EMPLOYMENT, TRAINING, AND YOUTH AFFAIRS, 2008, p. 7, tradução nossa)

21 Trecho original: "Goal 1: Australian schooling promotes equity and excellence; Goal 2: All young Australians become: - successful learners; - confident and creative individuals; - active and informed citizens".

Na perspectiva da Meta 2 da referida declaração, o currículo da educação infantil apresenta três grandes pilares: *belonging*, relacionado à noção de pertencimento; *being*, relacionado à noção de ser; e *becoming*, relacionado às transformações, ao tornar-se.

Em razão da diversidade contextual e cultural da Austrália e, portanto, de suas crianças, Weston e Tayler (2016, p. 33, tradução nossa) assinalam que o currículo se apoia "na natureza social da aprendizagem de crianças pequenas e na importância de construir relacionamentos fortes", afastando-se de "teorias de desenvolvimento e domínios de desenvolvimento".²² Segundo elas, o currículo exige "resultados abrangentes e observáveis para as crianças".

22 Trecho original: "The diverse cultural contexts of young children in Australia supported moving away from developmental theories and domains of development, to a focus on the social nature of young children's learning and the importance of building strong relationships".

Os resultados são compreendidos como "capacidades gerais" e foram incluídos no currículo

[...] com base em evidências sobre o preenchimento das lacunas de aprendizagem, cognitivas e sociais, para as crianças desfavorecidas [...] e para apoiar um 'ensino intencional' dos conceitos de alfabetização, ciências e matemática.²³ (WESTON; TAYLER, 2016, p. 34, tradução nossa)

23 Trecho original: "The outcomes were included based on evidence about closing the cognitive and social learning gaps for disadvantaged children [...] and to support 'intentional teaching' of literacy, science and mathematics concepts".

Há ênfase na importância de documentar a aprendizagem e promover práticas avaliativas que levem em conta os conhecimentos, capacidades e entendimentos das crianças, "como parte de um ciclo contínuo de planejamento,

24 Trecho original: "The EYLF also promotes assessment practices that gather and analyse information about what 'children know, can do and understand' [...] as part of an ongoing cycle of planning, documenting and evaluating children's learning".

documentação e avaliação da aprendizagem das crianças"²⁴ (WESTON; TAYLER, 2016, p. 34, tradução nossa).

O EYLF orienta os educadores a agregarem exemplos de seu contexto às suas análises. Destacamos que as informações relativas à avaliação da aprendizagem sugerem que esse procedimento deve ser realizado pelos estabelecimentos escolares, não havendo indicações de que o mesmo deva ser feito por órgãos governamentais.

O NATIONAL QUALITY STANDARD

O *Quality Improvement and Accreditation System* (QIAS) – antigo sistema de credenciamento e avaliação da educação infantil da Austrália – foi criado em 1993.²⁵ Em 2009, passou por um processo de reformulação, passando a se denominar *National Quality Standard* (NQS). Antes dessa reformulação, cada estado/território possuía sua própria legislação e podia considerar ou não as indicações do sistema.

A concepção de qualidade presente nas legislações locais e estaduais da Austrália, à época do QIAS, era a de que as instituições deveriam adotar padrões mínimos que pudessem ser facilmente mensurados. Esses padrões referiam-se a “espaço, variedade de equipamentos, número e idade das crianças, tamanho da equipe e duração de seu treinamento”²⁶ (COMMONWEALTH OF AUSTRALIA, 1993, p. III, tradução nossa).

O NQS foi desenvolvido a partir de 2009, tendo sido aplicado, pela primeira vez, em 2012. Está consubstanciado em aspectos de qualidade definidos pela OCDE, quais sejam:

[...] conceito educacional e prática, qualidade estrutural, interações entre educadores e crianças e serviços voltados para atender as necessidades das famílias e das comunidades locais.²⁷ (AUSTRALIAN CHILDREN'S EDUCATION AND CARE QUALITY AUTHORITY - ACECQA, 2011, p. 8, tradução nossa)

De acordo com o *Guide to the National Quality Standard* (ACECQA, 2011a), o NQS é composto por sete áreas de qualidade, que contêm, ao todo, 18 padrões. Para cada padrão, há uma série de aspectos a serem analisados, totalizando 52.²⁸

25 A versão QIAS do sistema de credenciamento e avaliação da educação infantil australiana foi estudada e utilizada por Piotto *et al.* (1998), de forma adaptada, para avaliar quatro creches localizadas no município paulista de Ribeirão Preto e região.

26 Trecho original: "such as space, range of equipment, number and ages of children, number of staff and length of their training".

27 Trecho original: "These aspects include educational concept and practice, structural quality, interactions between educators and children and targeting services to meet the needs of families and local communities".

28 Para saber quais são os aspectos analisados, ver Pimenta (2017).

No Quadro 1, a seguir, registramos as áreas avaliadas e os padrões concernentes a cada uma delas.

QUADRO 1 – Padrões de qualidade e áreas avaliadas na Ecec da Austrália

ÁREA	PADRÕES
Programa educacional e prática	<ul style="list-style-type: none"> - Uma estrutura de aprendizagem aprovada informa o desenvolvimento de um currículo que melhora a aprendizagem e o desenvolvimento de cada criança. - Educadores e coordenadores são focados, ativos e reflexivos na concepção e materialização do programa para cada criança.
Saúde e segurança das crianças	<ul style="list-style-type: none"> - A saúde de cada criança é promovida. - Alimentação saudável e atividade física são incorporadas ao programa para crianças. - Cada criança está protegida.
Ambiente físico	<ul style="list-style-type: none"> - O desenho e a localização das instalações são apropriados para a realização do serviço. - O ambiente é inclusivo, promove competência, exploração independente e aprendizado através da brincadeira. - O serviço tem um papel ativo no cuidado de seu ambiente e contribui para um futuro sustentável.
Arranjos de pessoal	<ul style="list-style-type: none"> - Arranjos de pessoal melhoram a aprendizagem e o desenvolvimento das crianças e garantem a sua segurança e bem-estar. - Educadores, coordenadores e membros da equipe são respeitosos e éticos.
Relações com as crianças	<ul style="list-style-type: none"> - Relações respeitosas e justas são desenvolvidas e mantidas com cada criança. - Cada criança é apoiada para construir e manter relações delicadas e sensíveis com outras crianças e adultos.
Parcerias de colaboração com as famílias e comunidades	<ul style="list-style-type: none"> - Relações de respeito e apoio são desenvolvidas e mantidas com as famílias. - As famílias são apoiadas no seu papel parental e os seus valores e crenças sobre a criação dos filhos são respeitados. - O serviço colabora com outras organizações e prestadores de serviços para melhorar a aprendizagem e o bem-estar das crianças.
Liderança e gestão de serviços	<ul style="list-style-type: none"> - A liderança eficaz promove uma cultura organizacional positiva e constrói uma comunidade profissional de aprendizagem. - Há um compromisso com a melhoria contínua. - Sistemas administrativos permitem a gestão eficaz de um serviço de qualidade.

Fonte: Elaboração da autora com base em ACECQA (2011, p. 10, tradução nossa).

Há um processo de classificação e avaliação da qualidade,

[...] que reflete uma abordagem nacional para a avaliação e comunicação de qualidade dos serviços de educação e cuidados em toda a variedade de configurações de serviço existentes.²⁹ (ACECQA, 2011, p. 12, tradução nossa)

Acredita-se que a disponibilização dessas informações tem o potencial de promover “a transparência e a prestação de contas”, contribuindo para que as famílias façam “escolhas informadas sobre a qualidade da educação e dos cuidados em um serviço”³⁰ (ACECQA, 2011, p. 12, tradução nossa).

No entanto, de acordo com a diretora de um serviço de educação infantil, entrevistada por nós em outubro de 2016, alguns serviços, especialmente nas regiões suburbanas, não

29 Trecho original: “The National Quality Standard is accompanied by a national quality rating and assessment process that reflects a national approach to the assessment and reporting of the quality of education and care services across the variety of service settings”.

30 Trecho original: “The availability of this information promotes transparency and accountability and will help parents make informed choices regarding the quality of education and care at a service”.

têm o mesmo padrão de qualidade daquele oferecido no serviço dirigido por ela – um *Long Day Care* [Centro de Cuidados ao Longo do Dia] pertencente à Universidade de Melbourne, cujas taxas semanais podem chegar a 650 dólares australianos por cinco dias de frequência da criança ao serviço. Dessa forma, ao que tudo indica, famílias mais pobres não conseguem acessar serviços mais caros e provavelmente com mais qualidade em termos de insumos, formação de professores, etc. Todavia, vale destacar que o governo australiano presta auxílio financeiro às famílias, por meio de programas em que elas podem receber um desconto no pagamento de taxas, tais como o Child Care Benefit (CCB) e o Child Care Rebate (CCR). Nesses programas, o governo subsidia as famílias para ajudá-las a pagar as taxas dos serviços de Ecec.

Além disso, a ACECQA disponibiliza, para as famílias, educadores e sociedade em geral, a classificação dos estabelecimentos educacionais nas áreas de qualidade do NQS, por meio do *site MyChild*.³¹ No Quadro 2, a seguir, registramos o sistema de classificação da avaliação realizada para aferir o atendimento dos serviços de Ecec ao NQS.

31 A avaliação de cada serviço está disponível em: <<https://www.mychild.gov.au/>>. Os resultados são públicos para qualquer cidadão.

QUADRO 2 – Níveis de classificação da avaliação dos serviços de Ecec. NQS. Austrália

NÍVEL		SIGNIFICADO
01	Excelente	Indica que um serviço demonstra excelência e é reconhecido como um líder do setor.
02	Excede o Padrão Nacional de Qualidade	Indica que um serviço é superior ao Padrão Nacional de Qualidade.
03	Atende ao Padrão Nacional de Qualidade	Indica que um serviço está atendendo ao Padrão Nacional de Qualidade.
04	Trabalha para atender ao Padrão Nacional de Qualidade	Indica que um serviço está trabalhando no sentido de cumprir o Padrão Nacional de Qualidade.
05	Necessitando de melhoria significativa	Indica que um serviço não está cumprindo o Padrão Nacional de Qualidade e que o regulador está trabalhando estreitamente com o serviço para melhorar imediatamente a sua qualidade (caso contrário, a aprovação do serviço para operar será retirada).

Fonte: Elaboração da autora com base em ACECQA (2011, p. 12, tradução nossa).

O Guia explicita, ainda, como a classificação geral dos serviços é determinada ao fim do processo de avaliação, como registramos no Quadro 3, a seguir:

QUADRO 3 – Nível e Processo de Classificação Geral do NQS, na Austrália

NÍVEL		CLASSIFICACAO GERAL
01	Excelente	A classificação excelente é dada pela ACECQA.
02	Excede o Padrão Nacional de Qualidade	<p>O serviço atende a todas as normas e recebe uma classificação "Excede o Padrão Nacional de Qualidade" em pelo menos quatro áreas de qualidade, incluindo ao menos duas das seguintes áreas:</p> <ul style="list-style-type: none"> • programa educacional e prática; • relações com as crianças; • parcerias de colaboração com as famílias e comunidades; • liderança e gestão de serviços. <p>Um serviço que oferece educação e cuidados para crianças em idade pré-escolar só pode receber uma classificação "Excede o Padrão Nacional de Qualidade" para a Área de Qualidade 1 (Programa educacional e prática) se o serviço:</p> <ul style="list-style-type: none"> • fornece um programa pré-escolar; ou • tem um acordo documentado com outro serviço aprovado para fornecer um programa de pré-escola e informa os pais sobre esse regime.
03	Atende ao Padrão Nacional de Qualidade	O serviço recebe uma classificação de que atende no mínimo ao Padrão Nacional de Qualidade, em cada área de qualidade. Ele pode receber uma classificação "Excede o Padrão Nacional de Qualidade" em uma ou mais áreas de qualidade, mas não satisfaz os requisitos para a classificação "Excede o Padrão Nacional de Qualidade".
04	Trabalhando para atender ao Padrão Nacional de Qualidade	O serviço recebe uma classificação de que está trabalhando com o Padrão Nacional de Qualidade para uma ou mais áreas de qualidade (mas não recebe qualquer classificação "Necessitando de melhoria significativa").
05	Necessitando de melhoria significativa	O serviço recebe uma classificação "Necessitando de melhoria significativa" para uma ou mais áreas de qualidade.

Fonte: Elaboração da autora com base em ACECQA (2011, p. 13, tradução nossa).

Se um serviço não atende ao Padrão Nacional de Qualidade e é classificado como "Necessitando de melhoria significativa" (questões sérias) ou "trabalhando para avançar", mas ainda não atende aos padrões, isso

[...] indica que um serviço não está cumprindo o NQS e que o regulador [ACECQA] está trabalhando em estreita colaboração com o serviço para melhorar imediatamente a sua qualidade (caso contrário, a aprovação do serviço para operar será revogada).³² (ACECQA, 2011, p. 12, tradução nossa)

De acordo com informações prestadas pela Professora Jane Page,

[...] se um serviço falhar e não puder mais atender as crianças que estão sob sua responsabilidade, o governo estadual deverá trabalhar com o governo local para colocar as crianças e as famílias em novos serviços.³³ (tradução nossa)

A Professora Collette Tayler complementou a informação:

Se um serviço é classificado como "Necessitando de melhoria significativa", ele fica sob a supervisão rigorosa da autoridade reguladora e é obrigado a resolver imediatamente

32 Trecho original: "indicates that a service is not meeting the National Quality Standard and that the regulator is working closely with the service to immediately improve its quality (otherwise the service's approval to operate will be withdrawn)".

33 Trecho original: "if a service fails and can no longer meet the children under its responsibility, the state government should work with local government to put children and families in new services".

quaisquer questões de segurança, etc. Eles [os serviços] são visitados com muita frequência e recebem ajuda para melhorar. Se, dentro do prazo alocado pela autoridade para atender às mudanças, eles não o fizerem, a autoridade poderá (por meio da ação governamental) encerrar o serviço.³⁴ (Entrevista concedida em outubro de 2016, tradução nossa)

34 Trecho original: "If a service is rated 'Significant Improvement required' they are under the close supervision of the Regulatory Authority and are required to address immediately any issues of safety etc; they are visited very frequently and they are given help to improve. If, within the timeframe allocated by the Authority to meet the changes, they fail to do so then the Authority can (through government action) close down the service".

35 Trecho original: "Few are forced to close as the conditions required, along with supports given usually work to keep the service open. The Regulatory Authorities are the places where the forced closures evidence resides".

36 Trecho original: "This varies according to each state and territory budget, workload and priorities. Mostly, the NQS Authorised Officers and local professional associations and networks provide professional development support to the centres, and ACECQA provides learning and development workshops. The level of professional development budgets is now beginning to decline, after major investment between 2009 and 2016 to bring the NQF/NQS into place".

37 Trecho original: "NQS requires them to develop their Improvement Plans specifically to address the area/s that have not been demonstrated as at the Standards".

Ainda de acordo com ela,

[...] poucos [serviços] são forçados a fechar, pois as condições exigidas, juntamente com o apoio dado [pelas autoridades reguladoras], geralmente trabalham para manter o serviço aberto³⁵ (Entrevista concedida em outubro de 2016, tradução nossa).

Taylor asseverou que, embora haja investimentos do governo para apoiar esses serviços,

Isso varia de acordo com cada orçamento de estado e território, carga de trabalho e prioridades. Os funcionários autorizados do NQS, as associações e redes locais de profissionais fornecem apoio de desenvolvimento profissional aos centros, e o ACECQA oferece *workshops* de aprendizagem e desenvolvimento. O nível dos orçamentos de desenvolvimento profissional está começando a diminuir, após grandes investimentos realizados entre 2009 e 2016, para implementar o NQF e o NQS.³⁶ (Entrevista concedida em outubro de 2016, tradução nossa)

Ao ser questionada sobre o que ocorre com um serviço quando este piora ou não avança de uma avaliação para a outra, Taylor informou que é exigido que os serviços “desenvolvam seus planos de melhoria para a área que não está atendendo aos padrões”³⁷ (Entrevista concedida em outubro de 2016, tradução nossa).

De fato, à época da publicação do guia do NQS (ACECQA, 2011), o documento indicava que os regulamentos nacionais seriam avaliados e que as autoridades reguladoras de cada estado/território deveriam “informar a revisão dos serviços (conformidade histórica, classificação e história de

avaliação)”,³⁸ bem como rever o Plano de Melhoria da Qualidade do serviço (Quality Improvement Plan) – mencionado pela Professora Collette – que existia até então (ACECQA, 2011, p. 13, tradução nossa).

A avaliação e a subsequente classificação dos serviços são realizadas pela “entidade reguladora competente em cada estado e território”³⁹ (ACECQA, 2011, p. 12, tradução nossa), que envia um ou mais funcionários autorizados para realizar o processo, a depender do tipo e tamanho do serviço a ser avaliado. Para tanto, utiliza-se o Instrumento de Avaliação e Classificação do NQS, em inglês, *National quality standard assessment and rating instrument* (ACECQA, 2012). De acordo com as informações fornecidas por Tayler,

As avaliações começaram em 2012 (com cerca de 15.000 serviços a serem avaliados). Elas [as avaliações] levam tempo e nem todos os serviços foram avaliados ainda, lembrando que novos serviços são adicionados ao sistema todos os anos. Alguns serviços (aqueles classificados como “*Trabalhando para atender ao Padrão de Qualidade*” receberam uma segunda avaliação – isso acontece um ano após o primeiro). O calendário para visitar e avaliar cada serviço está sob o controle de cada autoridade reguladora do estado e do território. A ACECQA, em uma série de exercícios de consistência e para garantir que cada estado e território estejam acompanhando o seu trabalho de avaliação, publica relatórios de avaliação sobre o progresso de cada estado e território no *National Quality Standard*, quatro vezes por ano.⁴⁰ (Entrevista concedida em outubro de 2016, tradução nossa)

Um conjunto de documentos e ferramentas acompanha o NQS para orientação, tanto dos educadores como dos serviços e autoridades que realizam a avaliação e o monitoramento da qualidade, quais sejam: os quadros curriculares; guias curriculares para os educadores; guia do Padrão Nacional de Qualidade; guia para o desenvolvimento do Plano de Melhoria da Qualidade; e o Instrumento de Avaliação e Classificação.

Com o objetivo de auxiliar

38 Trecho original: “review service information (compliance history, rating and assessment history)”.

39 Trecho original: “The relevant Regulatory Authority in each state and territory undertakes the assessment and rating process [...]”.

40 Trecho original: “Evaluations started in 2012 (with about 15.000 services to evaluate). They take time and not all services are evaluated as yet, bearing in mind that new services are added to the system every year. Some services (those rated as “Working Towards” have received a second assessment – this happens one year after the first. The timetable for visiting and assessing each service is in the control of each state and territory Regulatory Authority. Together the Regulatory Authorities work with ACECQA on a range of consistency exercises, and to ensure that each state and Territory is keeping up with their assessment work. ACECQA publicly reports on the progress of each state and territory’s NQS assessments 4 times per year”.

[...] os serviços a concluir o processo de planejamento de melhoria da qualidade, de autoavaliação e de avaliação de suas práticas atuais e identificar quais práticas podem ou devem melhorar.⁴¹ (ACECQA, 2011, p. 16, tradução nossa)

41 Trecho original: "The purpose of this guide is to assist services to complete the self-assessment and quality improvement planning process by evaluating their current practices and identifying which practices they can or should improve".

o *Guide to the NQS* (ACECQA, 2011) orienta como educadores e assessores das autoridades que realizam a avaliação devem proceder em relação a cada aspecto avaliado pelos padrões de qualidade, descrevendo cada uma das sete áreas de qualidade do NQS, incluindo:

- uma declaração introdutória para cada área de qualidade, que fornece o contexto e os fundamentos, bem como uma lista dos padrões e elementos que se enquadram na área de qualidade;
- uma lista das seções relevantes da Lei Nacional e das Regulações Nacionais que se aplicam à área de qualidade;
- uma descrição de cada padrão e uma explicação sobre como ele contribui para educação e cuidados de qualidade para as crianças;
- perguntas reflexivas para o serviço considerar quando trabalha na direção de cada padrão;
- um guia para a prática de cada elemento, o qual descreve como o elemento pode ser colocado em prática no serviço e como o elemento pode ser avaliado. Este consiste de orientação aplicável a todos os tipos de serviços e crianças de todas as idades, seguido por qualquer orientação específica identificada para o tipo de serviço ou idade das crianças.⁴² (ACECQA, 2011, p. 16, tradução nossa)

42 Trecho original: "an introductory statement for each quality area, which provides context and the rationale, as well as a list of the standards and elements that fall within the quality area; a list of the relevant sections of the National Law and National Regulations that apply to the quality area; a description of each standard and an explanation about how it contributes to quality education and care for children; reflective questions for the service to consider when working towards each standard; a guide to practice for each element, which describes how the element might be put into practice at the service and how the element may be assessed. This consists of guidance applicable to all service types and children of all ages, followed by any specific guidance identified for the service type or age of the children".

43 Disponível em: <<http://www.acecqa.gov.au/assessment-and-ratings-process>>. Acesso em: 10 dez. 2016.

O processo avaliativo tem duração de aproximadamente dois meses, entre o aviso da ACECQA para a instituição até sua finalização, com a pontuação e classificação do serviço. Cabe informar que, durante o processo, os avaliadores elaboraram um relatório preliminar com suas considerações sobre as áreas avaliadas e, dentro do período de avaliação, o serviço pode solicitar a revisão dessas considerações, desde que com argumentos fundamentados e documentados.⁴³ Para a diretora de um Jardim de Infância visitado por nós, o período de avaliação não deveria ser avisado aos serviços. Em suas palavras,

A única coisa que eu mudaria é que a classificação e a avaliação deveriam ocorrer de forma aleatória, sem preparação. Acredito que devemos visar à alta qualidade em todos os momentos e que os avaliadores devem observar a prática diária, em vez dos serviços. Dessa forma, teríamos uma verdadeira indicação de onde o serviço está alinhado com os padrões, com oportunidades para continuar a melhorar.⁴⁴ (Entrevista concedida em outubro de 2016, tradução nossa)

Ainda assim, ela considera que o NQS contribui sobremaneira para elevar a qualidade da educação infantil na Austrália:

O NQS garante a segurança, a saúde e o bem-estar de todas as crianças que frequentam o nosso serviço e, com cada elemento, apoia os professores e educadores para melhorar os resultados educacionais e de desenvolvimento para as crianças. [...] Com o Quality Improvement Plan do serviço, somos capazes de refletir continuamente sobre nossa prática e prestação de serviços em consonância com as sete principais áreas de qualidade e regulamentos, identificar problemas e estabelecer metas para trabalhar em cada área relevante para o nosso serviço.⁴⁵ (Entrevista concedida em outubro de 2016)

Até o mês de dezembro de 2016, dos 15.434 serviços de educação infantil aprovados na Austrália, 13.312 (86%) deles tinham sido avaliados. Destes, 5.537 (42%) foram classificados como *Atende ao Padrão Nacional de Qualidade*. Outros 4.045 (30%) receberam a classificação *Excede o Padrão Nacional de Qualidade (Exceeding NQS)* e 3.644 (27%) foram classificados como *Trabalhando para atender ao Padrão Nacional de Qualidade*. As classificações *Necessitando de melhoria significativa* e *Excelente* foram atribuídas para menos de 1% dos serviços: 33 e 53, respectivamente (ACECQA, 2017).

Entre 2012 e 2016, 1.705 serviços foram reavaliados, sendo que 1.053 (62%) deles obtiveram uma melhor classificação, 579 (34%) mantiveram o resultado e 73 (4%) ficaram abaixo da classificação anterior. Da primeira para a segunda

44 Trecho original: "The only thing I would change is that the rating and assessment should take place randomly without preparation, as I believe we should aim for high quality at all times and assessors should observe everyday practice rather than services have time to prepare. This way we would have a true indication of where the service is aligned with the standards with opportunities to continue to improve".

45 Trecho original: "NQF ensures the safety, health and wellbeing of all children attending our service, as well as with each element supporting teachers and educators to improve the educational and developmental outcomes for children. [...] With the Quality Improvement Plan in place we are able to continuously reflect on our practice and service delivery in line with the 7 key quality areas and regulations identifying issues and setting goals to work on each area relevant to our service".

avaliação, o número de serviços classificados como *Necessitando de melhoria significativa* caiu de 28 para nove. Já a quantidade de serviços que obtiveram classificação *Trabalhando para atender ao Padrão Nacional de Qualidade* caiu de 1.365 para 495. Em contrapartida, o número de serviços que atingiu classificação *Atende ao Padrão Nacional de Qualidade* aumentou de 236 para 806 e 395 alcançaram a classificação *Excede o Padrão Nacional de Qualidade* no processo de reavaliação, contra 76 classificados nesse nível na avaliação anterior (ACECQA, 2017).

Segundo Tayler (2014, p. 146), os resultados iniciais dessa avaliação indicaram que:

[...] o Padrão ainda se aplica a uma minoria de instituições que continuam trabalhando para atingi-lo. O programa e a prática educacional representam a Área de Qualidade que parece ser a mais difícil de equiparar ao Padrão exigido. A área requer a utilização de um currículo aprovado; um currículo baseado na brincadeira e nas ideias e interesses das crianças; experiências e rotinas inclusivas e cognitivamente desafiadoras; e uma demonstração de foco, atividade e reflexão, por parte dos educadores, em suas práticas e interações com as crianças.

Corroborando essa visão, Weston e Tayler (2016, p. 36, tradução nossa), com base em estudo de caso que analisou resultados da avaliação relativos à área *Programa educacional e prática*, revelam que “uma percentagem significativa dos serviços foi avaliada no nível *Trabalhando para ao Padrão Nacional de Qualidade*”,⁴⁶ ou seja, tais serviços ainda não atingiram os padrões estabelecidos, o que, segundo elas, levanta questões sobre o que precisa ser feito após o processo de avaliação.

A professora Tricia Eadie, da Universidade de Melbourne, reiterou, por meio de entrevista em outubro de 2016, que as avaliações têm demonstrado que os padrões relativos às práticas educativas são aqueles que necessitam de mais atenção. Para ela, os padrões concernentes à *Liderança e gestão de serviços* também merecem um olhar mais apurado.

Observe-se que tais áreas – *Programa educacional e prática e Liderança e Gestão de Serviços* – estão diretamente relacionadas

46 Trecho original: “a significant percentage of assessed services at the ‘working towards NQS’ level”.

aos recursos humanos e, portanto, aos processos que ocorrem no interior das instituições, podendo ser considerados como aspectos *menos tangíveis*, ou seja, mais difíceis de serem captados, como problematizam Logan e Sumsion (2010, p. 44, tradução nossa, grifos dos autores):

Aspectos *tangíveis* de qualidade são semelhantes aos aspectos estruturais da qualidade e incluem aspectos facilmente mensuráveis, tais como proporção entre funcionários-crianças, tamanhos de grupo, qualificações do pessoal, materiais de aprendizagem e capital físico. Aspectos *menos tangíveis* de qualidade incluem entusiasmo, liderança, motivação, interações pessoais da criança, crenças e a complexidade da tomada de decisão que contribui para julgamentos profissionais. [...] Aspectos menos tangíveis de qualidade [...] abrangem uma gama mais ampla do que aspectos do processo e são difíceis de medir por sistemas tradicionais e escalas de avaliação.⁴⁷

Ao que parece, o desafio que se coloca para a Austrália, nesse momento, é conseguir ir além da constatação realizada pelo processo de avaliação. Cabem reflexões e a construção de ações mais diretas nos estabelecimentos educacionais, de formação de pessoal e de escuta das percepções de qualidade dos profissionais que atuam na educação infantil.

Nessa perspectiva, Mentha, Church e Page (2015), ao analisarem as percepções de professores quanto à participação e atuação das crianças em seu processo de aprendizagem (lembrando que a participação das crianças é um dos aspectos avaliados pelo NQS), asseveram a importância de evidenciar os conceitos de qualidade existentes entre os professores para compreender as “complexidades” inerentes ao conceito e, assim, poder implementar “políticas mais consistentes e claras”.

Segundo Tayler (2014, p. 146), há uma expectativa, por parte dos implementadores das políticas e do NQS, de que os educadores sejam

[...] reconhecidos por suas habilidades profissionais, por sua contribuição ao desenvolvimento e à aprendizagem das crianças, e que recebam apoio para aprimorarem sua

47 Trecho original: “Tangible aspects of quality are similar to structural aspects of quality and include readily measurable aspects, such as child-staff ratios, group sizes, qualifications of staff, learning materials and physical capital. Less tangible aspects of quality include enthusiasm, leadership, motivation, child-staff interactions, beliefs, and the complexity of decision making that contributes to professional judgements. [...] Less tangible aspects [...] encompass a broader range than process aspects and are difficult to measure by traditional systems and rating scales”.

especialização profissional, mantendo-se em consonância com as intenções do NQS.

A autora defende que

[...] o fato de todas as unidades aprovadas estarem sendo avaliadas com base no NQS tem sido considerado positivo para crianças, famílias, educadores, prestadores de serviços e governos. (TAYLER, 2014, p. 146)

Ela assinala que a expectativa é a de que esse sistema de avaliação “estimule melhorias na qualidade dos programas e das experiências,” beneficiando “o desenvolvimento e a aprendizagem das crianças”, bem como informando melhor as famílias sobre a qualidade das instituições.

Para a professora Tricia Eadie, o NQS é considerado bom pelas famílias e crianças. Entretanto, “não está claro se elas estão conscientes da diferença entre o atingir o padrão mínimo *versus* exceder os padrões de qualidade”⁴⁸ (entrevista concedida em outubro de 2016, tradução nossa). Ou seja, a noção de qualidade das famílias parece ser uma questão a ser problematizada.

Em relação aos prestadores de serviços, Tayler (2014, p. 147) assinala que o objetivo é que os mesmos tenham “sua carga regulatória reduzida (um único conjunto de regras e menos restrições)” e possam “demonstrar o valor e os benefícios de seus serviços à comunidade, à luz de um padrão comum”. Por fim, destaca a autora, “espera-se que os governos estejam mais bem posicionados para determinar o valor de seu investimento nos serviços de educação infantil”.

CONTRIBUIÇÕES DA EXPERIÊNCIA AUSTRALIANA PARA O BRASIL: ALGUMAS CONSIDERAÇÕES

A experiência australiana de implementação de políticas voltadas para a educação infantil, entre elas a avaliação da etapa, sugere para o Brasil a necessidade de se aprimorar o regime de colaboração entre os entes federados. O estudo constatou a importância da articulação entre o governo nacional e os estados/territórios australianos, que assumiram

⁴⁸ Trecho original: “unclear whether they are aware of difference between minimum versus exceeds ratings”.

o compromisso de levar a cabo a aplicação do Padrão Nacional de Qualidade para todos os serviços de Ecec, bem como de cuidar das demais políticas voltadas para essa etapa.

Não obstante nossas diferenças com a Austrália, no que tange à estrutura federativa, tipo de governo e estrutura educacional, o Plano Nacional de Educação (2014-2024) e os planos municipais para o mesmo decênio colocam a colaboração entre as três esferas de governo como uma questão imprescindível para garantir uma oferta de qualidade na educação infantil, haja vista que, sendo os municípios os principais provedores da etapa, há que se considerar que estes são a esfera economicamente mais vulnerável da federação e, portanto, necessitam de aportes financeiros das demais esferas.

A referida experiência reforça, ainda, a importância de considerar diferentes dimensões para mensurar a qualidade da educação oferecida e a necessidade de reflexão sobre os procedimentos avaliativos, com vistas a abranger diferentes demandas, entre elas, as visões sobre qualidade existentes no interior dos estabelecimentos educacionais.

No que tange ao foco da avaliação, os resultados alcançados nessa investigação evidenciam que a iniciativa australiana avalia dimensões importantes da qualidade esperada para a educação infantil brasileira, defendidas pela legislação e por documentos nacionais, norteadores da etapa, a despeito das diferenças contextuais existentes em ambos os países, e tem potencial para auxiliar na reflexão sobre as experiências em curso no Brasil.

Cabe demarcar que a avaliação do desenvolvimento/aprendizagem das crianças não é considerada no desenho avaliativo nacional da Austrália. O sistema de avaliação dos serviços de educação infantil não tem as crianças como foco, mas sim as práticas educacionais existentes nos estabelecimentos e aspectos da gestão. Dessa maneira, busca evidenciar se o estabelecimento educacional e seus profissionais estão levando em conta o que é determinado pelo *National Quality Framework*.

O referido quadro de qualidade prevê a avaliação e o acompanhamento da aprendizagem das crianças no âmbito dos serviços educacionais, realizados pelos professores, de

modo a promover a aprendizagem das crianças e obter resultados abrangentes, sem, contudo, levá-los em conta para aferir a qualidade dos serviços prestados. Esse movimento parece-nos interessante, pois valoriza o processo de aprendizagem das crianças, sem restringi-lo a determinadas habilidades/competências. Ao mesmo tempo, indica, aos serviços e seus profissionais, que a aprendizagem é fruto de ações intencionais e planejadas pelos professores.

Ressalvamos que a avaliação das crianças é fundamental. Afinal, espera-se que o resultado de ações de melhoria das/nas instituições de educação infantil seja o pleno desenvolvimento/aprendizagem de cada uma delas, dimensões fundamentais da qualidade que se deseja alcançar. Todavia, defendemos que tomá-la como instrumento de definição das políticas para a educação infantil pode levar-nos a desconsiderar aspectos estruturais e processuais necessários para a melhoria da qualidade em nosso país.

Por fim, a experiência australiana indica a necessidade de se aprimorar o processo avaliativo para abranger diferentes demandas. Isso porque esse processo nunca será totalmente perfeito, dado que a dinâmica educacional nem sempre se enquadra em padrões pré-estabelecidos, embora a existência de padrões não seja vista por nós como algo negativo, ainda mais em um país com dimensões continentais como o Brasil. Garantir o atendimento a padrões é assegurar, de alguma maneira, equidade educacional.

REFERÊNCIAS

ADRIÃO, Theresa; GARCIA, Teise; BORGHI, Raquel; ARELARO, Lisete. Uma modalidade peculiar de privatização da educação pública: a aquisição de “sistemas de ensino” por municípios paulistas. *Educação e Sociedade*, Campinas, v. 30, n. 108, p. 799-818, out. 2009.

AUSTRALIAN CHILDREN'S EDUCATION AND CARE QUALITY AUTHORITY. *Guide to the National Quality Standard*. Australia: ACECQA, 2011a.

AUSTRALIAN CHILDREN'S EDUCATION AND CARE QUALITY AUTHORITY. *National Quality Framework*. Australia: ACECQA, 2011b. Disponível em: <<http://www.acecqa.gov.au/>>. Acesso em: 07 set. 2012.

AUSTRALIAN CHILDREN'S EDUCATION AND CARE QUALITY AUTHORITY. *National quality standard assessment and rating instrument*. Australia: ACECQA, 2012.

AUSTRALIAN CHILDREN'S EDUCATION AND CARE QUALITY AUTHORITY. *NQF Snapshot Q4 2016: a quarterly report*. Australia: ACECQA, Feb. 2017.

BAUER, Adriana; HORTA NETO, João Luiz; SOUSA, Sandra Zákia; PIMENTA, Cláudia Oliveira. *Avaliação e gestão educacional em municípios brasileiros: mapeamento e caracterização de iniciativas em curso: relatório final*. São Paulo: FCC/Inep, 2016. Mimeo.

BRASIL. Lei n. 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, n. 120-A, 26 de junho de 2014, Seção 1, p. 1-7. Edição extra.

COMMONWEALTH OF AUSTRALIA. National Childcare Accreditation Council. *Putting children first: quality improvement and accreditation system handbook*. First Edition. Sydney: NCAC, 1993.

COUNCIL OF AUSTRALIAN GOVERNMENTS. *Belonging, being & becoming: the early years learning framework for Sydney*: COAG, 2009. Disponível em: <https://www.coag.gov.au/sites/default/files/early_years_learning_framework.pdf>. Acesso em: 15 ago. 2016.

FENECH, Marianne; GIUGNI, Miriam; BOWN, Kathryn. Critical analysis of the National Quality Framework: mobilising for a vision for children beyond minimum standards. *Australasian Journal of Early Childhood*, Australian Capital Territory, v. 37, n. 4, Dec. 2012.

GRIFFITH, Gareth. Managerial federalism: COAG and the states. New South Wales: NSW Parliamentary Library, 2009. (Briefing Paper 10/09).

LEVY, Evelyn. Eficácia na contratualização intergovernamental na Austrália. *Revista Gestão e Políticas Públicas*, São Paulo, v. 1, n. 2, p. 87-112, jul./dez. 2011.

LOGAN, Helen; SUMSION, Jennifer. Early childhood teachers' understandings of and provision for quality. *Australasian Journal of Early Childhood*, Australian Capital Territory, v. 35, n. 3, Sept. 2010.

MENTHA, Sue; CHURCH, Amelia; PAGE, Jane Page. Teachers as brokers: perceptions of “participation” and agency in early childhood education and care. *International Journal of Children's Rights*, Koninklijke Brill NV, Leiden, n. 23, 2015.

MINISTERIAL COUNCIL ON EDUCATION, EMPLOYMENT, TRAINING AND YOUTH AFFAIRS. *Melbourne Declaration on Educational Goals for Young Australians*. Melbourne: December 2008.

NASCIMENTO, Maria Letícia Barros Pedroso. As políticas públicas de educação infantil e a utilização de sistemas apostilados no cotidiano de creches e pré-escolas públicas. *Revista Brasileira de Educação*, Rio de Janeiro, v. 17, n. 49, jan./abr. 2012.

PIMENTA, Cláudia Oliveira. *Avaliações municipais da educação infantil: contribuições para a garantia do direito à educação das crianças brasileiras* 670 p. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2017.

PIOTTO, Débora Cristina *et al.* Promoção da qualidade e avaliação na educação infantil: uma experiência. *Cadernos de Pesquisa*, São Paulo, n. 105, p. 52-77, nov. 1998.

SOUSA, Sandra Zákia. Avaliação da educação infantil: propostas em debate no Brasil. *Interações*, Lisboa, v. 10, n. 32, p. 68-88, 2014.

SOUSA, Sandra Zákia; PIMENTA, Cláudia Oliveira. Avaliação da educação infantil: aportes de iniciativas estrangeiras. *Estudos em Avaliação Educacional*, São Paulo, v. 27, n. 65, p. 376-406, maio/ago. 2016.

TAYLER, Collette. Avaliação da qualidade da educação infantil na Austrália. Tradução de Peter Laspina. *Estudos em Avaliação Educacional*, São Paulo, v. 25, n. 58, p. 6-7, maio/ago., 2014.

VICTORIA. State Government. Parliament. *Education and Care Services National Law Act 2010*. Victoria, AU, 2010. Disponível em: <HYPERLINK "http://www.legislation.vic.gov.au" www.legislation.vic.gov.au>. Acesso em: 20 set. 2016.

WESTON, Karen; TAYLER, Collette. A policy frame on early learning and teaching. In: PAGE, Jane; TAYLER, Collette. *Learning and teaching in the early years*. Melbourne, Cambridge University, 2016.

CLÁUDIA OLIVEIRA PIMENTA

Doutora em Educação pela Faculdade de Educação da Universidade de São Paulo (FE/USP). Professora da Rede Estadual de Educação de São Paulo, São Paulo, São Paulo, Brasil
claudiapimenta68@gmail.com

Recebido em: 21 NOVEMBRO 2017

Aprovado para publicação em: 11 JANEIRO 2018

AVALIAÇÃO DA APRENDIZAGEM PROFISSIONAL NO ESTÁGIO SUPERVISIONADO: ESTUDO EXPLORATÓRIO NA EDUCAÇÃO INFANTIL

**SARA BARROS ARAÚJO
ANA PEREIRA ANTUNES**

RESUMO

O presente estudo visa analisar processos de avaliação da aprendizagem profissional na formação prática, considerando o perfil específico de desempenho docente na educação infantil. Trata-se de um estudo exploratório em que participaram estudantes que frequentaram a componente de estágio de um mestrado de natureza profissionalizante em contexto português. Os dados foram recolhidos através de um instrumento de avaliação composto por 25 dimensões organizadas em quatro domínios (Observação, Planificação, Ação e Reflexão), aplicado em dois momentos: avaliação reguladora e avaliação final. Os resultados evidenciaram evolução positiva e estatisticamente significativa em 21 das 25 dimensões analisadas, entre os dois momentos de avaliação. Os dados permitem retirar ilações acerca de percursos de aprendizagem de competências profissionais e das potencialidades de uma abordagem autêntica e coparticipada na avaliação da formação prática e na compreensão endógena dos processos formativos.

PALAVRAS-CHAVE AVALIAÇÃO DA APRENDIZAGEM • DESENVOLVIMENTO PROFISSIONAL • ESTÁGIO SUPERVISIONADO • EDUCAÇÃO INFANTIL.

EVALUACIÓN DEL APRENDIZAJE PROFESIONAL EN EL PRÁCTICUM SUPERVISADO: ESTUDIO EXPLORATORIO EN LA EDUCACIÓN INFANTIL

RESUMEN

El presente estudio tiene el propósito de analizar procesos de evaluación del aprendizaje profesional en la formación práctica, considerando el perfil específico de desempeño docente en la educación infantil. Se trata de un estudio exploratorio en el que participaron estudiantes que frecuentaron el prácticum de una maestría de formación profesional en contexto portugués. Los datos se recogieron a través de un instrumento de evaluación compuesto por 25 dimensiones organizadas en cuatro dominios (Observación, Planificación, Acción y Reflexión), aplicado en dos momentos: evaluación reguladora y evaluación final. Los resultados pusieron de manifiesto una evolución positiva y estadísticamente significativa en 21 de las 25 dimensiones analizadas, entre los dos momentos de evaluación. Los datos permiten retirar ilaciones acerca de recorridos de aprendizaje de competencias profesionales y de las potencialidades de un abordaje auténtico con coparticipación en la evaluación de la formación práctica y en la comprensión endógena de los procesos formativos.

PALABRAS CLAVE EVALUACIÓN DEL APRENDIZAJE • DESARROLLO PROFESIONAL • PRÁCTICUM SUPERVISADO • EDUCACIÓN INFANTIL.

ASSESSMENT OF PROFESSIONAL LEARNING IN SUPERVISED PRACTICUM: EXPLORATORY STUDY IN EARLY CHILDHOOD EDUCATION

ABSTRACT

This study aims to analyze professional learning assessment procedures considering the specific teaching performance profile in early childhood education. A cohort of Portuguese students involved in a practicum in early childhood education, at a master level, participated in this exploratory study. Data were collected using an assessment instrument composed of 25 dimensions organized in four domains (Observation, Planning, Action and Reflection), applied at two different instances: regulatory assessment and final assessment. The results showed a statistically significant and positive evolution in 21 of the 25 dimensions examined, between the two instances of assessment. Data allow conclusions to be drawn regarding the learning of professional skills, and the potentialities of an authentic and co-participative approach in the assessment of practical training and in the endogenous understanding of training processes.

KEYWORDS LEARNING ASSESSMENT • PROFESSIONAL DEVELOPMENT • SUPERVISED PRACTICUM • EARLY CHILDHOOD EDUCATION.

INTRODUÇÃO

O Processo de Bolonha trouxe, em Portugal, mudanças significativas no regime jurídico de habilitação para a docência na educação infantil, regulamentada, numa primeira fase, através do Decreto-lei n. 43/2007, de 22 de fevereiro (PORTUGAL, 2007), entretanto substituído pelo Decreto-lei n. 79/2014, de 14 de maio (PORTUGAL, 2014). Dentre as várias mudanças registadas, é de salientar que a formação para docência na educação infantil passou a ser realizada num modelo sequencial de dois ciclos de estudos: uma licenciatura de três anos em educação básica, destinada a assegurar a formação de base na área da docência, e um mestrado profissionalizante destinado a aprofundar e reforçar a formação acadêmica, particularmente nas didáticas específicas e na prática supervisionada. Dentre as cinco componentes de formação¹ previstas para os ciclos de estudos de habilitação profissional para a docência, é de assinalar a “Iniciação à prática profissional”, que inclui a prática supervisionada na sala de atividades ou sala de aula. Esta componente era concebida, no articulado do Decreto-lei n. 43/2007,

¹ Os ciclos de estudos que visam a habilitação profissional para a docência integram as seguintes componentes de formação: (i) área de docência; (ii) área educacional geral; (iii) didáticas específicas; (iv) área social, cultural e ética; (v) iniciação à prática profissional (PORTUGAL, 2014, art. 7º).

[...] numa perspectiva de desenvolvimento profissional dos formandos visando o desempenho como futuros docentes e promovendo uma postura crítica e reflexiva em relação aos desafios, processos e desempenhos do quotidiano profissional. (PORTUGAL, 2007, art. 14º)

No Decreto-lei n. 79/2014, este articulado sofre alterações, passando a elencar, no art. 11º, dentre outros, os seguintes princípios de organização da Iniciação à prática profissional:

[...] d) É concebida numa perspectiva de formação para a articulação entre o conhecimento e a forma de o transmitir visando a aprendizagem; e) É concebida numa perspectiva de desenvolvimento profissional dos formandos e promove nestes uma atitude orientada para a permanente melhoria da aprendizagem dos seus alunos. (PORTUGAL, 2014, art. 11º)

Apesar de inflexões relevantes na forma de concetualizar esta componente de formação, nomeadamente no nexo entre educação/ensino e aprendizagem introduzido na versão revista do regime jurídico, bem como na retração ao nível da assunção de uma postura crítica e reflexiva, manteve-se a tónica numa perspectiva de desenvolvimento profissional associado a esta componente formativa.

O presente estudo centra-se na formação prática de educadores/as de educação infantil e, muito particularmente, no estágio desenvolvido no âmbito de um mestrado profissionalizante, reconhecendo-o como espaço e tempo de relevância central na formação destes docentes. Gomes (2013) considera o estágio uma oportunidade de ação profissional qualificada, de diálogo com o mundo do trabalho através de sua análise crítica, articulando dialeticamente o saber teórico e o saber prático (desenvolvimento da práxis). Em convergência surgem as ideias de Calderano (2012, p. 270), que afirma:

[...] o estágio curricular não se reduz à prática, no sentido operacional. O estágio precisa se constituir em uma oportunidade de estreitar as relações entre reflexões teórico-práticas, visualização ampliada do processo educacional, descoberta e delineamento de desafios encontrados para,

numa perspectiva de síntese, atingir a construção conjunta de alternativas pedagógicas para os problemas identificados.

Por seu turno, Dayan (2008) adianta que o estágio, no âmbito dos programas de formação de educadores/as de infância, envolve os estudantes num processo desenvolvimental de observação e experimentação de práticas de trabalho com crianças, bem como de aprendizagem de competências, conhecimentos, filosofias e atitudes em contextos educativos.

Pela sua relevância no percurso formativo dos estudantes, reconhece-se a necessidade de processos de avaliação que, numa lógica de monitorização e melhoria das práticas, preservem a natureza complexa da atividade docente enquanto atividade intelectual, técnica, moral e relacional (FORMOSINHO, 2002), sejam abertos a um perfil polifacetado de competências (FORMOSINHO; MACHADO, 2010) e facilitadores da construção de conhecimento profissional de tipo praxeológico (FORMOSINHO; MACHADO, 2009). Formosinho (2009) considera que as formas de avaliar e regular as aprendizagens fazem parte da dimensão manifesta de um currículo de processos na formação de docentes, constituindo práticas particularmente marcantes de um currículo oculto e não problematizado. No âmbito mais circunscrito da formação de profissionais da educação infantil, Jensen (2015) adianta que a identificação de modos adequados para avaliar a aprendizagem dos estudantes nos contextos da prática constitui um desafio aos programas de formação. No presente trabalho, assume-se que os instrumentos intervenientes e procedimentos implicados na avaliação do desempenho dos estagiários deverão ser alvo de análise, não só pelo impacto que poderão ter nos percursos de aprendizagem profissional, mas também pela influência na compreensão dos processos formativos.

AVALIAÇÃO DE DESEMPENHO NA FORMAÇÃO PRÁTICA INICIAL

De acordo com Caughlan e Jiang (2014), a avaliação do desempenho específica o que os docentes ou candidatos a docentes

devem saber e saber fazer em contextos reais de ensino. Esta avaliação tem vindo a ser associada à melhoria das práticas de ensino dos futuros docentes (CAUGHLAN; JIANG, 2014) e nos casos em que se reconhece sua natureza formativa como influente nos seus processos de aprendizagem e reflexividade (WEI; PECHEONE, 2010).

Darling-Hammond (2001) considera que a avaliação de desempenho de futuros docentes deve ser baseada em visões complexas e holísticas do ensino e em padrões de desempenho profissional validados, numa linha de avaliação autêntica sensível aos contextos, longitudinal e individualizada. A autora considera que uma avaliação desta natureza oferece medidas mais válidas do conhecimento e competências dos professores que os testes tradicionais (DARLING-HAMMOND, 2006; DARLING-HAMMOND; NEWTON; WEI, 2013). Na ótica de Darling-Hammond e Snyder (2000), sob o escopo da avaliação autêntica incluem-se oportunidades para o desenvolvimento e análise do pensamento e ações dos professores em situações baseadas na experiência e orientadas para a resolução de problemas. Os autores salientam que esta abordagem autêntica à avaliação dos futuros docentes os apoia a deslocarem-se de generalizações acerca da prática para instâncias de aprendizagem aparentemente mais idiossincráticas, contextualizadas e responsivas à multidimensionalidade dos problemas e possibilidades de cada aluno. Paralelamente, uma abordagem autêntica à avaliação poderá inspirar alterações nos programas de formação, pelo providenciar de informação relevante acerca das características dos candidatos a docentes, aspeto crítico no âmbito de uma agenda de avaliação com preocupações dirigidas, em paralelo à documentação e melhoria da formação docente (DARLING-HAMMOND, 2006).

Na revisão da literatura realizada por Alves e Cabral (2015), a avaliação assim concebida enquadra-se numa quarta geração da avaliação pedagógica que adota uma perspectiva construtivista, participativa e negocial no âmbito do processo avaliativo. Especificamente no que concerne ao perfilhar de uma abordagem participativa, Wiggins (1989) salienta que a avaliação torna-se mais rigorosa, justa e equitativa quando envolve a opinião e o diálogo, integrando o

input e feedback daquele que está a ser avaliado. Assim, a responsividade aos contextos estende-se à responsividade aos indivíduos, reconhecendo-se a justiça e o rigor de práticas democráticas concretizadas através do exercício de deliberação coletiva (MOSS; DALHERG, 2008). Caughlan e Jiang (2014) propõem o conceito de “profissionalismo agêntico” para se referirem à importância de a escolha e utilização de instrumentos de avaliação do desempenho considerarem aspetos como o posicionamento, poder e agência dos estudantes ou a criação de oportunidades de tomada de decisão. De facto, os autores salientam que estes instrumentos não são neutros, mas refletem os valores dos programas que os utilizam; argumentam, ainda, que os programas deverão manter vigilância relativamente à consistência entre a teoria de aprendizagem preconizada e praticada e a teoria de avaliação, uma vez que ambas se encontram, frequentemente, em dissonância.

Concentrada na formação de educadores de infância, Dayan (2008) defende uma abordagem humanístico-democrática no âmbito dos processos supervisivos, em que estudantes, orientadores cooperantes e supervisores institucionais se possam envolver em parcerias e processos democráticos de deliberação e discussão. São corolários desta abordagem o desenvolvimento da autonomia do estagiário e a promoção do sentido de responsabilidade pela sua aprendizagem e pelo seu crescimento profissional. Numa linha semelhante, Dalrymple, Kemp e Smith (2014) sustentam a relevância de um modelo triádico em que o estudante, o facilitador do contexto de trabalho e o tutor da universidade detêm papéis interdependentes e harmonizados. Os autores referem que as configurações mais eficazes do ponto de vista da aprendizagem em contextos de trabalho (como é o caso dos estágios profissionalizantes) envolvem a confluência destes três atores no processo de aprendizagem e a criação de oportunidades, entre outras, para que os estudantes se tornem coprodutores de conhecimento. Reconhecem, ainda, o papel dos supervisores no âmbito daquilo que designam *facilitação responsiva*, caracterizada centralmente pelo encorajamento à identificação e avaliação crítica pelos

estudantes das suas compreensões e práticas. Por seu turno, Calderano (2013, p. 48) aponta, no seu extenso estudo acerca do estágio curricular na formação inicial e continuada, as possibilidades de uma “conceção marcada pela organicidade entre escola e universidade, entre teoria e prática, entre formação e trabalho docente”. A autora salienta a relevância da conceção e exercício de uma docência compartilhada no âmbito dos estágios, em que os professores da universidade e da escola se envolvem num trabalho colegial (CALDERANO, 2012, 2013). Nesta conceção, o estatuto epistemológico do estagiário é revisto, cabendo-lhe uma participação ativa no processo formativo e, por inerência, na própria avaliação (CALDERANO, 2013).

O presente estudo inscreve-se numa avaliação autêntica e coparticipada do desempenho na formação prática inicial de educadores infantis. Tem como objetivo analisar a evolução de estudantes de um mestrado profissionalizante ao nível de competências profissionais relevantes no que respeita ao perfil específico de desempenho docente, com base numa avaliação participada e individualizada. Visa, igualmente, numa lógica de monitorização interna, servir de base à compreensão endógena dos processos formativos, procurando sua permanente adequação a necessidades identificadas (por exemplo, ao nível dos processos supervisivos).

MÉTODO

PARTICIPANTES

Participaram no estudo 41 estudantes (40 mulheres e um homem) que frequentaram o estágio de um mestrado profissionalizante em educação infantil numa instituição de ensino superior do norte de Portugal, nos anos letivos de 2012/2013 (19 estudantes, 18 mulheres e um homem) e 2013/2014 (22 estudantes mulheres).

INSTRUMENTO

Utilizou-se a Grelha de Avaliação sobre os Processos de Desenvolvimento na Prática Pedagógica Supervisionada (GAPD_PPS). A utilização desta grelha, de base observacional

e reflexiva, tem a dupla função de servir à avaliação reguladora e avaliação final do processo de desenvolvimento de conhecimentos e competências profissionais. Sua construção tomou em consideração o perfil específico de desempenho profissional do educador de infância (PORTUGAL, 2001), os princípios orientadores da formação na prática pedagógica da instituição formadora e, ainda, referentes teóricos relevantes no âmbito da formação, designadamente os modelos curriculares para a educação de infância (OLIVEIRA-FORMOSINHO, 2007). A grelha encontra-se organizada em quatro domínios que integram dimensões centrais do desempenho docente: (i) Observação (cinco dimensões); (ii) Planificação (seis dimensões); (iii) Ação (nove dimensões); e (iv) Reflexão (cinco dimensões), como se pode verificar no Quadro 1.

QUADRO 1 - Domínios e dimensões da GAPD_PPS

DOMÍNIOS E DIMENSÕES DA GAPD_PPS			
OBSERVAÇÃO	PLANIFICAÇÃO	AÇÃO	REFLEXÃO
D1: Conhecimentos teóricos e de enquadramento legal	D6: Conhecimento do currículo e dos conteúdos	D12: Organização do espaço e materiais	D21: Reflexão sobre, na e para a ação
D2: Observação da criança e dos grupos	D7: Coerência com o projeto curricular do grupo	D13: Organização do tempo	D22: Mobilização de referentes teóricos e legais
D3: Práticas de registo	D8: Estratégias e recursos	D14: Interações pedagógicas	D23: Postura indagadora e crítica
D4: Recolha de informação sobre os contextos	D9: Articulação, objetivos, estratégias e avaliação	D15: Estratégias diferenciadas	D24: Articulação de diversos pontos de vista, valores, saberes e crenças
D5: Auto e hétero-observação	D10: Integração de propostas das crianças e situações imprevistas	D16: Envolvimento das crianças em projetos	D25: Avaliação crítica acerca do seu papel e funções
	D11: Natureza integradora e diferenciação pedagógica	D17: Integração de situações de jogo espontâneo	
		D18: Construção de parcerias com família e comunidade	
		D19: Interações colaborativas com equipas de sala e centro educativo	
		D20: Postura democrática, ética e deontológica	

Fonte: Elaboração das autoras.

Nota: D: Dimensão.

PROCEDIMENTOS

A GAPD_PPS foi utilizada em dois momentos do estágio: numa fase intermédia, com propósitos de monitorização do processo de aprendizagem profissional, e, no final do estágio, com objetivo de avaliação final, com consequências ao nível da avaliação sumativa. Num momento inicial, a GAPD_PPS foi apresentada às/aos estagiárias e orientadoras cooperantes (educadoras de infância que acompanham o estágio em contextos da prática), clarificando-se dúvidas relativamente às dimensões que a compõem. Posteriormente, sua utilização, em cada um dos momentos de avaliação, concretizou-se em três fases: (i) autoavaliação, compreendendo uma reflexão retrospectiva e prospetiva da/o estagiária/o acerca do seu desempenho em cada uma das dimensões; (ii) reunião reflexiva entre as/os estagiárias/os (o estágio é realizado em pares de formação) e respetiva orientadora cooperante centrada em processos de auto e heteroavaliação dos desempenhos; (iii) reunião entre as/os estagiárias/os, orientadora cooperante e supervisor/a institucional realizada nos contextos de estágio, visando a análise colaborativa e coparticipada dos processos de aprendizagem profissional.

No caso da avaliação reguladora, da reunião de auto e heteroavaliação e reflexão resultou um conjunto de intenções e prioridades ao nível da melhoria do desempenho docente em cada uma das dimensões, explicitadas no espaço da grelha dedicado a “observações”. Destas reuniões resultou também a atribuição consensualizada de um nível de desempenho em cada uma das dimensões, numa escala com a seguinte correspondência: A: excelente; B: muito bom; C: bom; D: suficiente; E: insuficiente (sendo que, em alguns dos casos, recorreu-se a uma valoração intermédia como, por exemplo, B+: muito bom, mas ainda não ao nível do excelente). Esta atribuição pretende a trasladação dos resultados de um processo reflexivo para uma abordagem tendentemente mais quantitativa, facilitadora da leitura da análise evolutiva da/o estudante.

No caso da reunião de avaliação final, esta atribuição associa-se à necessidade de classificação inerente à avaliação de tipo sumativo para a qual concorre este momento reflexivo entre formadoras/es e estagiárias/os. O presente estudo faz uso destas atribuições para analisar a evolução das/os

estudantes nas dimensões da GAPD_PPS. Mais concretamente, procedeu-se a uma equiparação, para efeitos deste estudo, dos valores qualitativos a valores quantitativos, a saber: às classificações de A, B, C, D e E fez-se a correspondência das classificações 5, 4, 3, 2 e 1, respetivamente.

RESULTADOS

Nesta secção apresentam-se os resultados obtidos através da utilização da GAPD_PPS em cada um dos seus domínios (Observação, Planificação, Ação e Reflexão). Dado o número de participantes no estudo não ser muito alargado (41 sujeitos) e a escala de medida das variáveis em estudo poder ser considerada ordinal, optou-se por uma análise estatística para comparação de resultados através de testes não paramétricos (MARTINS, 2011), mais concretamente o teste não paramétrico de Wilcoxon (*signed-rank*) (FIELD, 2009).

Ao iniciar-se a apresentação dos resultados pelo domínio Observação, a consulta à Tabela 1 permite constatar diferenças estatisticamente significativas em todas as dimensões avaliadas neste domínio ($p < 0,05$), traduzindo uma diferenciação dos resultados dos alunos nos dois momentos de avaliação, nos quais se regista uma melhoria dos resultados no momento de avaliação final.

TABELA 1 - Apresentação dos resultados obtidos na avaliação reguladora e na avaliação final no que se refere às dimensões do domínio Observação

DIMENSÃO	MOMENTO	N	MEDIANA	MÉDIA	DESVIO-PADRÃO	MÍN.	MÁX.	TESTE DE WILCOXON	
								Z	p
1	I	41	2,5	2,7	0,69	1	4,5	-4,371	0,000
	II	41	3	3,4	0,81	1	4,5		
2	I	41	3	3,1	0,52	2	4	-3,452	0,001
	II	41	3,5	3,4	0,67	2	4		
3	I	41	3	2,8	0,81	1	4	-3,369	0,001
	II	41	3	3,2	0,72	1	4		
4	I	41	3	3,1	0,61	2	4	-3,147	0,002
	II	41	3,5	3,5	0,76	2	4,5		
5	I	41	3	2,9	0,53	1	4	-3,257	0,001
	II	41	3	3,3	0,67	1	4		

Fonte: Elaboração das autoras.

Momento: I - Avaliação reguladora; II - Avaliação final.

Em relação ao domínio Planificação, os resultados encontrados seguem a tendência registrada previamente no domínio Observação, ou seja, assiste-se a uma melhoria estatisticamente significativa no desempenho dos alunos, favorável ao momento de avaliação final ($p < 0,05$), conforme se constata na Tabela 2.

TABELA 2 – Apresentação dos resultados obtidos na avaliação reguladora e na avaliação final no que se refere às dimensões do domínio Planificação

DIMENSÃO	MOMENTO	N	MEDIANA	MÉDIA	DESVIO-PADRÃO	MÍN.	MÁX.	TESTE DE WILCOXON	
								Z	p
6	I	41	3	3,0	0,61	2	4	-3,507	0,000
	II	41	3,5	3,4	0,70	1	4,5		
7	I	41	3	3,3	0,55	2	4	-3,139	0,002
	II	41	4	3,6	0,68	2	4,5		
8	I	41	3	3,0	0,55	2	4,5	-4,036	0,000
	II	41	3,5	3,6	0,59	2	4,5		
9	I	41	3	2,9	0,61	1	4	-3,821	0,000
	II	41	3,5	3,6	0,74	1	4,5		
10	I	41	3	3,1	0,70	2	4	-3,293	0,001
	II	41	4	3,5	0,82	1	4,5		
11	I	41	3	2,9	0,54	2	4,5	-3,403	0,001
	II	41	3	3,4	0,52	2,5	4,5		

Fonte: Elaboração das autoras.

Momento: I - Avaliação reguladora; II - Avaliação final.

Uma análise mais detalhada da Tabela 2 permite verificar que na dimensão 6, na avaliação final, houve uma queda no valor mínimo atribuído (desceu de 2 para 1), pelo que se decidiu proceder à análise individual dos sujeitos, procurando compreender este resultado, o que permitiu verificar que este se deve à descida de desempenho do sujeito 8. Também na dimensão 10, na avaliação final, regista-se a diminuição da classificação mínima (desceu de 2 para 1), sendo que esta se deve igualmente à descida de desempenho por parte do sujeito 8.

No que diz respeito ao domínio Ação, os resultados referentes às dimensões avaliadas são apresentados na Tabela 3, em função do momento de avaliação, sendo que os alunos não parecem ter evoluído de forma estatisticamente significativa em todas as dimensões consideradas.

TABELA 3 - Apresentação dos resultados obtidos na avaliação reguladora e na avaliação final no que se refere às dimensões do domínio Ação

DIMENSÃO	MOMENTO	N	MEDIANA	MÉDIA	DESVIO-PADRÃO	MÍN.	MÁX.	TESTE DE WILCOXON	
								Z	p
12	I	41	3	2,8	0,66	2	4,5	-4,280	0,000
	II	41	3,5	3,5	0,58	2	4,5		
13	I	41	3	3,1	0,71	1	4,5	-1,923	0,054
	II	41	3,5	3,4	0,66	2	4,5		
14	I	41	3	3,1	0,75	1	4	-3,131	0,002
	II	41	4	3,4	0,76	1	4,5		
15	I	41	3	3,0	0,54	1	4,5	-4,127	0,000
	II	41	4	3,5	0,65	1,5	4,5		
16	I	27	3	2,7	0,87	1	4	-3,488	0,000
	II	41	3,5	3,6	0,58	2,5	4,5		
17	I	41	3,5	3,4	0,71	1	4,5	-1,503	0,133
	II	41	4	3,5	0,67	2	4,5		
18	I	40	2	2,4	0,92	1	5	-4,686	0,000
	II	41	3	3,4	0,62	2	4,5		
19	I	41	3	3,3	0,73	1	5	-2,628	0,009
	II	41	4	3,6	0,65	2	4,5		
20	I	41	4,5	4,3	0,72	3	5	-0,258	0,796
	II	41	5	4,3	0,79	3	5		

Fonte: Elaboração das autoras.

Momento: I - Avaliação reguladora; II - Avaliação final.

A análise da Tabela 3 permite constatar que, entre a avaliação reguladora e a avaliação final, registam-se dimensões em que ocorreu subida no valor das médias, sendo que a diferença na distribuição de resultados revela-se estatisticamente significativa ($p < 0,05$) para a dimensão 12. A dimensão 13 regista apenas uma diferenciação marginalmente significativa ($p = 0,054$). No entanto, para a dimensão 17 e para a dimensão 20 não se registam diferenças estatisticamente significativas entre os dois momentos de avaliação ($p > 0,05$).

Verifica-se, ainda, que nas dimensões 18 e 19 houve diminuição do valor máximo atribuído (desceu de 5 para 4,5) na avaliação final, pelo que se decidiu analisar os desempenhos individuais dos alunos, constatando-se que o aluno 17, que teve, na avaliação reguladora, a classificação de 5, desceu para 3 na avaliação final, enquanto o aluno 34, que na dimensão 19 teve, na avaliação reguladora, a classificação de 5, desceu para 3,5 na avaliação final.

O domínio Reflexão, o último em análise (Tabela 4), caracteriza-se pela diferenciação estatisticamente significativa dos resultados obtidos nas duas avaliações ($p < 0,05$), apontando para resultados superiores na avaliação final em quatro das dimensões avaliadas, ou seja, na dimensão 21, na dimensão 22, na dimensão 23 e na dimensão 25. Todavia, para a dimensão 24 não se registam diferenças estatisticamente significativas entre os dois momentos de avaliação ($p > 0,05$).

TABELA 4 – Apresentação dos resultados obtidos na avaliação reguladora e na avaliação final no que se refere às dimensões do domínio Reflexão

DIMENSÃO	MOMENTO	N	MEDIANA	MÉDIA	DESVIO-PADRÃO	MÍN.	MÁX.	TESTE DE WILCOXON	
								Z	p
21	I	41	3	2,9	0,67	1	4	-3,641	0,000
	II	41	3,5	3,4	0,75	1	4,5		
22	I	41	3	3,0	0,85	1	4,5	-3,565	0,000
	II	41	3,5	3,5	0,82	1	4,5		
23	I	41	3	2,9	0,68	1	4,5	-3,618	0,000
	II	41	3,5	3,4	0,79	1	4,5		
24	I	41	3	3,3	0,62	2	4,5	-1,732	0,083
	II	41	4	3,5	0,75	1	4,5		
25	I	31	3	2,7	0,58	2	4	-3,230	0,001
	II	41	3	3,3	0,8	1	4,5		

Fonte: Elaboração das autoras.

Momento: I – Avaliação reguladora; II – Avaliação final.

A partir da análise da Tabela 4, verifica-se, nas dimensões 24 e 25, a diminuição da avaliação mínima atribuída (desceu de 2 para 1) na avaliação final, pelo que mais uma vez procedeu-se à análise dos desempenhos individuais, permitindo constatar que novamente o aluno 8 surge como o responsável pelo registo destes valores, sendo que, na dimensão 24, desce de 2 para 1, e, na dimensão 25, seu desempenho não permitiu a atribuição de uma classificação na avaliação reguladora, o qual, na avaliação final, é avaliado com 1 valor.

DISCUSSÃO DOS RESULTADOS

As análises efetuadas permitiram verificar que há diferenças estatisticamente significativas nos quatro domínios analisados entre os dois momentos de avaliação.

No que concerne às dimensões que compõem a GAPD_PPS, as análises realizadas evidenciaram uma evolução estatisticamente significativa em 21 das 25 dimensões dos quatro domínios avaliados (Observação, Planificação, Ação e Reflexão). Mais concretamente, regista-se uma evolução do grupo de estudantes ao nível das dimensões da Observação e da Planificação. Em quatro dimensões (D13, D17, D20, D24) da GAPD_PPS, não se verificou evolução estatisticamente significativa, apesar de se ter registado, nestas dimensões, um aumento na média e na mediana quando se comparam a avaliação reguladora e a avaliação final. Três destas dimensões enquadram-se no domínio Ação (D13, D17 e D20) e a restante (D24) enquadra-se no domínio Reflexão.

Perante estes resultados, procedeu-se à análise de cada um dos itens em concreto. Assim, para D13 (organização do tempo), ainda que a diferença possa ser considerada marginalmente significativa ($p=0,054$), sendo um domínio relacionado com a capacidade de organização e gestão do tempo, os resultados podem traduzir o facto de o grupo de estagiários, apesar da evolução que conseguiu, não ter evoluído de forma homogênea, ao mesmo ritmo e no mesmo período de tempo, condicionando assim o valor da mediana e o valor médio registados no grupo. Aliás, esta explicação parece-nos que também se poderá aplicar em relação ao D17, domínio relacionado com a integração de situações de jogo espontâneo, bem como ao D20, referente à postura democrática, ética e deontológica, e ao D24, articulação de diversos pontos de vista, valores, saberes e crenças. Em relação ao D20, acrescenta-se o facto de os alunos, na avaliação reguladora e na avaliação final, registarem os mesmos valores médios, sendo que a mediana na avaliação reguladora já se situava num valor bastante elevado, correspondendo à apreciação de muito bom+, alcançando esta, na avaliação final, o nível excelente. Nos outros três domínios, os valores da mediana são mais baixos e remetem para uma evolução grupal que se situa entre os níveis bom e muito bom. Mais especificamente, para D13, regista-se a evolução de bom para bom+; para D17, a evolução de bom+ para muito bom; e, para D24, a evolução de bom para muito bom.

Considerando os dados apresentados, estas serão dimensões a merecer particular atenção nos processos avaliativos, para melhor compreender as tendências evolutivas observadas e a análise de casos individuais. Um dos aspetos a tomar em consideração refere-se ao diferencial grau de exigência ao nível da aprendizagem profissional destas dimensões/competências. De facto, levanta-se a hipótese de algumas destas dimensões representarem um maior desafio ao desenvolvimento profissional, designadamente no caso em análise, a D13 (organização do tempo) e a D24 (articulação de diversos pontos de vista, valores, saberes e crenças).

Detendo-nos na análise dos desempenhos individuais, foi possível verificar que 40 dos 41 participantes evidenciaram evolução positiva entre os dois momentos de avaliação. Um dos participantes (8) registou resultados mais baixos na avaliação final, comparativamente à avaliação reguladora. A análise mais específica desta situação permitiu constatar tratar-se de um estudante que não atingiu os objetivos propostos do ponto de vista do desenvolvimento de competências profissionais durante o estágio no ano letivo 2012/13, tendo realizado com sucesso o estágio no ano letivo seguinte.

Verificamos, ainda, que nas dimensões 16 (envolvimento das crianças em projetos), 18 (construção de parcerias com família e comunidade) e 25 (avalia reflexivamente o seu papel e funções considerando princípios de participação e responsabilidade social) não foram registadas respostas para todos os participantes na avaliação reguladora. Relativamente às duas primeiras dimensões, o preenchimento da dimensão não se revelava aplicável atendendo ao caso concreto do aluno e do local de estágio. Este facto sugere uma alteração à própria grelha, introduzindo-se uma coluna com a designação “não aplicável”, por forma a permitir um registo mais rigoroso de circunstâncias desta natureza. Em relação à dimensão 25, as dificuldades iniciais de interpretação do enunciado sugerem a necessidade de revisão da sua operacionalização.

Retomando o objetivo central do presente estudo, estes dados permitem retirar ilações acerca dos percursos evolutivos das/os estudantes deste ciclo de estudos, reconhecendo

a existência de evidência acerca do desenvolvimento de conhecimentos e competências profissionais em domínios estruturantes da profissionalidade docente. Tratando-se de um processo formativo, esta evolução das/os estudantes seria expectável, fruto do conhecimento prévio das intenções formativas e da direção desejável ao nível de um percurso bem-sucedido de aprendizagem profissional. Sublinhamos, neste apartado, alguns procedimentos adotados para minorar a possibilidade de influência destes fatores, designadamente as três instâncias de avaliação utilizadas (autoavaliação; auto e heteroavaliação entre estagiárias/os e orientadora cooperante; auto e heteroavaliação entre todos os participantes implicados, estagiárias/os, orientadoras cooperantes e supervisoras/es institucionais), bem como o facto de o estudo se basear num conjunto de atribuições consensualizadas entre as/os formadoras/es e as/os estagiárias/os, o que poderá contribuir para robustecer a confiabilidade nestes dados. Refira-se, ainda, que estes procedimentos seguem práticas de qualidade internacionalmente reconhecidas ao nível da avaliação de estágios na educação infantil, nomeadamente o foco numa abordagem processual e criticamente reflexiva e a responsabilidade partilhada pelas instituições de ensino superior e de estágio, que têm vindo a caracterizar as práticas avaliativas na Itália (BALDUZZI; LAZZARI, 2015) e na Dinamarca (JENSEN, 2015).

CONSIDERAÇÕES FINAIS

Como sublinhado por Formosinho (2009), a avaliação da aprendizagem de futuros docentes faz parte da dimensão manifesta de um currículo de processos que, apesar da sua relevância, é descurada em termos de problematização. O caso particular da avaliação da aprendizagem profissional durante o estágio profissionalizante segue esta tendência, visível na escassez de estudos que a tenham reclamado enquanto objeto. De facto, e a título exemplificativo, a análise de um recente número temático da revista *Early Years: An International Research Journal*, dedicado à aprendizagem em contexto de trabalho e respetivas práticas de acompanhamento na

formação de educadoras/es de infância em vários países europeus (OBERHUEMER, 2015), permitiu constatar a atenção muito marginal dada à avaliação da aprendizagem profissional no âmbito de processos de estágio na educação infantil.

O presente artigo apresenta uma proposta de avaliação da aprendizagem de competências profissionais estruturantes no que diz respeito ao perfil específico de desempenho profissional na educação infantil, em contexto português. Esta proposta parte, por um lado, da necessidade de tomar em consideração padrões de desempenho profissional validados e, por outro, de um conjunto de princípios e procedimentos que visam manter a natureza reflexiva, autêntica e participada dos processos avaliativos. Esta opção por uma avaliação autêntica, contextualizada e individualizada (DARLING-HAMMOND, 2001, 2006), envolvendo vários atores relevantes ao processo avaliativo, incluindo a/o própria/o estagiária/o (CALDERANO, 2013; CAUGHLAN; JIANG, 2014; DAYAN, 2008; WIGGINS, 1989), pretende traduzir-se em benefício para várias partes. Em primeira instância, benefícios para a/o estudante que, estando ativamente envolvida/o num processo coletivo de reflexão e deliberação, torna-se agente ativo do seu próprio processo de desenvolvimento profissional, possibilitando um afastamento de uma avaliação dirigida por lógicas burocráticas e corporativas e mais profissional (FORMOSINHO; MACHADO, 2010). De facto, é possível afirmar que o processo avaliativo é promotor de competências que ele próprio avalia, garantindo a coerência entre a teoria de aprendizagem e a teoria de avaliação preconizadas. Outro efeito positivo diz respeito ao robustecimento da confiabilidade nos processos e resultados dos processos avaliativos e investigativos, resultante de análises e deliberações participadas.

O estudo aqui relatado permitiu constatar a evolução positiva das/os estudantes deste mestrado profissionalizante em educação infantil do ponto de vista da aprendizagem profissional. Simultaneamente, e reconhecendo que a avaliação de desempenho na formação prática inicial tem vindo a ser assinalada como relevante na avaliação da qualidade dos programas de formação de docentes (DARLING-HAMMOND, 2006; WEI; PECHEONE, 2010), sublinha-se também o contributo

deste estudo na compreensão endógena dos processos formativos. Neste caso concreto, parece-nos possível afirmar a eficácia percebida da formação realizada no âmbito deste estágio profissionalizante, reconhecendo-se a necessidade de levar a cabo uma análise das estratégias formativas que mais efeitos estarão a ter ao nível da aprendizagem profissional.

REFERÊNCIAS

ALVES, José Matias; CABRAL, Ilídia. Os demónios da avaliação: memórias de professores enquanto alunos. *Estudos de Avaliação Educacional*, São Paulo, v. 26, n. 63, p. 630-662, set./dez. 2015.

BALDUZZI, Lucia; LAZZARI, Arianna. Mentoring practices in workplace-based professional preparation: a critical analysis of policy developments in the Italian context. *Early Years: An International Research Journal*, Abingdon, v. 35, n. 2, p. 124-138, 2015.

CALDERANO, Maria da Assunção. O estágio supervisionado para além de uma atividade curricular: avaliação e proposições. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 53, p. 250-278, set./dez. 2012.

CALDERANO, Maria da Assunção. *Docência compartilhada entre universidade e escola: formação inicial e continuada através do estágio curricular*. São Paulo: FCC; CNPQ, 2013. Disponível em: <<http://www.fcc.org.br/fcc/images/pesquisa/estagios/projetos/MariaAssuncaoCalderano.pdf>>. Acesso em: 21 mar. 2016.

CAUGHLAN, Samantha; JIANG, Heng. Observation and teacher quality: critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education*, Washington, v. 65, n. 5, p. 375-388, 2014.

DALRYMPLE, Roger; KEMP, Chris; SMITH, Patrick. Characterising work-based learning as a triadic learning endeavor. *Journal of Further and Higher Education*, London, v. 38, n. 1, p. 75-89, 2014.

DARLING-HAMMOND, Linda. Standard setting in teaching: changes in licensing, certification, and assessment. In: RICHARDSON, Virginia (Ed.). *Handbook of research on teaching*. 4. ed. Washington: American Educational Research Association, 2001. p. 751-776.

DARLING-HAMMOND, Linda. Assessing teacher education: the usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education*, Washington, v. 57, n. 2, p. 120-138, 2006.

DARLING-HAMMOND, Linda; NEWTON, Stephen; WEI, Ruth Chung. Developing and assessing beginning teacher effectiveness: the potential of performance assessments. *Educational Assessment, Evaluation and Accountability*, Berlin, v. 25, n. 3, p. 179-204, 2013.

DARLING-HAMMOND, Linda; SNYDER, Jon. Authentic assessment of teaching in context. *Teaching and Teacher Education*, v. 16, p. 523-545, 2000.

DAYAN, Yael. Towards professionalism in early childhood practicum supervision: a personal journey. *European Early Childhood Education Research Journal*, Birmingham, v. 16, n. 2, p. 153-170, 2008.

FIELD, Andy. *Discovering statistics using SPSS*. 3. ed. London: Sage, 2009.

FORMOSINHO, João. A universidade e a formação de educadores de infância: potencialidades e dilemas. In: MACHADO Maria Lúcia (Org.). *Encontros e desencontros em educação infantil*. São Paulo: Cortez, 2002. p. 169-188.

FORMOSINHO, João. A formação prática dos professores: da prática docente na instituição de formação à prática pedagógica nas escolas. In: FORMOSINHO, João (Coord.). *Formação de professores: aprendizagem profissional e acção docente*. Porto: Porto, 2009. p. 93-117.

FORMOSINHO, João; MACHADO, Joaquim. Professores nas escolas de massas: novos papéis, nova profissionalidade. In: FORMOSINHO, João (Coord.). *Formação de professores: aprendizagem profissional e acção docente*. Porto: Porto, 2009. p. 143-164.

FORMOSINHO, João; MACHADO, Joaquim. Desempenho, mérito e desenvolvimento. Para uma avaliação mais profissional dos professores. In: FORMOSINHO, João; MACHADO, Joaquim; OLIVEIRA-FORMOSINHO, Júlia. *Formação, desempenho e avaliação de professores*. Mangualde: Pedago, 2010. p. 97-118.

GOMES, Marineide Oliveira. *Formação de professores na educação infantil*. 2. ed. São Paulo: Cortez, 2013.

JENSEN, Jytte J. Placement supervision of pedagogue students in Denmark: the role of university colleges and early childhood centres. *Early Years: An International Research Journal*, Abingdon, v. 35, n. 2, p. 154-167, Apr. 2015.

MARTINS, Carla. *Manual de análise de dados quantitativos com recurso ao IBM SPSS*. Braga: Psiquilíbrios, 2011.

MOSS, Peter; DAHLBERG, Gunilla. Beyond quality in early childhood education and care: languages of evaluation. *New Zealand Journal of Teachers' Work*, Auckland, v. 5, n. 1, p. 3-12, 2008.

OBERTHUEMER, Pamela. Editorial – Seeking new cultures of cooperation: a cross-national analysis of workplace-based learning and mentoring practices in early years professional education/training. *Early Years: An International Research Journal*, Abingdon, v. 35, n. 2, p. 115-123, Apr. 2015.

OLIVEIRA-FORMOSINHO, Júlia. *Modelos curriculares para a educação de infância*. 3. ed. Porto: Porto, 2007.

PORTUGAL. *Decreto-lei n. 241/2001, de 30 de agosto*. Aprova os perfis específicos de desempenho profissional do educador de infância e do professor do 1º ciclo do ensino básico. Lisboa: ME, 2001.

PORTUGAL. Decreto-lei n. 43/2007, de 22 de fevereiro. Aprova o regime jurídico da habilitação profissional para a docência na educação pré-escolar e nos ensinos básico e secundário. Lisboa: ME, 2007.

PORTUGAL. Decreto-lei n. 79/2014, de 14 de maio. Aprova o regime jurídico da habilitação profissional para a docência na educação pré-escolar e nos ensinos básico e secundário. Lisboa: MEC, 2014.

WEI, Ruth Chung; PECHEONE, Raymond. Assessment for learning in preservice teacher education: performance-based assessments. In: KENNEDY, Mary (Ed.). *Teacher assessment and the quest for teacher quality: a handbook*. San Francisco: Jossey Bass, 2010. p. 60-103.

WIGGINS, Grant. A true test: toward more authentic and equitable assessment. *Phi Delta Kappan*, Thousand Oaks, CA, v. 70, n. 9, p. 703-713, 1989.

SARA BARROS ARAÚJO

Professora-adjunta do Instituto Politécnico do Porto (IPP),
Porto, Portugal
saraujo@ese.ipp.pt

ANA PEREIRA ANTUNES

Professora auxiliar da Universidade da Madeira (UMa),
Funchal, Portugal
aantunes@uma.pt

Recebido em: 10 DEZEMBRO 2016

Aprovado para publicação em: 16 AGOSTO 2017

DADOS AUSENTES EM AVALIAÇÕES EDUCACIONAIS: COMPARAÇÃO DE MÉTODOS DE TRATAMENTO

**LUIS GUSTAVO DO AMARAL VINHA
JACOB ARIE LAROS**

RESUMO

Os dados ausentes são comuns nas avaliações educacionais. Por isso, o uso de métodos adequados torna-se fundamental para reduzir o impacto da perda de informação. O objetivo deste estudo é comparar o desempenho de quatro métodos de tratamentos de dados ausentes (imputação pela média, listwise deletion, máxima verossimilhança e imputação múltipla), tendo como base o uso de modelos de regressão aplicados aos dados da avaliação educacional realizada no estado do Ceará. Foram utilizadas informações de 7.000 estudantes, simulando-se diversos cenários de acordo com o percentual e o tipo de ausência. A imputação pela média apresentou o pior desempenho em todos os cenários simulados e os demais métodos mostraram resultados semelhantes entre si. Verificou-se ainda que o uso de variáveis auxiliares na estimação por máxima verossimilhança e imputação múltipla reduziu o viés das estimativas de parâmetros importantes do modelo quando a ausência simulada não é ao acaso.

PALAVRAS-CHAVE TRATAMENTO DE DADOS AUSENTES • AVALIAÇÃO DA EDUCAÇÃO • DESEMPENHO ACADÊMICO • SIMULAÇÃO.

DATOS AUSENTES EN EVALUACIONES EDUCACIONALES: COMPARACIÓN DE MÉTODOS DE TRATAMIENTO

RESUMEN

Los datos ausentes son comunes en las evaluaciones educacionales. Por ello el uso de métodos adecuados se hace fundamental para reducir el impacto de la pérdida de información. El objetivo de este estudio es comparar el desempeño de cuatro métodos de tratamientos de datos ausentes (imputación por el promedio, listwise deletion, máxima verosimilitud e imputación múltiple), en base al uso de modelos de regresión aplicados a los datos de la evaluación educacional realizada en el estado de Ceará. Se utilizaron informaciones de 7.000 estudiantes, simulando diversos escenarios de acuerdo al porcentual y al tipo de ausencia. La imputación por el promedio presentó el peor desempeño en todos los escenarios simulados y los demás métodos mostraron resultados similares entre ellos. También se verificó que el uso de variables auxiliares en la estimación por máxima verosimilitud e imputación múltiple redujo el sesgo de las estimaciones de parámetros importantes del modelo cuando la ausencia simulada no se debe a la casualidad.

PALABRAS CLAVE TRATAMIENTO DE DATOS AUSENTES • EVALUACIÓN DE LA EDUCACIÓN • DESEMPEÑO ACADÉMICO • SIMULACIÓN.

MISSING DATA IN EDUCATIONAL ASSESSMENT: A COMPARISON OF DATA TREATMENT METHODS

ABSTRACT

Missing data are common in educational assessments. Using the appropriate methods has, therefore, become essential to reduce the impact of the loss of information. The present study aims to compare the performance of four methods for dealing with missing data (mean imputation, listwise deletion, maximum likelihood and multiple imputation), all based on regression models applied to the educational assessment of data collected in the State of Ceará. Information about 7,000 students was used, simulating various scenarios according to the percentage and the type of the missing data. The mean imputation method showed the worst performance in all simulated scenarios and the other methods showed similar results among themselves. Moreover, the use of auxiliary variables in the estimation by maximum likelihood and multiple imputation proved to reduce the bias of estimates of some important parameters of the model, when the simulated missing data is not random.

KEYWORDS TREATMENT OF MISSING DATA • EDUCATION ASSESSMENT • ACADEMIC PERFORMANCE • SIMULATION.

INTRODUÇÃO

Apesar de a literatura estatística já abordar o tratamento de dados ausentes há décadas, o assunto ainda é um enigma na pesquisa social aplicada. Muitos pesquisadores não utilizam as técnicas adequadas por falta de familiaridade e, em geral, empregam métodos simples de eliminação ou substituição, que na maioria das situações não são apropriados (COX et al., 2014; MCKNIGHT et al., 2007; PEUGH; ENDERS, 2004). Trata-se, sem dúvida, de um assunto delicado, pois a ausência da informação pode ser causada por diversos fatores, apresentar diferentes padrões e distorcer os resultados da pesquisa, uma vez que a maioria das técnicas estatísticas foi desenvolvida para dados completos.

A preocupação com o tratamento de dados ausentes na pesquisa social pode ser observada em algumas revisões de literatura. Peugh e Enders (2004) revisaram estudos publicados nas áreas de educação e psicologia aplicada com o objetivo de inventariar como os pesquisadores dessas áreas reportavam a ausência de informação e os procedimentos usados nas análises. Os autores verificaram que a presença

e o tratamento de dados ausentes eram raramente reportados. Na maioria dos casos, os valores ausentes tinham que ser identificados pela comparação do tamanho da amostra e o número de graus de liberdade nas diversas análises realizadas. Além disso, em quase todos os trabalhos em que foi identificada a presença de valores ausentes, foram utilizados métodos tradicionais de eliminação ou substituição. Segundo os autores, entre 1999 e 2003, aumentou o número de estudos em que os valores ausentes foram reportados, entretanto, técnicas mais sofisticadas baseadas na estimação por máxima verossimilhança ou imputação múltipla foram pouco utilizadas.

Resultados semelhantes foram encontrados por Rousseau *et al.* (2012), que apresentam uma revisão rigorosa de artigos publicados no *British Educational Research Journal*, no período de 2003 a 2007. Dos 68 estudos selecionados, mais de um terço não indicava qualquer informação sobre dados ausentes. Na metade dos trabalhos em que a presença de valores ausentes foi reportada, o tratamento adotado não estava explícito e, entre os que mencionaram o tratamento, a maioria utilizou métodos de eliminação de observações.

Discutindo também quais seriam os motivos que levam a tamanha variação nos relatos dos estudos analisados, Rousseau *et al.* (2012) acreditam que muitos pesquisadores não mencionam os valores ausentes por ignorar a natureza desses dados, como eles influenciam os resultados, quais métodos estão disponíveis para análise e como reportá-los. Os autores afirmam ainda que não existe na área de educação um procedimento que possa ser usado pelos pesquisadores – como acontece na psicologia com os manuais da APA (*American Psychology Association*) – e que a literatura estatística relacionada ao tema é muito técnica para esse público. Eles ressaltam também que os pesquisadores não mencionam a existência de valores ausentes, ou reportam resumidamente, em função do aumento na complexidade das interpretações e dos efeitos negativos nos resultados gerados pela falta de parte da informação.

No Brasil, considerando os estudos em avaliação educacional, verifica-se também que poucos pesquisadores mencionam a presença de valores ausentes nos dados analisados.

Pode-se destacar o estudo de Oliveira, Belluzzo e Pazello (2013), que indicam a retirada das observações com dados faltantes e admitem que esse procedimento pode comprometer os resultados observados. Em Xerxenevsky (2012), as observações incompletas também são excluídas, mas nesse caso a autora afirma que isso não impactaria nos resultados, em função do tipo de dado ausente. Em alguns trabalhos foram utilizadas variáveis indicadoras no ajuste dos modelos como uma forma de reduzir o impacto das informações ausentes (MACEDO, 2004; RODRIGUES; RIOS-NETO; PINTO, 2011; SOARES; ALVES, 2003).

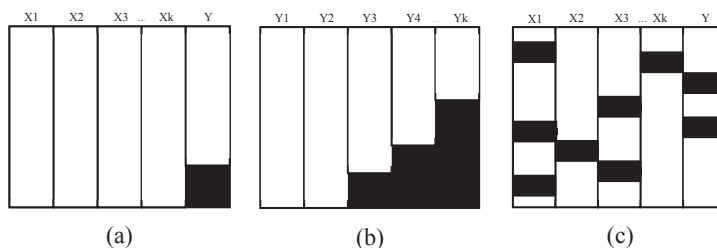
O presente estudo visa a contribuir para o melhor entendimento do impacto dos valores ausentes nos resultados de avaliações educacionais. Para isso, além de discutir conceitos relacionados e apresentar alguns métodos de tratamento e análise de dados com informações faltantes, o artigo procura comparar o desempenho de diferentes procedimentos de análise em diferentes situações, tendo como base dados reais oriundos de uma avaliação educacional de larga escala realizada no Brasil. Espera-se que este trabalho seja útil para o melhor entendimento do problema, mesmo para leitores que não estão familiarizados com análise de dados e conceitos estatísticos.

TIPOS DE DADOS AUSENTES

Algumas classificações relacionadas a dados ausentes são encontradas na literatura. Por exemplo, McKnight *et al.* (2007) utilizam uma classificação de ausência de acordo com a fonte: casos ausentes, variáveis ausentes ou ocasiões ausentes. Os casos ausentes acontecem quando o participante não fornece qualquer informação, o que ocorre com a ausência do indivíduo selecionado no momento da coleta de dados. Variáveis ausentes são observadas quando o participante não fornece parte da informação requerida, por exemplo, quando ele não responde um ou mais itens do questionário. Ocasiões ausentes são comuns em estudos longitudinais e ocorrem quando o participante não está presente em todos os momentos de coleta de dados.

Os dados ausentes podem também ser classificados de acordo com o padrão de não resposta (SCHAFER; GRAHAM, 2002). No padrão univariado apenas uma variável apresenta valores ausentes. Por exemplo, em um modelo de regressão, a ausência ocorre somente na variável dependente (Figura 1a). O padrão monótono é geralmente observado em estudos longitudinais em decorrência do abandono de participantes ao longo das avaliações (Figura 1b). Nesse caso, supondo-se que Y seja a característica acompanhada ao longo do tempo, os indivíduos com valor ausente em Y_t também apresentarão valores ausentes em Y_{t+1} , Y_{t+2} e assim por diante, para qualquer t . Por fim, como mostrado na Figura 1c, no padrão arbitrário os valores ausentes podem ocorrer em uma ou mais variáveis para qualquer observação.

FIGURA 1 - Padrões de não resposta: (a) univariado, (b) monótono e (c) arbitrário



Fonte: Adaptado de Schafer e Graham (2002).

Entre as diversas classificações encontradas na literatura, aquela proposta por Rubin (1976, 1987) é a mais importante. Segundo o autor, os valores ausentes são gerados por três mecanismos distintos que relacionam a propensão de ausência aos dados observados: valores ausentes completamente ao acaso (MCAR, do inglês *Missing Completely at Random*); ausentes ao acaso (MAR, *Missing at Random*); e ausentes não ao acaso (MNAR, *Missing not at Random*).

Os dados são classificados como ausentes completamente ao acaso (MCAR) quando a ocorrência não está relacionada a qualquer variável observada no estudo ou à própria variável que apresenta os valores faltantes. Dessa forma, o mecanismo gerador desses dados não está relacionado a qualquer

característica observada; daí a denominação de ocorrência completamente ao acaso. Esse tipo de ausência poderia ser interpretado como a retirada de uma amostra aleatória de observações do banco de dados completo. Por exemplo, valores faltantes na proficiência em leitura em uma avaliação educacional podem ser consequência de eventos diversos, como adoecimento do estudante.

Os dados faltantes são classificados como MAR quando a ocorrência está relacionada aos valores observados de outras variáveis, mas independe do valor da variável em questão. Por exemplo, pode-se supor que alunos com baixa renda familiar apresentam maiores taxas de ausência nas provas, no entanto, para uma determinada faixa de renda familiar, não é observada qualquer relação entre a ausência e o desempenho (PEUGH; ENDERS, 2004). Nesse caso, o nome dado ao mecanismo pode gerar dúvida (COLLINS; SCHAFER; KAM, 2001; GRAHAM, 2009), uma vez que a ausência de dados não é de fato ao acaso. Pode-se, então, interpretar da seguinte forma: a taxa de valores ausentes está relacionada à renda, mas, quando controlada a faixa de renda, a ausência passa a ser aleatória (a incidência é aleatória depois do controle de uma variável observada). Na prática não é possível confirmar se a ocorrência de valores ausentes está apenas relacionada com as outras variáveis e não com os valores da própria variável, pois não se conhecem os valores faltantes. No exemplo, entre os indivíduos de baixa renda, não é possível verificar se os valores ausentes são dos alunos com menor proficiência (ENDERS, 2010).

Por fim, os dados faltantes são identificados como MNAR quando a ocorrência está relacionada aos valores da própria variável analisada. Esses dados são mais difíceis de serem identificados, já que a ocorrência está relacionada a valores não observados. Na avaliação educacional, esse fato pode ser exemplificado pela maior taxa de valores faltantes em testes de proficiência entre os alunos com menor desempenho (mesmo depois de controladas outras variáveis).

A classificação proposta por Rubin está diretamente relacionada ao impacto da ausência de informação e à escolha da abordagem mais apropriada para análise dos dados. Os que menos influenciam as análises e os resultados são os

dados ausentes do tipo MCAR, uma vez que a amostra de valores completos pode ser vista como representativa da população. Quando os dados faltantes são do tipo MAR, a ausência pode ser considerada ignorável, pois se faz necessária a modelagem adicional do mecanismo de ausência de dados no processo de estimação. Por fim, os dados MNAR também são chamados de não ignoráveis, dado que o mecanismo gerador de ausência deve ser modelado para que sejam obtidas boas estimativas dos parâmetros de interesse (ALLISON, 2002).

MÉTODOS DE TRATAMENTO E ANÁLISE DE DADOS INCOMPLETOS

Nesta seção, são apresentados métodos tradicionais baseados na eliminação e substituição de valores, além de dois procedimentos mais sofisticados baseados em métodos de estimação.

MÉTODOS DE ELIMINAÇÃO E IMPUTAÇÃO SIMPLES

De acordo com as revisões de literatura mencionadas anteriormente, os procedimentos baseados na *eliminação de observações* são utilizados na maior parte dos estudos em que os dados ausentes são identificados (CHEEMA, 2014; PEUGH; ENDERS, 2004; ROUSSEAU et al., 2012). Acredita-se também que muitos pesquisadores acabam utilizando esses métodos inconscientemente, por ser o procedimento padrão dos pacotes estatísticos (ROUSSEAU et al., 2012). Os procedimentos baseados na eliminação de observações são *listwise deletion* (ou *complete-case analysis*) e *pairwise deletion* (*available case analysis*).

LISTWISE DELETION

Neste procedimento, todos os casos com um ou mais valores ausentes nas variáveis observadas são retirados; logo, consideram-se apenas as observações completas. Como consequência, o poder dos testes estatísticos se reduz com a perda de informação e pode gerar estimativas viesadas dos parâmetros quando os dados não são MCAR.

Entretanto, é importante ressaltar que esse método apresenta vantagens que o tornam atrativo em algumas situações. Trata-se de um método simples que não gera maior

complexidade nas análises e interpretação de resultados e, como mencionado anteriormente, é o procedimento padrão dos pacotes estatísticos. Em algumas situações o *listwise deletion* mostra desempenho semelhante a técnicas mais sofisticadas, como, por exemplo, quando os dados faltantes são do tipo MCAR, o tamanho da amostra é grande e o número de faltantes é relativamente pequeno (ENDERS, 2010). Ainda, no ajuste de modelos de regressão, quando os valores ausentes dependem de uma variável presente no modelo (tipo MAR), o procedimento tem bom desempenho (ALLISON, 2002; SCHAFER; GRAHAM, 2002).

PAIRWISE DELETION

A utilização do procedimento *listwise deletion* gera a perda de uma parcela considerável da informação contida no banco de dados, principalmente quando o número de variáveis envolvidas no estudo aumenta. O procedimento *pairwise deletion* surge então como alternativa para reduzir essa perda de informação.

Nesse caso, a parcela de dados utilizada nos cálculos é maior, uma vez que são consideradas observações completas para pares de variáveis e não todo o conjunto de informações de um indivíduo. Esse procedimento é muito comum quando são empregadas técnicas baseadas em correlações, dado que a estimação da correlação depende apenas de um par de valores. No entanto, tal abordagem é criticada porque as amostras utilizadas nos diferentes cálculos não são as mesmas, o que pode gerar problemas de estimação e inconsistências de resultados. Ainda, esse procedimento também pode gerar viés se os dados ausentes não são MCAR (ALLISON, 2002).

As *técnicas de imputação simples* podem ser mais atrativas do que os métodos de eliminação, uma vez que não existe descarte de informação, mas são muito criticadas. Com a utilização de técnicas de imputação simples, os valores imputados são tratados como conhecidos nas análises, o que pode distorcer os resultados. Além disso, alguns pesquisadores evitam o uso da imputação por acreditar que estariam “construindo os dados” (GRAHAM, 2009). A seguir

são apresentados os principais métodos de imputação simples encontrados na literatura.

IMPUTAÇÃO PELA MÉDIA

Essa técnica consiste na substituição dos valores ausentes pela média da variável, obtida a partir dos valores válidos observados na amostra. Segundo Enders (2010), trata-se de um procedimento antigo, cuja autoria é frequentemente atribuída a Wilks (1932). A imputação pela média é ainda muito empregada, mas produz distorções mesmo quando os dados são MCAR. Com a utilização desse procedimento, a variabilidade dos dados pode ser subestimada e, como consequência, há um aumento da chance de se rejeitarem hipóteses nulas em testes de significância, além de afetar a estimação de medidas de associação.

O método de imputação pela média é possivelmente o pior tratamento de dados ausentes (ENDERS, 2010) e fornece estimadores viesados para todos os parâmetros, com exceção da média, para qualquer que seja o tipo de dado ausente analisado.

IMPUTAÇÃO PELA REGRESSÃO

A imputação dos dados pela regressão foi proposta por Buck (1960) e, em geral, oferece melhores resultados do que a imputação pela média. Nesse caso, o valor ausente é substituído pelo valor obtido pela equação de regressão utilizando as demais variáveis que apresentam valores completos como preditoras. Esse método pode ser indicado quando poucas variáveis possuem observações faltantes. Já nas situações em que muitas variáveis têm dados ausentes, tal técnica perde a atratividade, pois o número de equações a serem estimadas é grande (ENDERS, 2010).

Como a substituição é feita por um valor predito baseado nas relações entre as variáveis, o valor substituído está exatamente na linha que descreve a relação, o que pode resultar na superestimação das covariâncias entre as variáveis (PEUGH; ENDERS, 2004). Por outro lado, a subestimação da variabilidade é menos acentuada do que a observada com a utilização da imputação pela média e o procedimento gera

estimativas não viesadas quando os dados são do tipo MAR.

Para atenuar os possíveis problemas decorrentes desse método, pode-se utilizar substituição pela regressão estocástica (*stochastic regression imputation*). Nesse caso o valor a ser substituído é composto pelo valor predito pela regressão mais uma parcela aleatória, proveniente da distribuição dos resíduos do modelo. Esse método também pode ser encontrado na literatura como *single random imputation* (CHEEMA, 2014).

HOT DECK IMPUTATION

Desenvolvido pelo *US Census Bureau* (ENDERS, 2010), o procedimento *hot deck imputation* propõe a substituição do valor ausente por um valor observado em unidades similares (ANDRIDGE; LITTLE, 2010). Para uma observação com ausência de informação são identificados os possíveis doadores, observações que apresentam valores válidos para a variável a ser tratada e com respostas semelhantes para as demais variáveis coletadas. Por exemplo, o valor observado de apenas um doador selecionado aleatoriamente pode ser usado nessa substituição, ou pode-se utilizar o dado do doador mais próximo, de acordo com alguma métrica preestabelecida. Em outras versões do método, os valores referentes a um conjunto de doadores são combinados, por exemplo, pela média aritmética, sendo que o resultado é então usado na substituição.

Uma característica importante desse método refere-se ao fato de a substituição ser feita sem a suposição de modelo. Porém, como Andridge e Little (2010) ressaltam, o procedimento depende da medida usada e das variáveis envolvidas na escolha de doadores. O método apresenta bons resultados quando os dados são do tipo MCAR e o número de observações é elevado.

VARIÁVEIS INDICADORAS

O uso de variáveis indicadoras de ausência no ajuste dos modelos foi proposto por Cohen e Cohen (1985). Esse método tem sido empregado pelos pesquisadores da área da educação (COX et al., 2014) e também pode ser encontrado na

literatura brasileira em avaliação educacional (MACEDO, 2004; RODRIGUES; RIOS-NETO; PINTO, 2011; SOARES; ALVES, 2003). O procedimento visa a reduzir a perda de informação e consiste na imputação dos valores faltantes pela média e inclusão de variáveis que indicam a ausência de informação. Com essa reparametrização um novo intercepto para categorias de dados ausentes é criado e, sob a hipótese de que os dados são MCAR, as estimativas dos parâmetros de interesse não são afetadas (RODRIGUES; RIOS-NETO; PINTO, 2011). Contudo, quando os valores ausentes não são MCAR, o procedimento pode gerar viés nas estimativas dos parâmetros de interesse (COX et al., 2014).

MÉTODOS DE ESTIMAÇÃO E IMPUTAÇÃO MÚLTIPLA

As abordagens que utilizam métodos de estimação para o tratamento de dados incompletos são apresentadas a seguir. Segundo Peugh e Enders (2004), diversos estudos têm apontado a superioridade desses métodos em relação aos tradicionais apresentados anteriormente. Ainda, Graham (2009) afirma que os pesquisadores deveriam empregar os procedimentos baseados na máxima verossimilhança e na imputação múltipla, pois são os melhores métodos disponíveis e baseiam-se em conceitos fortes e tradicionais da estatística.

MÁXIMA VEROSSIMILHANÇA

O tratamento dos valores ausentes pelo método da máxima verossimilhança é semelhante ao utilizado para dados completos (ENDERS, 2010). O método consiste na escolha de valores para os parâmetros do modelo que maximizam a função de verossimilhança, a qual expressa a probabilidade de se observarem os valores obtidos na amostra para um determinado modelo escolhido. Esse procedimento requer a suposição de que a ausência é do tipo MAR, além de uma suposição relacionada à distribuição. Em geral assume-se que os dados têm distribuição normal multivariada.

A estimação dos parâmetros é realizada através da maximização da função de verossimilhança, o que, na maior parte das situações, não pode ser realizado analiticamente.

Faz-se necessário o uso de métodos numéricos, como, por exemplo, o algoritmo EM. Esse método de maximização é muito popular quando os dados considerados na estimação não estão completos (ALLISON, 2002).

O algoritmo EM é um procedimento iterativo que consiste na repetição de dois passos: estimação (E) e maximização (M). O processo se inicia com uma estimação do vetor de médias e da matriz de covariâncias, utilizando apenas os dados completos. No passo E os elementos do vetor de médias e da matriz de covariâncias são utilizados para construir um conjunto de equações de regressão usadas para estimar os valores ausentes com base nas variáveis observadas. Em seguida, no passo M, o vetor de médias e a matriz de covariâncias são reestimados por meio dos estimadores de máxima verossimilhança com base em dados completos (valores observados e os substituídos no passo anterior). Em uma nova etapa E, as equações de regressão são reestimadas e os valores ausentes substituídos por novos valores. Na etapa M seguinte, esse novo banco de dados é usado para reestimar o vetor de médias e a matriz de covariâncias, sendo que o processo é repetido até que as estimativas de médias e variâncias não mudem mais (usando um critério de convergência preestabelecido).

O algoritmo EM apresenta algumas vantagens importantes que o tornam muito atrativo (ALLISON, 2002). Trata-se de um procedimento de fácil implementação, não requer a definição de muitos parâmetros e está disponível em diversos pacotes estatísticos. Entretanto, considerando o ajuste de modelos de regressão, os erros padrão associados aos coeficientes não são obtidos diretamente (ENDERS, 2010). Como alternativa, pode-se utilizar o procedimento FIML (*full information maximum likelihood*) que, assim como o algoritmo EM, gera estimativas não viesadas para os coeficientes da regressão e ainda estima diretamente os erros associados (ENDERS, 2001a).

A estimação pelo método da máxima verossimilhança possibilita a utilização de variáveis auxiliares, que não fazem parte do modelo principal de análise, mas estão associadas ao mecanismo gerador de valores ausentes ou à variável a ser

tratada. A inclusão de variáveis auxiliares aumenta a chance de satisfazer a suposição de ausência do tipo MAR, e assim melhorar a estimação (BARALDI; ENDERS, 2010; COLLINS; SCHAFER; KAM, 2001). É importante ressaltar que a inclusão de variáveis auxiliares de forma adequada não altera a interpretação dos parâmetros do modelo principal.

De forma geral, a estimação por máxima verossimilhança apresenta diversas propriedades interessantes no tratamento de dados ausentes, entretanto, não se trata de um método infalível. Entre os problemas relacionados, pode-se destacar que a suposição de distribuição normal multivariada geralmente não está satisfeita, o que pode gerar viés nas estimativas e erros associados. Nas situações em que a ausência não é ao acaso (MNAR), esse procedimento pode distorcer os resultados, apesar de apresentar desempenho superior aos métodos tradicionais, principalmente quando são utilizadas variáveis auxiliares (GRAHAM, 2009).

IMPUTAÇÃO MÚLTIPLA

Na imputação múltipla, proposta por Rubin (1987), os valores ausentes são repetidamente substituídos por valores obtidos a partir da simulação da distribuição condicional de probabilidade, tendo como resultado múltiplas versões do banco de dados. Cada versão do banco de dados é analisada de acordo com as técnicas usuais e os resultados são combinados gerando estimativas pontuais dos parâmetros de interesse (ROSE; FRASER, 2008).

Esse procedimento é composto por três etapas: imputação, análise e combinação. Na etapa de imputação, composta por dois passos, são gerados m novos bancos de dados. No primeiro passo (passo I), o vetor de médias e a matriz de covariâncias são estimados e é construído um sistema de equações de regressão para imputação dos valores ausentes, como no método de substituição pela regressão estocástica. No segundo passo (passo II), o vetor de média e a matriz de covariâncias são estimados novamente, e a partir dessas estimativas são geradas novas estimativas com a adição de um termo aleatório (esses novos valores correspondem à retirada de uma amostra da distribuição *a posteriori* da matriz

de covariâncias e do vetor de médias). Essas novas estimativas de médias e covariâncias são usadas no passo I seguinte, sendo que o processo se repete até que os m conjuntos de dados completos sejam criados.

Algumas considerações importantes são necessárias nessa etapa. Primeiro, a decisão de quais variáveis serão incluídas na etapa de imputação é um aspecto relevante para o bom desempenho da imputação múltipla (ENDERS, 2010). As variáveis presentes no modelo principal devem sempre ser incluídas. Variáveis auxiliares podem ser consideradas nessa etapa, sem correr o risco de introduzir viés nos resultados, cuidando apenas para não ser introduzido um número muito elevado de variáveis, o que poderia causar problemas de convergência. Segundo, deve-se considerar um número de iterações antes da retirada do primeiro banco de dados e entre as retiradas dos outros bancos de dados. As iterações iniciais são necessárias para estabilização da distribuição dos parâmetros e as iterações entre as retiradas asseguram a independência entre os dados gerados. Por fim, a eficiência do procedimento está relacionada com o número de bancos de dados gerados (m) e, de forma geral, quanto maior o percentual de valores ausentes, maior deve ser m (para mais detalhes ver GRAHAM; OLCHOWSKI; GILREATH, 2007).

A etapa de análise consiste na aplicação das técnicas estatísticas usuais aos m conjuntos de dados gerados na etapa anterior, de acordo com os objetivos da pesquisa. Como resultado dessa etapa, m conjuntos de estimativas para os parâmetros de interesse são gerados. Por fim, na última etapa são calculadas as estimativas dos parâmetros de interesse a partir da combinação das m estimativas. A média aritmética das m estimativas pode ser usada para gerar a estimativa combinada de um parâmetro, porém essa combinação será válida quando a distribuição do estimador se aproxima da normal. Alguns estimadores têm distribuições assimétricas, como os estimadores de variância e covariâncias, e, nesses casos, transformações podem ser utilizadas para melhorar as estimativas (ENDERS, 2010).

Os procedimentos de imputação múltipla e máxima verossimilhança são baseados nas mesmas suposições de distribuição normal multivariada e mecanismo gerador de dados

ausentes do tipo MAR. Em geral, os resultados obtidos por meio desses métodos são semelhantes, especialmente quando o tamanho da amostra é grande. Segundo Collins, Schafer e Kam (2001), os procedimentos apresentam resultados muito próximos quando é usado o mesmo banco de dados, com as mesmas suposições de relações entre as variáveis e suas distribuições.

ESTUDO COMPARATIVO

O estudo apresentado neste manuscrito visa a comparar métodos de tratamento de dados ausentes, tendo como base um conjunto de dados reais de uma avaliação educacional. Foram selecionados para essa comparação dois procedimentos tradicionais e que ainda são muito utilizados pelos pesquisadores (PEUGH; ENDERS, 2004) – a imputação simples pela média (Me) e o *listwise deletion* (LD) – e os métodos baseados na estimação por máxima verossimilhança (MV) e na imputação múltipla (IM). Também foi avaliada a utilização de variáveis auxiliares nos procedimentos MV e IM. Este estudo baseia-se na análise dos dados por meio de modelos de regressão linear múltipla, tendo como variável resposta o desempenho escolar.

DADOS

Os dados trabalhados são provenientes do Sistema Permanente de Avaliação da Educação Básica (Spaee), disponibilizados pela Secretaria de Educação do Estado do Ceará. Entre outros levantamentos, esse sistema monitora os alunos do ensino médio da rede pública cearense, tendo como medidas de desempenho escolar as proficiências em Língua Portuguesa e Matemática. As proficiências são estimadas por meio da Teoria da Resposta ao Item (TRI) (CEARÁ, 2011). Além dos testes, os alunos respondem a questionários contextuais com itens relacionados a dados socioeconômicos, hábitos de estudo e clima em sala de aula.¹

O banco de dados original disponibilizado contém informações relativas a centenas de milhares de estudantes, com grande incidência de valores ausentes. Entretanto, para

¹ O questionário contextual pode ser acessado em <<http://www.spaee.caedufjf.net/downloads/2009-2/>>.

a simulação realizada neste estudo, foi utilizada uma amostra composta apenas por estudantes com dados completos para as variáveis utilizadas (Quadro 1). Essa amostra contém informação referente a 7.000 estudantes matriculados no 1º ano do ensino médio em 2009 e que estavam presentes na avaliação em 2010. Esses estudantes selecionados foram avaliados por meio das provas de proficiência nos dois anos e responderam ao questionário no primeiro ano. As variáveis utilizadas no estudo são apresentadas no Quadro 1.

QUADRO 1 - Variáveis utilizadas no estudo

VARIÁVEIS	DESCRIÇÃO / CODIFICAÇÃO
MAT10	Desempenho em Matemática em 2010 - 2º ano do ensino médio / Variável contínua.
MAT09	Desempenho em Matemática em 2009 - 1º ano do ensino médio / Variável contínua.
LPO9	Desempenho em Língua Portuguesa em 2009 - 1º ano do ensino médio/Variável contínua.
MASC	Sexo do aluno / Variável binária: 0, se feminino; 1, se masculino.
SUP	Pretensão de ingresso no ensino superior. Obtida a partir da questão relativa aos planos dos alunos após a conclusão do ensino médio / Variável binária: 1, se pretende; e 0, se tem outros planos.
REPETÊNCIA	Números de vezes que o aluno repetiu um ano escolar, avaliado em 2009 / Variável ordinal: 0, se nunca repetiu; 1, uma repetência; 2, duas repetências; 3, três ou mais repetências.
IDADE	Idade reportada em 2009 (em anos).
MANHÃ	Turno em que o aluno frequenta as aulas. Variável binária que assume o valor 1 se o aluno estudava de manhã em 2009 e 0 se estudava à tarde ou à noite.

Fonte: Secretaria de Educação do Estado do Ceará/Spaeece (elaboração dos autores).

Entre os 7.000 estudantes selecionados, 45,3% são do sexo masculino, 43,8% pretendiam ingressar no ensino superior, 32,6% reportaram uma ou mais repetências e 42,0% estudavam no período matutino em 2009. A idade média observada foi de 15,7 anos com desvio-padrão de 1,1 anos e as notas médias em Matemática foram 246,5 e 257,1 pontos em 2009 e 2010, respectivamente, com desvios-padrão de 46,1 e 46,9 pontos. Já a nota média em Língua Portuguesa em 2009 foi de 247,5 pontos com desvio-padrão de 39,9 pontos.

PROCEDIMENTOS

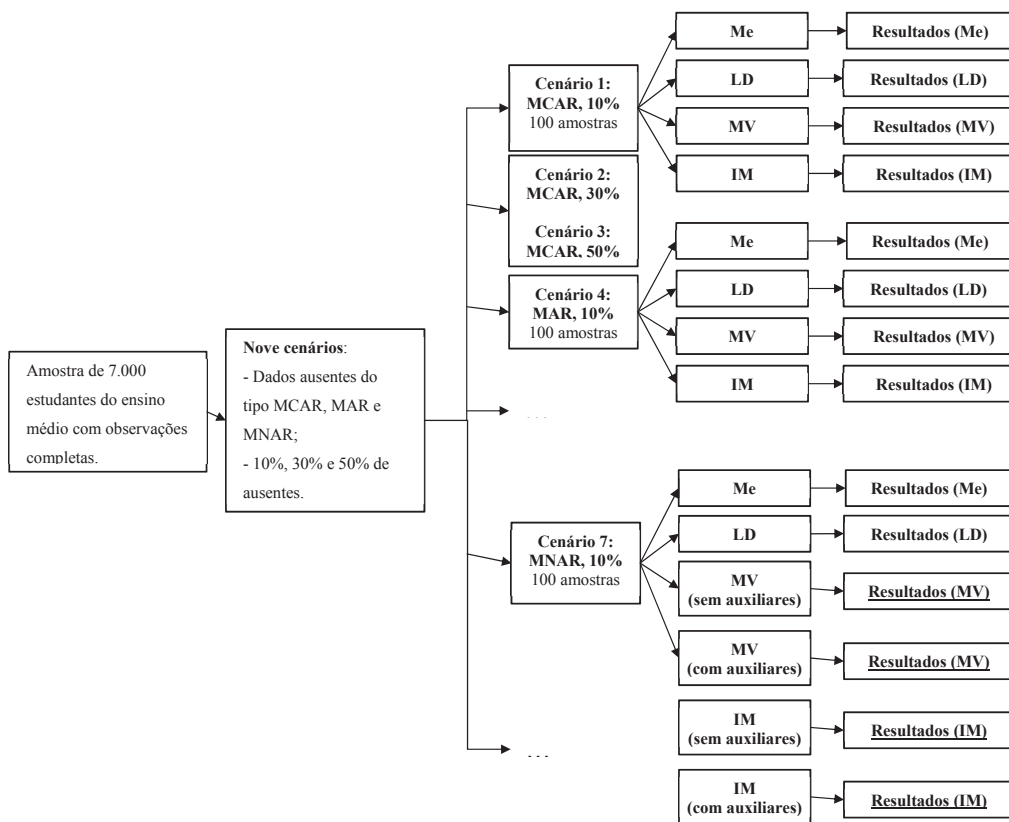
A comparação dos métodos foi realizada supondo-se que essa avaliação seja usada para comparar o desempenho em Matemática no segundo ano do ensino médio, por sexo e intenção de ingresso no ensino superior. Dessa forma, o modelo principal de análise é dado por:

$$\text{MAT10} = \beta_0 + \beta_1\text{MAT09} + \beta_2\text{MASC} + \beta_3\text{SUP} + \varepsilon \quad (1)$$

A variável MAT09 foi incorporada ao modelo como uma variável de controle para que sejam consideradas as diferenças entre os alunos por meio de uma medida inicial de desempenho. Dessa forma, os coeficientes β_2 e β_3 representam as diferenças entre as proficiências médias dos grupos, controladas pela proficiência no ano anterior.

A Figura 2 apresenta o plano utilizado para a simulação. Foram gerados nove cenários diferentes de acordo com o percentual (10%, 30% e 50%) e o tipo de ausência de dados (MCAR, MAR e MNAR). Os dados ausentes foram simulados apenas na variável resposta (MAT10), ou seja, padrão univariado. Para cada cenário foram geradas 100 amostras e cada amostra foi analisada por meio dos quatro procedimentos (Me, LD, MV e IM). Os métodos MV e IM foram implementados com e sem variáveis auxiliares para os cenários com dados ausentes do tipo MNAR. Com isso, foram obtidas as estimativas médias dos coeficientes da regressão referentes aos quatro procedimentos em cada cenário.

FIGURA 2 - Plano do estudo de simulação



Fonte: Dados da pesquisa (elaboração dos autores).

O primeiro mecanismo utilizado gera valores ausentes completamente ao acaso (MCAR). Foram sorteadas amostras aleatórias de indivíduos com valores completos. Para esses indivíduos o valor da variável resposta passa a ser ausente. No segundo mecanismo, os dados ausentes simulados são do tipo MAR e, nesse caso, a ausência estava relacionada à variável MAT09, de tal forma que a ocorrência de dados faltantes é maior entre os alunos com menor proficiência em Matemática no primeiro ano do ensino médio. Por fim, no terceiro mecanismo foram simulados valores ausentes (MNAR) cuja taxa depende de outras variáveis não presentes no modelo apresentado na equação (1) e, nesse caso, a ausência simulada está relacionada com as variáveis LP09, IDADE, REPETÊNCIA e MANHÃ (ver Quadro 1). Nesse último

cenário o percentual de valores ausentes é maior entre os alunos com menor proficiência em Língua Portuguesa, mais velhos, com maior número de repetência e que estudam no período vespertino ou noturno.

O terceiro mecanismo descrito acima pode ser considerado MNAR, dado que existe uma dependência residual entre a ausência e a variável que apresenta os valores ausentes (MAT10). Essa dependência deve-se à relação existente entre LP09, IDADE, REPETÊNCIA e MANHÃ e a variável MAT10, considerando as demais variáveis do modelo. Tais variáveis foram escolhidas por estarem relacionadas com a probabilidade de ausência dos alunos no segundo ano do ensino médio (VINHA, 2016).

ANÁLISE DE DADOS

As análises foram realizadas utilizando o *software* estatístico SAS (*Statistical Analysis System*), versão 9.4. Algumas funções específicas foram usadas:

- para a análise dos dados completos e para o procedimento LD foi empregada apenas a PROC REG. Essa função tem como padrão o procedimento *listwise deletion* no ajuste do modelo para dados incompletos. Os modelos foram estimados pelo método dos mínimos quadrados ordinários;
- a estimação por MV foi realizada através do algoritmo EM. A inclusão das variáveis auxiliares (LP09, IDADE, REPETÊNCIA e MANHÃ) foi feita pelo método de estimação em dois estágios,² sendo que, para isso, utilizaram-se a opção EM da PROC MI e a PROC REG;
- para o procedimento IM foi empregada a PROC MI com os seguintes parâmetros: $m = 50$ para os cenários com 10% e 30% de valores ausentes; e $m = 100$ para os cenários com 50% de ausentes, 500 iterações iniciais antes da retirada da primeira amostra e 200 iterações entre as amostras. A PROC MIANALYZE foi usada para gerar as estimativas finais a partir das m amostras.

² Esse método foi proposto por Savalei e Bentler (2009). No primeiro estágio são incorporadas todas as variáveis (de interesse no estudo e auxiliares) na estimação do vetor de médias e matriz de covariâncias pela máxima verossimilhança e, no segundo estágio, o modelo principal é ajustado considerando somente as variáveis de interesse no estudo.

RESULTADOS

As Tabelas 1, 2 e 3 apresentam os coeficientes estimados a partir dos dados completos e as estimativas médias relativas aos diferentes cenários de ausência de dados. Essas tabelas mostram também os desvios observados entre as estimativas médias e as estimativas obtidas com os dados completos, de acordo com a equação:

$$\text{Desvio (\%)} = \left(\frac{bi - bc}{bc} \right) \times 100\% \quad (2)$$

Onde *bi* e *bc* correspondem às estimativas obtidas para os dados incompletos e completos, respectivamente.

Considerando os resultados da amostra completa, pode-se observar que quanto maior é a proficiência em Matemática em 2009, maior tende a ser a proficiência em 2010 (o coeficiente estimado é de 0,677 pontos). Verifica-se que, controlado o desempenho anterior, os meninos têm nota média 3,017 pontos acima da nota média das meninas, e os que pretendem ingressar no ensino superior registram 6,995 pontos a mais, em média. Os resultados apresentados a seguir mostram as variações em relação a esses valores.

Primeiramente, observa-se que a imputação simples pela média mostra o pior desempenho, independentemente do mecanismo gerador e do percentual de ausência (Tabelas 1, 2 e 3). Mesmo quando a ausência é completamente ao acaso, constata-se, pela Tabela 1, que as estimativas obtidas por esse método são distantes dos valores encontrados com os dados completos.

Na Tabela 1 são apresentados os demais resultados para os dados ausentes do tipo MCAR. Com exceção da imputação pela média, os resultados são semelhantes, ou seja, as estimativas dos parâmetros do modelo de regressão mostram pequena variação em relação ao valor observado no ajuste para os dados completos (desvios de 4,1% ou menos). Verifica-se também que, mesmo com percentual elevado de valores ausentes, a variação observada nos coeficientes ainda é pequena. Em geral, o coeficiente relativo à variável MASC registra maiores desvios.

TABELA 1 – Estimativas dos coeficientes do modelo de regressão aplicado a dados completos e a dados com ausência de informação do tipo MCAR (simulada), acompanhados do desvio percentual, segundo o cenário de ausência e de tratamento

DADOS COMPLETOS		DADOS AUSENTES POR SIMULAÇÃO: TIPO MCAR			
		TRATAMENTO			
		MÉDIA	LD	MV	IM
10% AUSENTES					
Intercepto	85,85	102,79 (19,7)	85,90 (0,1)	85,88 (0,0)	85,86 (0,0)
MAT09	0,677	0,61 (-9,9)	0,68 (0,4)	0,68 (0,4)	0,68 (0,4)
MASC	3,017	2,71 (-10,2)	3,01 (-0,2)	3,02 (0,1)	3,00 (-0,6)
SUP	6,995	6,31 (-9,8)	7,02 (0,4)	7,02 (0,4)	7,01 (0,2)
30% AUSENTES					
Intercepto	85,85	136,51 (59,0)	85,92 (0,1)	85,85 (0,0)	85,72 (-0,2)
MAT09	0,677	0,48 (-29,1)	0,68 (0,4)	0,68 (0,4)	0,68 (0,4)
MASC	3,017	2,08 (-31,1)	2,99 (-0,9)	2,97 (-1,6)	2,94 (-2,6)
SUP	6,995	4,95 (-29,2)	7,00 (0,1)	7,01 (0,2)	7,03 (0,5)
50% AUSENTES					
Intercepto	85,85	170,53 (98,6)	85,89 (0,0)	85,88 (0,0)	85,62 (-0,3)
MAT09	0,677	0,34 (-49,8)	0,68 (0,4)	0,68 (0,4)	0,69 (1,9)
MASC	3,017	1,51 (-50,0)	3,14 (4,1)	3,14 (3,4)	3,12 (3,4)
SUP	6,995	3,53 (-49,5)	7,14 (2,1)	7,13 (1,9)	7,14 (2,1)

Fonte: Dados da pesquisa (elaboração dos autores).

Nota: Os valores em negrito do desvio percentual correspondem às estimativas com desvio maior que 10% em relação ao coeficiente da aplicação da regressão aos dados completos.

Pela Tabela 2 observa-se que os procedimentos LD, MV e IM têm desempenhos semelhantes quando os dados ausentes estão relacionados com a variável MAT09 presente no modelo. Em comparação com os cenários apresentados na Tabela 1, em geral, os dados ausentes do tipo MAR têm maior impacto na estimação dos coeficientes da regressão. Verifica-se que quanto maior é o percentual de ausência, maiores são os desvios observados, e os coeficientes mais afetados estão relacionados à variável MAT09 e ao intercepto do modelo. Considerando os coeficientes de MASC e SUP, variáveis não relacionadas diretamente com a ausência de dados, os desvios são pequenos (até 2,6%) quando o percentual de ausência é de 10% e 30%, e não ultrapassa 12,8% quando a ausência é simulada em metade das amostras.

TABELA 2 - Estimativas dos coeficientes do modelo de regressão aplicado a dados completos e a dados com ausência de informação do tipo MAR (simulada), acompanhados do desvio percentual, segundo o cenário de ausência e de tratamento

DADOS COMPLETOS		DADOS AUSENTES POR SIMULAÇÃO: TIPO MAR			
		TRATAMENTO			
		MÉDIA	LD	MV	IM
10% AUSENTES					
Intercepto	85,85	99,0 (15,3)	82,79 (-3,6)	82,75 (-3,6)	82,92 (-3,4)
MAT09	0,677	0,63 (-6,9)	0,689 (1,8)	0,689 (1,8)	0,688 (1,6)
MASC	3,017	3,14 (4,1)	2,974 (-1,4)	2,987 (-1,0)	2,999 (-0,6)
SUP	6,995	5,62 (-19,7)	7,017 (0,3)	7,019 (0,3)	7,034 (0,6)
30% AUSENTES					
Intercepto	85,85	127,7 (48,7)	75,3 (-12,3)	75,17 (-12,4)	75,77 (-11,7)
MAT09	0,677	0,54 (-20,2)	0,717 (5,9)	0,717 (5,9)	0,715 (5,6)
MASC	3,017	3,05 (1,1)	2,940 (-2,6)	3,008 (-0,3)	3,039 (0,7)
SUP	6,995	3,62 (-48,2)	7,049 (0,8)	7,074 (1,1)	7,081 (1,2)
50% AUSENTES					
Intercepto	85,85	168,00 (95,7)	61,67 (-28,2)	61,35 (-28,5)	62,65 (-27,0)
MAT09	0,677	0,39 (-42,4)	0,767 (13,3)	0,768 (13,4)	0,763 (12,7)
MASC	3,017	2,48 (-17,8)	2,631 (-12,8)	2,711 (-10,1)	2,796 (-7,3)
SUP	6,995	2,03 (-71,0)	6,937 (-0,8)	6,934 (-0,9)	6,984 (-0,2)

Fonte: Dados da pesquisa (elaboração dos autores).

Nota: Os valores em negrito do desvio percentual correspondem às estimativas com desvio maior que 10% em relação ao coeficiente da aplicação da regressão aos dados completos.

Os dados ausentes do tipo MNAR simulados neste estudo têm maior impacto nos resultados (Tabela 3). Para os procedimentos LD, MV e IM, o coeficiente da variável MASC é o mais afetado quando a ausência não é ao acaso. Nota-se que, quando não são consideradas variáveis auxiliares, os métodos MV e IM apresentam resultados praticamente iguais aos observados para o método LD.

No presente estudo, a utilização de variáveis auxiliares é importante para reduzir o impacto dos valores ausentes na estimação dos efeitos das variáveis MASC e SUP (Tabela 3). Considerando 30% e 50% de valores ausentes, verifica-se que os desvios observados para o coeficiente da variável MASC são reduzidos aproximadamente pela metade. Por exemplo, para o método MV, o desvio passa de 31,9% para 16,3%,

quando o percentual de ausentes é de 30%, e de 55,8% para 27,9%, quando 50% estão ausentes. O impacto da ausência de informação na variável SUP é menor, com desvios de 12,9% e 15,0% para os maiores percentuais de ausência, e com a utilização das variáveis auxiliares os desvios passam para 2,5% ou menos. Por outro lado, pode-se observar que a utilização de variáveis auxiliares gera maiores desvios na estimação dos coeficientes da variável MAT09 e do intercepto do modelo.

TABELA 3 - Estimativas dos coeficientes do modelo de regressão aplicado a dados completos e a dados com ausência de informação do tipo MNAR (simulada), acompanhados do desvio percentual, segundo o cenário de ausência e de tratamento

DADOS COMPLETOS		DADOS AUSENTES POR SIMULAÇÃO: TIPO MNAR					
		TRATAMENTO					
		MÉDIA	LD	MV		IM	
				SEM AUXILIARES	COM AUXILIARES	SEM AUXILIARES	COM AUXILIARES
10% AUSENTES							
Intercepto	85,85	100,1 (16,6)	84,3 (-1,8)	84,3 (-1,8)	83,9 (-2,3)	84,4 (-1,7)	83,9 (-2,3)
MAT09	0,677	0,63 (-6,9)	0,68 (0,4)	0,68 (0,4)	0,69 (1,9)	0,68 (0,4)	0,69 (1,9)
MASC	3,017	3,14 (4,1)	3,15 (4,4)	3,15 (4,4)	3,12 (3,4)	3,15 (4,4)	3,12 (3,4)
SUP	6,995	5,55 (-20,7)	6,96 (-0,5)	6,96 (-0,5)	6,97 (-0,4)	6,96 (-0,5)	6,97 (-0,4)
30% AUSENTES							
Intercepto	85,85	131,4 (53,1)	79,5 (-7,4)	79,5 (-7,4)	74,3 (-13,5)	79,4 (-7,5)	74,7 (-13,0)
MAT09	0,677	0,53 (-21,7)	0,71 (4,9)	0,71 (4,9)	0,72 (6,4)	0,71 (4,9)	0,72 (6,4)
MASC	3,017	3,00 (-0,6)	3,98 (31,9)	3,98 (31,9)	3,51 (16,3)	3,99 (32,3)	3,57 (18,3)
SUP	6,995	3,40 (-51,4)	6,09 (-12,9)	6,09 (-12,9)	6,82 (-2,5)	6,09 (-12,9)	6,83 (-2,4)
50% AUSENTES							
Intercepto	85,85	164,7 (91,8)	75,3 (-12,3)	75,3 (-12,3)	67,2 (-21,7)	75,3 (-12,3)	67,3 (-21,6)
MAT09	0,677	0,40 (-40,9)	0,73 (7,8)	0,73 (7,8)	0,75 (10,8)	0,73 (7,8)	0,75 (10,8)
MASC	3,017	2,28 (-24,4)	4,70 (55,9)	4,70 (55,8)	3,86 (27,9)	4,70 (55,8)	3,86 (27,9)
SUP	6,995	2,62 (-62,5)	5,95 (-14,9)	5,95 (-14,9)	7,06 (0,9)	5,94 (-15,1)	7,05 (0,8)

Fonte: Dados da pesquisa (elaboração dos autores).

Nota: Os valores em negrito do desvio percentual correspondem às estimativas com desvio maior que 10% em relação ao coeficiente da aplicação da regressão aos dados completos.

DISCUSSÃO

O estudo comparativo apresentado confirma resultados observados por outros autores. Por exemplo, verificou-se que a imputação pela média tem desempenho inferior na maioria das circunstâncias, o que corrobora a afirmação de McKnight *et al.* (2007) de que esse método não deveria ser utilizado pelos pesquisadores.

O desempenho dos procedimentos LD, MV e IM para análise de dados com ausência do tipo MCAR encontrado neste estudo é semelhante ao observado nos trabalhos de Enders (2001a) e de Schafer e Graham (2002). O primeiro autor relata que os coeficientes estimados para os quatro métodos avaliados (LD, MV, IM e *pairwise deletion*) apresentam pequenas variações em relação aos verdadeiros parâmetros. Já Schafer e Graham (2002) avaliaram apenas os métodos MV e IM e também verificaram pequena variação nas estimativas em relação aos parâmetros quando a ausência é completamente ao acaso.

No estudo de Enders (2001a) também foram simulados cenários com dados ausentes do tipo MAR. De forma geral, as estimativas dos coeficientes do modelo geradas pelos métodos avaliados mostram desvios relativamente pequenos. O autor afirma que o método MV apresentou melhor desempenho, porém a diferença entre os métodos é significativa apenas na estimação de um dos coeficientes do modelo. No estudo de Schafer e Graham (2002), não foram observados desvios expressivos nas estimativas dos coeficientes da regressão (para os métodos MV e MI) quando a ausência é ao acaso. Vale ressaltar que esses autores utilizaram dados simulados nas comparações.

O maior impacto dos valores ausentes do tipo MNAR também é reportado por outros autores. No estudo de Schafer e Graham (2002), os procedimentos MV e MI registraram estimativas para os coeficientes da regressão distantes dos valores verdadeiros. Em Enders (2001a), desvios expressivos foram observados em apenas uma variável e, nesse caso, o procedimento LD alcançou os melhores resultados.

Em Langkamp, Lehman e Lemeshow (2010) foram simulados dados ausentes do tipo MNAR, com percentual de

ausência variando de 10% a 40%. Nesse estudo foram utilizados dados completos extraídos de uma pesquisa da área de saúde. Os métodos LD e IM mostram desempenhos similares quando o percentual de ausentes é igual a 10%, mas, quando o percentual é maior, o IM gera estimativas mais próximas dos resultados observados para dados completos.

Uma amostra de dados completos obtidos a partir de dados reais também foi empregada no estudo de Young e Johnson (2013). Os métodos IM, MV e LD apresentam viés semelhante na estimação dos coeficientes da regressão quando os dados ausentes são do tipo MNAR, porém os autores ressaltam que os métodos IM e MV têm melhor desempenho na estimação dos erros padrão.

Croninger e Douglas (2005) usaram dados simulados baseados em variáveis presentes em uma pesquisa realizada entre ex-alunos de uma universidade. Os dados considerados nas comparações apresentam padrão arbitrário de ausência, contendo ausentes do tipo MCAR, MAR e MNAR. Os métodos MV e IM registraram resultados superiores em relação a métodos tradicionais, entre eles o LD, com menores desvios em relação aos coeficientes observados para a amostra completa.

Os benefícios da utilização de variáveis auxiliares são relatados no trabalho de Collins, Schafer e Kam (2001). Esse estudo tinha como objetivo principal avaliar o uso de variáveis auxiliares nos procedimentos MV e IM e, para isso, foram simulados diversos cenários variando os tipos de dados ausentes e a relação entre a taxa de ausência e as variáveis auxiliares. Em geral, os autores identificaram que o uso dessas variáveis é importante quando o percentual de ausentes é superior a 25% e a relação entre a variável auxiliar (usada no mecanismo gerador de valores ausentes) e a variável que apresenta os valores ausentes é forte. No presente estudo, tendo em vista que o modelo de regressão foi empregado para a comparação do desempenho entre grupos de estudantes (segundo o sexo e a intenção de ingresso no ensino superior), verificou-se também o benefício do uso de variáveis auxiliares. Apesar do aumento do viés nas estimativas do intercepto do modelo e do coeficiente da variável utilizada como controle (MAT09), observou-se a redução dos desvios

nas estimativas dos coeficientes das variáveis usadas na comparação dos grupos (MASC e SUP).

CONCLUSÕES E RECOMENDAÇÕES

A ausência de informação faz parte de praticamente toda pesquisa quantitativa e, portanto, deve-se fazer o máximo para reduzir seu impacto e assim evitar conclusões equivocadas a partir dos resultados observados. Segundo Allison (2002), os pesquisadores que pretendem mitigar os riscos associados aos dados ausentes devem escolher a estratégia de análise com cuidado, devendo considerar as características da ausência de informação. Nesse contexto, o presente estudo tem como objetivo principal contribuir para o melhor entendimento do tratamento e análise de dados incompletos.

Neste trabalho foram apresentadas algumas classificações de dados ausentes encontradas na literatura, em especial a proposta por Rubin (1976, 1987). Essa classificação é amplamente utilizada e extremamente útil na escolha do tratamento a ser empregado, no entanto outras informações devem ser analisadas. O pesquisador deve também considerar o objetivo da pesquisa e a técnica principal de análise de dados, a fonte e o padrão de ausência de informação, o tamanho da amostra, o percentual de ausência e a disponibilidade de variáveis que podem servir como auxiliares.

Dada a diversidade de situações e tipos de dados ausentes, não existe um procedimento único e infalível para a análise de dados. Inúmeras abordagens têm sido desenvolvidas ao longo dos anos e neste artigo foram apresentados resumidamente alguns dos procedimentos mais encontrados na literatura. Buscou-se aqui uma apresentação simples, logo muitos detalhes foram omitidos. Para os leitores interessados, Enders (2010) e McKnight *et al.* (2007) são boas referências, apresentando conceitos e métodos de forma acessível. Para os que buscam uma leitura mais teórica, ver Rubin (1987) e Allison (2002).

Outros procedimentos para tratamento de dados ausentes podem ser encontrados na literatura. Por exemplo, é comum na área social o uso de escalas para mensuração de

traços latentes (ansiedade, qualidade de vida, motivação) e, nesses casos, os dados ausentes de um item da escala podem ser substituídos pela média dos demais, sob a suposição de que todos representam medidas válidas do mesmo traço latente (SCHAFER; GRAHAM, 2002). Nos estudos longitudinais, a substituição pode ser feita pelo último valor observado para o mesmo indivíduo (MCKNIGHT et al., 2007). Nas ocasiões em que existem casos ausentes (um ou mais indivíduos sem resposta para todo o questionário), podem-se utilizar procedimentos de reponderação (*reweighting*), em que são calculados novos pesos amostrais para os casos presentes na amostra (SCHAFER; GRAHAM, 2002). Ainda, quando os dados ausentes são do tipo MNAR, os modelos de seleção e de mistura de padrões são muito úteis (FITZMAURICE et al., 2009).

O estudo de comparação realizado teve como foco o ajuste de modelos de regressão aplicados aos dados de uma avaliação educacional. Em geral, os desvios observados nas estimativas dos coeficientes da regressão quando 10% dos dados da variável resposta estavam ausentes foram pequenos. O impacto da ausência de dados depende do tipo de ausência (menor para MCAR e maior para MNAR) e aumenta quando o percentual de ausência é maior. A substituição simples pela média mostrou resultados insatisfatórios em todos os cenários. O procedimento *listwise deletion* apresenta resultados semelhantes aos procedimentos baseados na máxima verossimilhança e imputação múltipla nos cenários simulados, e as variáveis auxiliares foram importantes para redução dos desvios.

Os resultados do estudo comparativo aqui exposto devem ser interpretados com cautela. Optou-se pela utilização de apenas uma técnica de análise de dados na comparação dos métodos, o que limita a extrapolação dos resultados. Foram empregados modelos de regressão por se tratar de uma técnica amplamente usada na pesquisa aplicada. Em geral, os sistemas de avaliação educacional coletam diversas informações dos estudantes, escolas, professores e diretores, no entanto poucas variáveis foram utilizadas para simplificar a apresentação dos resultados.

O impacto das informações ausentes e a comparação dos métodos foram feitos apenas com as estimativas médias

3 O EQM representa a média das diferenças entre as estimativas individuais e o parâmetro de interesse, elevadas ao quadrado. Dessa forma, tal medida considera, além do viés, a variância das estimativas (COLLINS; SCHAFFER; KAM, 2001).

dos coeficientes, porém outros aspectos poderiam ser considerados. Uma análise rigorosa deveria avaliar não somente o viés dos estimadores, mas também o erro quadrático médio³ (EQM), estimativas dos erros padrão, cobertura de intervalos de confiança, conclusões de testes de significância e medidas de qualidade de ajuste dos modelos (COLLINS; SCHAFFER; KAM, 2001; ENDERS, 2001b). Além disso, nas comparações apresentadas, os coeficientes estimados a partir dos dados completos foram tratados como parâmetros do modelo proposto. Deve-se considerar que existe incerteza associada a esses coeficientes e que eles podem ser afetados por problemas de especificação do modelo, uma vez que não foram estimados a partir de população e sim de uma amostra.

Uma vez que o presente estudo teve como base os dados reais de uma avaliação, com variáveis contínuas e categóricas, a suposição de normalidade multivariada não está satisfeita. No entanto, acredita-se que esse fato não comprometa os resultados obtidos, pois as estimativas dos coeficientes da regressão são pouco afetadas pela não normalidade (ALLISON, 2002).

Recomenda-se a utilização dos procedimentos baseados na máxima verossimilhança e imputação múltipla. Além de apresentar bons resultados na estimação dos coeficientes, esses procedimentos são mais apropriados para a estimação de erros padrão, o que resulta em testes de hipóteses mais confiáveis. Os procedimentos MV e IM geram resultados similares, no entanto, com o uso de pacotes estatísticos adequados, o MV pode ser mais vantajoso pela facilidade de implementação. O procedimento *listwise deletion* alcança resultados semelhantes aos procedimentos mais sofisticados quando as variáveis auxiliares não são incorporadas nas análises; logo pode ser atrativo dada a enorme simplicidade. Por fim, a imputação simples pela média deve ser evitada.

REFERÊNCIAS

ALLISON, Paul D. *Missing data*. Thousand Oaks, CA: Sage, 2002.

ANDRIDGE, Rebecca R.; LITTLE, Roderick J. A. A review of hot deck imputation for survey non-response. *International Statistical Review*, New Jersey, v. 78, n. 1, p. 40-64, 2010.

BARALDI, Amanda N.; ENDERS, Craig K. An introduction to modern missing data analyses. *Journal of School Psychology*, Amsterdam, v. 48, p. 5-37, 2010.

BUCK, S. F. A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society, Series B*, London, v. 22, n. 2, p. 302-306, 1960.

CEARÁ. Secretaria da Educação. SPAECE – 2011. Matemática, 3º ano: ensino médio. Fortaleza: SEE, UFJF, 2011. p. 1-22. (Boletim Pedagógico, v. 3).

CHEEMA, Jehanzeb R. A review of missing data handling methods in education research. *Review of Educational Research*, Thousand Oaks, CA, v. 20, n. 10, p. 1-20, 2014.

COHEN, Jacob; COHEN, Patricia. *Applied multiple regression and correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum, 1985.

COLLINS, Linda M.; SCHAFER, Joseph L.; KAM, Chi-Ming. A comparison of inclusive and restrictive strategies in modern missing-data procedures. *Psychological Methods*, Washington, v. 6, n. 4, p. 330-351, 2001.

COX, Bradley E. et al. Working with missing data in higher education research: a primer and real world. *The Review of Higher Education*, Baltimore, v. 37, n. 3, p. 377-402, Spring 2014.

CRONINGER, Robert G.; DOUGLAS, Karen M. Missing data and institutional research. In: UMBACH, P. D. (Ed.). *Survey research: emerging issues of technology, policy, and analysis*. San Francisco: Wiley Interscience Periodicals, 2005. p. 33-49.

ENDERS, Craig K. The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, Thousand Oaks, CA, v. 61, n. 5, p. 713-740, 2001a.

ENDERS, Craig K. The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, Washington, v. 6, n. 4, p. 352-370, 2001b.

ENDERS, Craig K. *Applied missing data analysis*. New York: Guilford, 2010.

FITZMAURICE, Garret et al. *Longitudinal data analysis*. Boca Raton: Chapman & Hall, 2009.

GRAHAM, John W. Missing data analysis: making it work in the real world. *Annual Review of Psychology*, Palo Alto, CA, v. 60, p. 549-576, 2009.

GRAHAM, John W.; OLCHOWSKI, Allison E.; GILREATH, Tamika D. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, Berlin, v. 8, p. 206-213, 2007.

LANGKAMP, Diane L.; LEHMAN, Amy; LEMESHOW, Stanley. Techniques for handling missing data in secondary analyses of large surveys. *Academic Pediatrics*, Amsterdam, v. 10, n. 3, p. 205-210, maio/jun. 2010.

MACEDO, Glaucia Alves. *Fatores associados ao rendimento escolar de alunos da 5ª série (2000): uma abordagem longitudinal do valor adicionado e da heterogeneidade*. 2004. 212f. Dissertação (Mestrado em Demografia) – Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

MCKNIGHT, Patrick E. et al. *Missing data: a gentle introduction*. New York: Guilford, 2007.

OLIVEIRA, Pedro Rodrigues; BELLUZZO, Walter; PAZELLO, Elaine Toldo. The public–private test score gap in Brazil. *Economics of Education Review*, Amsterdam, v. 35, p. 120-133, 2013.

PEUGH, James L.; ENDERS, Craig K. Missing data in educational research: a review of reporting practices and suggestions for improvement. *Review of Educational Research*, Thousand Oaks, CA, v. 74, n. 4, p. 525-556, Winter 2004.

RODRIGUES, Clarissa Guimarães; RIOS-NETO, Eduardo Luiz Gonçalves; PINTO, Cristine Campos de Xavier. Diferenças intertemporais na média e distribuição do desempenho escolar no Brasil: o papel do nível socioeconômico, 1997-2005. *Revista Brasileira de Estudos de População*, Belo Horizonte, v. 28, n. 1, p. 5-36, jan./jun. 2011.

ROSE, Roderick A.; FRASER, Mark W. A simplified framework for using multiple imputation in social work research. *Social Work Research*, Oxford, v. 32, n. 3, p. 171-178, 2008.

ROUSSEAU, Michel et al. Reporting missing data: a study of selected articles published from 2003-2007. *Quality & Quantity*, Berlin, v. 46, n. 5, p. 1393-1406, 2012.

RUBIN, Donald B. Inference and missing data. *Biometrika*, Oxford, v. 63, n. 3, p. 581-592, 1976.

RUBIN, Donald B. *Multiple imputation for nonresponse in surveys*. New York: Wiley, 1987.

SAVALEI, Victoria; BENTLER, Peter M. A two-stage approach to missing data: theory and application to auxiliary variables. *Structural Equation Modeling*, London, v. 16, n. 3, p. 477-497, 2009.

SCHAFER, Joseph L.; GRAHAM, John W. Missing data: our view of the state of the art. *Psychological Methods*, Washington, v. 7, n. 2, p. 147-177, 2002.

SOARES, José Francisco; ALVES, Maria Teresa Gonzaga. Desigualdades raciais no sistema brasileiro de educação básica. *Educação e Pesquisa*, São Paulo, v. 29, n. 1, p. 147-165, jan./jun. 2003.

VINHA, Luís Gustavo do Amaral. *Estudos longitudinais e tratamento de dados ausentes em avaliações educacionais*. 2016. 124f. Tese (Doutorado em Psicologia Social, do Trabalho e das Organizações) – Instituto de Psicologia, Universidade de Brasília, Brasília, DF, 2016.

WILKS, S. S. Moments and distributions of estimates of population parameters from fragmentary samples. *The Annals of Mathematical Statistics*, New York, v. 3, p. 163-195, 1932.

XERXENEVSKY, Lauren Lewis. *Programa Mais Educação: avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul*. 2012. 143f. Dissertação (Mestrado em Economia do Desenvolvimento) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.

YOUNG, Rebekah; JOHNSON, David. Methods for handling missing secondary respondent data. *Journal of Marriage and Family*, New Jersey, v. 75, n. 1, p. 221-234, 2013.

LUIS GUSTAVO DO AMARAL VINHA

Professor adjunto do Departamento de Estatística da Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil
lgvinha@gmail.com

JACOB ARIE LAROS

Professor associado do Instituto de Psicologia da Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil
jalaros@gmail.com

Recebido em: 09 MAIO 2017

Aprovado para publicação em: 11 OUTUBRO 2017

AUSTRÁLIA, BRASIL E CANADÁ: IMPACTO DAS AVALIAÇÕES NO ENSINO DE CIÊNCIAS

PAULO SÉRGIO GARCIA

XAVIER FAZIO

DEBRA PANIZZON

NELIO BIZZO

RESUMO

A cultura global de testes se expandiu rapidamente nas últimas décadas. Nessas avaliações, as áreas mais comumente testadas são as de Linguagem e Matemática (LM). O foco e as políticas de responsabilização têm trazido consequências para o ensino de Ciências. Neste estudo reúnem-se evidências do Brasil, do Canadá e da Austrália para compreender as implicações dessas políticas para o ensino de Ciências, com base em pesquisas recentemente publicadas. Os resultados mostraram que a cultura dos testes, com ênfase em LM, e as políticas de responsabilização têm gerado o descaso e a negligência com o ensino de Ciências, causado sobretudo pelos sistemas de ensino e pelos gestores escolares, que criam iniciativas, ações e projetos voltados para atingir os resultados e metas em LM.

PALAVRAS-CHAVE ENSINO DE CIÊNCIAS • TESTES • POLÍTICAS EDUCACIONAIS • RESPONSABILIZAÇÃO.

AUSTRALIA, BRASIL Y CANADÁ: IMPACTO DE LAS EVALUACIONES EN LA ENSEÑANZA DE CIENCIAS

RESUMEN

La cultura global de pruebas se expandió rápidamente en las últimas décadas. En estas evaluaciones, las áreas más comúnmente probadas son las de Lenguaje y Matemáticas (LM). El enfoque y las políticas de responsabilización han provocado consecuencias en la enseñanza de Ciencias. En este estudio se reúnen evidencias de Brasil, Canadá y Australia para comprender las implicaciones de tales políticas para la enseñanza de Ciencias, en base a investigaciones recientemente publicadas. Los resultados mostraron que la cultura de las pruebas, con énfasis en LM, así como las políticas de responsabilización, han generado desatención y negligencia en la enseñanza de Ciencias, sobre todo en función de los sistemas de enseñanza y los gestores escolares, que crean iniciativas, acciones y proyectos que se destinan a alcanzar los resultados y metas en LM.

PALABRAS CLAVE ENSEÑANZA DE CIENCIAS • PRUEBAS • POLÍTICAS EDUCACIONALES • RESPONSABILIZACIÓN.

AUSTRALIA, BRAZIL AND CANADA: THE IMPACT OF ASSESSMENT ON SCIENCE TEACHING

ABSTRACT

The global testing culture has expanded rapidly in recent decades. In these assessments, the most commonly tested areas are those of Language and Mathematics (LM). The focus and accountability policies have brought consequences for the teaching of the sciences. Based on recently published research, the present study gathers evidence collected in Brazil, Canada and Australia in order to understand the implications of these policies for the teaching of the sciences. The results showed that both the culture of testing, with its emphasis on LM, and accountability policies have generated a disregard and neglect of science teaching. This situation is caused mainly by the education systems and the school managers who create initiatives, actions and projects aimed at achieving LM results and targets.

KEYWORDS SCIENCE TEACHING • TESTS • EDUCATIONAL POLICIES • ACCOUNTABILITY.

INTRODUÇÃO

Com a importância da educação crescendo globalmente, o debate internacional tem mudado da questão de fornecer acesso à educação para a de garantir a eficiência e a equidade nos resultados educacionais. Para certificar que o processo educacional seja de alta qualidade, no contexto da expansão global do ensino básico universal, políticas de responsabilização sob a forma de testes acadêmicos foram lançadas em quase todas as partes do mundo (SMITH, 2014).

Nesse sentido, nas últimas décadas tem acontecido uma grande expansão, em nível global, da cultura dos testes. Um contexto em que estudantes, em várias localidades, regiões e países, têm sido expostos a testes padronizados de avaliações em larga escala (SONGER; RUIZ-PRIMO, 2012).

Essa cultura dos testes tem transpassado vários aspectos da educação, influenciando questões como: financiamento, currículo escolar, gestão da escola, responsabilização dos atores e crenças e práticas dos professores, aulas dos alunos, envolvimento dos pais, entre outras. É necessário afirmar também que essa tendência do uso dos testes não se

limita aos países industrializados, uma vez que aqueles que discutem e realizam as políticas educacionais insistem que a implantação desses sistemas de testes (nacionais, estaduais e locais) é uma estratégia importante, talvez a chave para melhorar a qualidade educacional (CHAPMAN; SNYDER, 2000).

Os resultados dessas avaliações de larga escala (ALE), dos testes, têm induzido muitos gestores, professores, pais e mídia à percepção e ao entendimento de que os resultados são sinônimos de qualidade educacional, como se a avaliação dos alunos em algumas áreas de ensino (Linguagem, Matemática) pudesse demonstrar e garantir um panorama de excelência escolar.

A utilização de testes no campo educacional não se constitui em uma novidade das últimas décadas. Os testes têm sido usados como estratégia para orientar o ensino e a avaliação e, sobretudo, para identificar e acompanhar o rendimento dos alunos já há muito tempo. No entanto, a partir da segunda metade do século XX, mudanças aconteceram no tipo de teste aplicado aos estudantes, deslocando, pelo menos de forma parcial, a responsabilidade pelos resultados dos alunos para as escolas e para os professores (SMITH, 2014). Tal situação ocasionou um processo de maior responsabilização (*accountability*) daqueles que atuam nas escolas, sobretudo professores e gestores, um fato que começou a acontecer, primeiramente, nos Estados Unidos e no Reino Unido (DORN, 2007) e depois se espalhou pelo mundo. A partir da década de 1980, de acordo com informações da rede Eurydice (EUROPEAN COMMISSION, 2009), quase todos os países europeus já tinham adotado algumas políticas e testes censitários nacionais.

Essa expansão pode ser vista entre os anos de 1995 a 2006. Nesse período, o número de países no mundo que participou de algum tipo de programa de teste anual mais que dobrou, passando de 28 para 67, de acordo com dados de Benavot e Tanner (2007). Na América Latina, atualmente 16 nações apresentam algum tipo de programa de avaliação de larga escala. Países como o Chile, México, Colômbia e Brasil avaliam seus alunos regularmente, no entanto há variações na forma de uso dos resultados. No Brasil, com a utilização do

Sistema de Avaliação da Educação Básica (Saeb) e da Prova Brasil, desde 1990, o número de testes aplicados em jovens brasileiros chega a, aproximadamente, 19 milhões (HORTA NETO, 2013).

Sem negar o valor da avaliação para o trabalho pedagógico, pois permite aos sistemas e às redes de ensino possibilidades de utilização – definir padrões e expectativas para o aprendizado dos alunos, orientar o trabalho das escolas e dos docentes, gerar dados para os professores para a tomada de decisões sobre o progresso dos jovens, entre outras (BAUER; ALAVARSE; OLIVEIRA, 2015) –, há hoje grande discussão sobre as limitações e as consequências dessa cultura dos testes: a forma como são construídos e validados, a orientação e a legitimação da política educacional que, nesse caso, desconsidera o processo educacional e responsabiliza a escola e os professores pelos resultados escolares, entre outros. Como sinaliza Freitas (2013), as avaliações direcionam políticas de responsabilização, sobretudo quando integram políticas de alto impacto (*high stakes tests*), tal como o pagamento de bônus aos professores. Além disso, esse processo de responsabilização do docente e das escolas pode, em alguns casos, levar a medidas punitivas e injustas, como a transferência ou a perda do emprego de um profissional (gestor ou professor), produzindo injustiças associadas às premiações das melhores escolas (BAUER, 2012; FREITAS, 2013).

O contexto da cultura dos testes tem induzido os sistemas educativos à busca por melhores resultados em vez de melhora da qualidade e da equidade educativa (IAIES, 2003, p. 18), um processo que tem interferido na autonomia dos professores e na forma como os conteúdos têm sido ensinados. Em muitos casos, os professores atuam preparando os alunos para os exames, utilizando metodologias para atingir maiores rendimentos nas ALE. Ou seja, a fim de conseguir resultados significativos, os docentes posicionam o conteúdo e as formas de ensino sobre a malha de conhecimentos que é cobrada nos exames, deixando de lado outros assuntos, habilidades e competências que são essenciais para a formação dos alunos (MADAUS; RUSSEL; HIGGINS, 2009; HYPÓLITO, 2013; SANTOS, 2013). Um processo claro de empobrecimento

e afunilamento curricular (BAUER, 2013; MADDAUS; RUSSELL; HIGGINS, 2009).

Outra consequência da cultura dos testes é que essas ALE possibilitam a criação de *rankings* que induzem a competição entre as escolas e, nesse caso, podem deslocar o foco do aprendizado para melhoria dos resultados. Essas avaliações podem também gerar um ensino que considere a difusão de “macetes” para potencializar as notas dos estudantes (SANTOS, 2013). Elas podem ainda induzir o aumento da desigualdade, desencadeando um processo no qual investe-se mais nos melhores alunos, como uma estratégia mais favorável na busca por melhores resultados, do que naqueles que apresentam dificuldades de aprendizado e podem comprometer o rendimento da escola (OLIVEIRA, 2013).

Há também, como consequência, a questão da indução de gestores, educacionais e escolares, a investir maiores recursos físicos, intelectuais e financeiros nas disciplinas-alvo dos testes padronizados, em detrimento das outras matérias. Trata-se do caso de que esses testes, com foco apenas em Linguagem e Matemática, deixam de lado conhecimentos e habilidades relacionadas a Ciências, por exemplo, o que, possivelmente, traz consequências para o seu ensino. Nesse sentido, este estudo pretende averiguar as consequências dessas políticas de testes padronizados e de responsabilização para o ensino de Ciências. Para tal, reuniram-se evidências de três realidades de países diferentes (Austrália, Brasil e Canadá), utilizando pesquisas recentemente publicadas nesses contextos. Trata-se de um estudo internacional que apresenta alguns dados empíricos.

POLÍTICAS NEOLIBERAIS, O CONTEXTO DA EDUCAÇÃO E A CULTURA DOS TESTES

Políticas neoliberais indicaram tendências na organização e na gestão de políticas sociais, em geral, e educacionais, em particular. Tais programas sociais e de educação, já existentes nos países desenvolvidos no final da década de 1970 e início de 1980, atingiram outras nações por meio dos processos de globalização e formas de recontextualização, a partir

da década de 1990 (América Latina e Brasil), iniciando um processo denominado de “paradigma da educação contábil” (LIMA, 1997, p. 43).

No campo da educação, verificou-se o surgimento de um grande número de iniciativas distintas – processo de privatização e de desregulação, descentralização e reorganização do sistema para flexibilizar a oferta educacional, indução de um mercado educacional, políticas de livre escolha, contratos de gestão, compra de materiais didático-pedagógicos do setor privado, políticas de avaliação institucional e de responsabilização, vigilância da qualidade educacional por meio do controle da eficácia e da eficiência das escolas, criação de sistema de méritos (meritocracia), novas estratégias de gestão escolar baseada em conceitos de qualidade total, reformulação do currículo, indução de lógica competitiva entre as escolas, alteração nas práticas pedagógicas dos professores, entre outros – que, de forma geral, instituíram um discurso ideológico e tecnocrático já consolidado no mundo empresarial. Tal discurso obscureceu a linguagem da pedagogia e vem colonizando a política educacional pelos imperativos da economia (BALL, 1999). Trata-se da influência da ciência econômica e da economia capitalista na educação, em que, em muitos processos, as decisões sobre as políticas educacionais têm como base a economia, e não mais a pedagogia. Essa visão mercadológica da educação se confronta com o ideal de democratização e de gestão escolar estabelecidos no Brasil no final da década de 1980, por meio do ordenamento jurídico da educação.

Nesse contexto, ocorreu a criação e a disseminação da cultura dos testes, que atenderam e atendem a grandes interesses econômicos de empresas nacionais e multinacionais que criam e aplicam os testes e também atuam na negociação e venda de materiais didáticos. Um grande mercado de consumo da educação (HAGOPIAN, 2014).

Para Arelaro (2003) e Esteban (2012), as avaliações em larga escala vêm responder às pressões para mudanças nos modos de administração e de controle das redes de ensino, estando aliadas a um “novo” modelo de gestão educacional pública. Tais pressões, oriundas de organismos multilate-

rais, visam a impor uma agenda de educação para os países subdesenvolvidos. As ALE apresentam, portanto, um papel político, que tem relevância na análise de suas finalidades e de seus efeitos sobre os sistemas educativos em que são aplicados (OZGA, 2000).

As discussões sobre a validade desses testes padronizados (*standards*) apresentam muitas controvérsias. Stecher (2002, p. 99-100) sinaliza que o efeito líquido dessas avaliações de alto impacto sobre a política e a prática ainda é duvidoso, considerando o alinhamento curricular, a alocação do tempo de aula para enfatizar os assuntos abordados nos exames, o treinamento excessivo e a trapaça.

Há também discussões sobre os aspectos técnicos dos testes (validade, adequação dos instrumentos e confiabilidade dos resultados). Há dúvidas ainda se esses testes conseguem medir, em termos de aprendizagem dos alunos, o centro daquilo para o que foram projetados e com a acuidade anunciada de seus resultados (CASASSUS, 2013). Outra sinalização diz respeito à seleção dos conteúdos e habilidades que embasam os testes e, ao mesmo tempo, a definição de níveis ou padrões de rendimento. Nesse sentido, questionam-se as diretrizes na elaboração das matrizes e a definição de parâmetros de avaliação (CASASSUS, 2013; IAIES, 2003).

Há também dúvidas quanto ao potencial dos testes com o propósito de melhorar a qualidade do ensino, pois apesar de muitos esforços realizados, sobretudo pelos países latino-americanos, os resultados mostram pouca alteração no quadro educacional da maioria dos países que utilizam as avaliações como instrumento de gestão (TEDESCO, 2003).

Mons (2009) e Ravitch (2010) sinalizam algumas consequências dessas políticas de avaliação: a tomada de decisões sobre os recursos financeiros a serem endereçados para as escolas, a bonificação de professores para estabelecimentos de ensino que atingirem os resultados, a indicação de caminhos para o gerenciamento das redes escolares, a contratação e a demissão de gestores escolares, entre outros.

Madaus, Russell e Higgns (2009) afirmam que alunos e professores, quando pressionados pelos testes, podem sofrer problemas de saúde (estresse), perder a motivação, ou criar

um sentimento negativo em relação à escola. Os mesmos autores afirmam também que essa pressão pela melhoria do rendimento escolar pode induzir os gestores a outras formas de gerenciamento dos tempos e conteúdos, focando nos resultados, e tal situação pode influenciar de maneira negativa tanto docentes quanto discentes.

Por fim, outra questão polêmica é a divulgação das notas das escolas nesses testes padronizados. Um processo que não atenta para a complexidade dos resultados, assim como para o contexto em que eles foram produzidos (BRUNNER, 2003). Dessa forma, banalizam e, o que é mais preocupante, distorcem a opinião pública. As avaliações não levam em consideração as condições de partida de cada escola, a infraestrutura, os profissionais, suas formações, entre outros.

UMA VISÃO GERAL DA REALIDADE BRASILEIRA

No Brasil, a responsabilidade por organizar e integrar as políticas educacionais é do Ministério da Educação (MEC). Trata-se de um órgão da administração federal direta, que, por meio de suas secretarias, coordena, entre outras coisas, as orientações para o currículo, a avaliação, a formação de professores relacionada à educação básica e ao ensino superior. Atua também na informação e na pesquisa educacional. Estados e municípios também estabelecem suas diretrizes para seus sistemas, baseados nas políticas e indicações do MEC.

No país, na década de 1990, foi criado o Saeb, com o objetivo de avaliar a qualidade do ensino das escolas brasileiras. A portaria n. 482, de 7 de junho de 2013, do MEC, dispõe sobre o Saeb, revelando que o sistema passa a ser composto pela Avaliação Nacional da Educação Básica (Aneb), Avaliação Nacional do Rendimento Escolar (Anresc) e Avaliação Nacional da Alfabetização (ANA).

Em 1990, ocorreu a primeira aplicação do Saeb, que contou com a participação, de forma amostral, de escolas públicas urbanas com primeira, terceira, quinta e sétima séries do ensino fundamental. Os jovens foram avaliados em Língua Portuguesa, Matemática e Ciências. Em 1995 foi implantada a metodologia

da Teoria da Resposta ao Item (TRI), permitindo a comparabilidade entre os resultados das avaliações ao longo do tempo. Nesse ano, os alunos avaliados foram aqueles das quarta e oitava séries do ensino fundamental e do terceiro ano do ensino médio, incluindo uma amostra da rede privada. No entanto, nesse ano não aconteceu a avaliação em Ciências. De 1997 até 1999, os estudantes das quarta e oitava séries foram avaliados em Língua Portuguesa, Matemática e Ciências, e os estudantes de 3º ano do ensino médio, em Língua Portuguesa, Matemática, Ciências, História e Geografia. A partir de 2001, o exame do Saeb passou a focar exclusivamente as áreas de Língua Portuguesa e Matemática. Uma configuração que se mantém até os dias atuais (HORTA NETO, 2007; FREITAS, 2007).

Em 2013, no entanto, foi sinalizado que a disciplina de Ciências passaria a fazer parte novamente do Saeb, e então foi publicado o primeiro caderno de “Inclusão de Ciências no Saeb”, um documento básico criado pelo Ministério de Educação. Ao mesmo tempo, nesse ano, alunos do ensino fundamental II¹ e Médio participariam do exame de forma experimental.

De acordo com o documento:

A proposta apresentada assume a natureza de matriz piloto a ser refletida, desenvolvida, testada em 2013 e aperfeiçoada com vistas a sua consolidação em 2015. Nesse sentido, a matriz desenvolvida refere-se apenas ao 9º ano do ensino fundamental, referenciando a aplicação piloto no final do ciclo. As matrizes para o 5º ano e para a 3ª série do ensino médio serão desenvolvidas em seguida, após a análise de resultados da aplicação piloto, de forma a servir de parâmetro para a construção e a revisão das demais matrizes do Saeb. (BRASIL, 2013, p. 12)

O documento sinalizou também que, a partir de 2015, o exame de Ciências seria introduzido no Brasil. No entanto, apesar de todo o esforço, a inclusão dessa disciplina não foi consolidada. Em 2015, as avaliações realizadas incidiram somente sobre as áreas de Português e Matemática (PM).

As disciplinas de Português e Matemática, exclusivamente, também são avaliadas em outras avaliações em larga

¹ Adotou-se essa nomenclatura “ensino fundamental II” para se referir aos anos finais do ensino fundamental: sexto, sétimo, oitavo e nono anos.

escala no Brasil. Os alunos são testados nacionalmente nessas disciplinas, no segundo ano, pela Provinha Brasil, uma avaliação diagnóstica aplicada duas vezes ao ano; no terceiro, pela Avaliação Nacional da Alfabetização, que se insere no contexto de atenção voltada à alfabetização; nos quinto e nono anos do ensino fundamental, pela Prova Brasil. Além dessas, estados e municípios têm suas próprias avaliações em PM.

Tal expansão dessas ALE ocorreu, principalmente, depois da criação do Índice de Desenvolvimento da Educação Básica (Ideb), em 2007, que combina o desempenho dos estudantes na Prova Brasil, para os municípios, e no Saeb, para os estados, acrescido ao fluxo escolar (BAUER, 2012; SOUSA, 2013; BONAMINO, 2013).

O estudo de Brooke, Cunha e Faleiros (2011) mostrou que, no país, já existiam iniciativas de ALE em 19 estados brasileiros. Nos municípios, a pesquisa de Bauer, Alavarse e Oliveira (2015), com 4.309 cidades, revelou que as cidades estão se incorporando às iniciativas federais – ANA, Provinha Brasil, Prova Brasil – e estaduais – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saesp)² – de ALE. Paralelamente, os municípios também estão construindo suas próprias ações e programas de avaliação em Português e Matemática, “sejam esses concebidos pelas secretarias de educação, ou por empresas contratadas pela administração municipal” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 329), permitindo a realização de estudos sobre as características e os usos dos resultados dessas avaliações para formulação e implementação das políticas educacionais municipais (SOUSA; PIMENTA; MACHADO, 2012; ALAVARSE; MACHADO; BRAVO, 2013; GIMENES et al., 2013).

Se, no contexto brasileiro, as avaliações em larga escala, após 2001, estão assentadas exclusivamente nas disciplinas Português e Matemática, no contexto internacional, no entanto, o Brasil tem participado do Programa Internacional de Avaliação de Estudantes (PISA), no qual está inserida a avaliação em Ciências. Trata-se de um programa contínuo que objetiva o desenvolvimento de informações para o monitoramento de conhecimentos e de habilidades dos jovens

2 Avaliação aplicada pela Secretaria da Educação do Estado de São Paulo com a finalidade de produzir um diagnóstico da situação da escolaridade básica e, ao mesmo tempo, orientar os gestores escolares e monitorar as políticas associadas à qualidade educacional.

em vários países e em diferentes subgrupos demográficos de cada nação (ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OCDE, 2016).

O PISA é uma avaliação trienal e amostral em Leitura, Matemática e Ciências de alunos com idade entre 15 anos e 3 meses e 16 anos e 2 meses completos no início do período de aplicação. Esses jovens necessitam estar matriculados e cursando no mínimo a sétima série/ano. No último exame participaram 35 países da OCDE e mais 35 de economias parceiras, incluindo o Brasil.

Essa avaliação utiliza instrumentos, testes e questionários que resultam em indicadores de um perfil de conhecimentos e habilidades dos jovens. O programa também tem indicadores provenientes de questionários, variáveis demográficas, econômicas, sociais e educacionais e de tendências que monitoram o rendimento dos alunos e dos sistemas de educação. Em 2015, ano da última avaliação, participaram jovens de 27 estados, perfazendo um total de 23.141 alunos de 841 escolas, e, pela primeira vez, a aplicação do exame foi realizada totalmente pelo computador.

Os resultados de Ciências, nas últimas edições, não foram animadores para a realidade brasileira. O Brasil obteve os mesmos 405 pontos em 2009 e 2012, e, em 2015, apresentou uma queda para 401. Este escore médio (401) foi um valor significativamente menor do que a média dos alunos dos países membros da OCDE (493). Jovens da rede de ensino estadual conseguiram 394 pontos, e os da rede federal obtiveram o melhor desempenho (517), superando a média nacional. O estado do Espírito Santo apresentou o maior desempenho (435), e o de Alagoas, o pior (360). O desempenho médio dos meninos foi maior do que o das meninas na maioria dos estados do Brasil.

UMA VISÃO GERAL DA REALIDADE CANADENSE

O Canadá não tem um departamento federal ou um sistema nacional integrado de educação. Os ministérios de educação das respectivas províncias e territórios, criados em 1967, formaram uma organização intergovernamental

denominada de Conselho de Ministérios da Educação do Canadá (CMEC). Tal organização tem o intuito de prover liderança nas questões educacionais em todo o país e representá-lo em nível internacional, nomeadamente no que se refere à realização de comparações realizadas por meio de testes (CANADÁ, 2016a).

Cada uma de suas jurisdições dispõe de um departamento ou ministério de educação que realiza a organização, provisão e avaliação educacional nos níveis elementar e secundário. Dessa forma, cada província ou território apresenta um currículo próprio para o ensino, inclusive de Ciência e Tecnologia. Cada província tem também seu próprio sistema de avaliação (KLINGER; DELUCA; MERCHANT, 2016).

No Canadá, o Pan-Canadian Assessment Program (PCAP) (CANADÁ, 2016b) é o único teste nacional para o desempenho dos alunos de 13 anos (anteriormente àqueles de 16 anos) em Ciências. As províncias têm autonomia para analisar seu currículo e avaliar as escolas de ensino elementar (alunos de 6 a 14 anos) e secundário (de 14 a 18 anos). No país, várias políticas foram desenvolvidas, e recursos significativos são usados para promover a melhoria das escolas com base nos resultados das avaliações em larga escala.

O Canadá participa de testes de comparação internacional para estudantes das escolas elementares. No ensino de Ciências, trata-se do PISA e do Trends in International Mathematics and Science Study (TIMSS) (INTERNATIONAL ASSOCIATION FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT – IAEEA, 2012). Dessas comparações provêm dados internacionais que são utilizados para monitorar as mudanças do sistema educacional ao longo do tempo. Ao mesmo tempo é possível acompanhar os resultados, fomentar discussões sobre eles e analisar as políticas educacionais. Na província de Ontário, o foco das ALE é Linguagem e Matemática, e não há exames na disciplina de Ciências. Os alunos que participam dos testes estão no terceiro e no sexto ano (EDUCATIONAL QUALITY AND ASSESSMENT OFFICE – EQAO, 2016). As políticas de responsabilização e as ALE, com foco em LM, tiveram grande impacto no cenário educacional. Muitos recursos são usados, e políticas, desenvolvidas para promover a melhoria

com base em notas de avaliação em larga escala para as áreas de LM (VOLANTE; BEN JAAFAR, 2008).

UMA VISÃO GERAL DA REALIDADE AUSTRALIANA

A responsabilidade pela educação na Austrália é compartilhada, e recai sobre os seis Estados e os dois Territórios, resultando em oito órgãos ou departamentos educacionais diferentes que se responsabilizam, entre outras coisas, por monitorar os currículos e as avaliações dos alunos e por padronizar as normas educacionais. No país, não existe um único departamento de educação responsável por todos os programas educacionais. Trata-se de um processo descentralizado de administração educacional, que possibilita que cada jurisdição realize seus atendimentos relacionados às necessidades das crianças e dos jovens, estruturando e organizando currículos compatíveis com o contexto educacional específico (REID, 2005). No entanto, tal sistema se transformou, ao longo do tempo, em oito currículos diferentes em vigor no país. Apesar de receber o apoio de muitos educadores, pesquisadores e gestores escolares, essa proposta descentralizada carece de consistência nacional.

Na Austrália, existe somente uma avaliação em larga escala (teste), obrigatória em todo território nacional para todos os estudantes das escolas elementares. Trata-se do Programa Nacional de Avaliação: Linguagem e Matemática (National Assessment Program: Literacy and Numeracy – NAPLAN). O exame é administrado todo ano aos alunos da terceira e da quinta série (escolas elementares, na maioria dos estados e territórios), sétima e nona série (escolas secundárias na maioria dos estados e territórios), muito embora os familiares possam não permitir, ou seja, recusar que seus filhos realizem tais testes.

No país, o exame realizado pelo NAPLAN avalia os alunos em quatro domínios diferentes, relacionados à leitura, à escrita, às convenções da linguagem (soletração, gramática) e à aritmética (Matemática). No entanto, não existe na Austrália um exame nacional e obrigatório da disciplina de Ciências para todos os estudantes das escolas elementares.

A Austrália conta com o Programa Nacional de Avaliação de Alfabetização Científica (*National Assessment Program – Science Literacy – NAP-SL*) para avaliar os alunos em relação aos processos científicos. Trata-se de um exame que foi inicialmente implantado em 2003 para avaliar a alfabetização científica de todos os jovens. Até a presente data, o NAP-SL é administrado, a cada três anos e de forma amostral, a alunos da sexta série (jovens com 11 e 12 anos da escola elementar) de todo o território australiano selecionados randomicamente. No entanto, há uma grande discussão no país para que esse exame seja associado e incorporado aos testes realizados pelo NAPLAN. Tal associação permitiria que o exame passasse a ser anual e censitário.

A Austrália, assim como o Canadá, participa do PISA e do TIMSS, avaliações internacionais do conhecimento dos alunos em Matemática e Ciências. Os resultados do TIMSS de 2015 foram publicados recentemente (THOMSON; BORTOLI; UNDERWOOD, 2016), com a participação do quarto ano da escola elementar e do oitavo ano da escola secundária na Austrália.

O CONTEXTO DO ENSINO DE CIÊNCIAS

Em uma sociedade que valoriza o conhecimento científico e as tecnologias, o ensino de Ciências tem um papel fundamental para a formação de um cidadão crítico e transformador. Trata-se de um espaço privilegiado para a reflexão sobre o mundo, os fenômenos da natureza e a ação do homem. Em tal espaço, os jovens podem avaliar explicações, possibilitando o desenvolvimento de uma postura reflexiva, crítica, questionadora e investigativa. Esse ensino também favorece à compreensão de conceitos sobre as questões éticas implícitas nas relações entre Ciência, Sociedade e Tecnologia. Além disso, sua abrangência inclui a compreensão da crise ambiental, a destruição dos seres humanos, as questões do consumismo, o destino dado ao lixo industrial e doméstico, a manipulação gênica, entre outros.

Em muitos países o Ensino de Ciência, e em graus diferentes a tecnologia, é um dos elementos-chave para a escolarização. No entanto, esse ensino tem enfrentado vários

desafios, entre eles o desinteresse dos alunos por essa disciplina. Há também a questão do baixo interesse dos jovens em seguir as carreiras científicas, que, segundo pesquisadores, está associado às aulas de Ciências (AIKENHEAD, 2005; BERNARDO et al., 2008; BROK et al., 2006; GOUW; BIZZO, 2016), ou seja, à forma como as aulas são ministradas, sobretudo baseadas no livro didático e na memorização.

Além disso, a conjuntura atual pode ser ainda mais agravada em consequência da criação das políticas de avaliação em larga escala e de responsabilização da escola e dos professores, que focam, no Brasil, exclusivamente, as áreas de Linguagem e Matemática, deixando o ensino de Ciências em segundo plano.

METODOLOGIA DO ESTUDO

Este estudo, parte de um projeto internacional de pesquisadores na Austrália, no Brasil e no Canadá, buscou evidências, em três realidades distintas, para compreender as implicações de um cenário marcado por políticas de avaliação em larga escala, com o foco em Linguagem e Matemática, e de responsabilização das escolas para a disciplina de Ciências.

Em estudos comparativos internacionais na área de educação, é necessário reconhecer as diferenças políticas, culturais e históricas de cada país para que se avalie, de fato, o que pode ser analisado, compartilhado e incorporado. Phillips (2000) realizou uma discussão relevante sobre os motivos para a realização de estudos comparativos na área de educação. Para o autor, a proficuidade desses estudos se deve, entre outras coisas, à descoberta de possíveis alternativas para a organização dos sistemas, das redes de ensino e das escolas; às análises das consequências das reformas, avaliando experiências similares em vários países; à avaliação da possibilidade da promoção de cooperação entre as nações, a partir da discussão social, política e cultural sobre as similaridades e as diferenças.

A seleção dos países participantes decorre do fato de o Canadá e a Austrália serem parte da Commonwealth, uma organização internacional com 54 membros independentes.

Tais países atuam na cooperação inseridos em um quadro de valores e objetivos confirmados na declaração de Singapura, incluindo a promoção da democracia, a busca pelos direitos humanos e pela liberdade individual, o comércio livre, o multilateralismo e a paz entre as nações. Esses dois países não apresentam sistemas federal ou nacional de educação. Tal situação resulta na criação, em cada estado ou jurisdição, de um departamento de educação independente, o que não acontece no Brasil, que dispõe de um ministério da educação nacional para coordenar as diretrizes políticas, de currículo e de avaliação em todo território. Dessa forma, embora possam existir algumas semelhanças entre o Canadá e a Austrália, há também muitas diferenças, propiciando uma comparação interessante com o Brasil, que não tem o mesmo histórico de educação que os outros dois países. Além disso, a seleção de tais contextos dá continuidade a uma série de estudos realizados pelos pesquisadores nas três realidades (FAZIO; GARCIA; PANIZZON, 2008; GARCIA; FAZIO; PANIZZON, 2011; GARCIA; BIZZO; FAZIO, 2012; GARCIA; BIZZO; FAZIO, 2014), tendo como marco a formação de professores de Ciências, os desafios da formação contínua a distância para docentes que atuam na área de Ciências e, nesse caso, a avaliação educacional.

Nos três países, os pesquisadores coletaram evidências a partir de estudos e pesquisas recentemente publicadas. Da Austrália, foram utilizadas evidências do relatório *TIMSS 2015: a first look at Australia's results* (THOMSON; BORTOLI; UNDERWOOD, 2016), do Conselho Australiano de pesquisa educacional (Australian Council for Educational Research – ACER). Trata-se de uma organização que visa a promover conhecimento baseado em pesquisas para ser usado com o intuito de melhorar a aprendizagem ao longo da vida.

Do Canadá, região de Ontário, foram utilizadas evidências reunidas na pesquisa *Exploring the Current State of Grades 4 to 8 Science Education in Ontario* (DOW, 2015), um estudo com metodologia mista com mais de 200 professores de Ciências da região, e o estudo de Fazio e Karrow (2013), *Science takes back seat*, que, além de entrevistar docentes, analisou também dados do TIMSS de 2015 para substanciar as análises.

Do Brasil, da região do grande ABC, foram utilizados dados dos estudos: *Educational policies and science education in Brazil: a case study* (GARCIA; BIZZO, 2015), que analisou a percepção dos jovens quanto à importância das disciplinas de Português, Matemática e Ciências estudadas cotidianamente em um contexto marcado por políticas públicas que enfatizam exclusivamente a Linguagem e a Matemática; *Efeitos Ideb na região do grande ABC sobre o ensino de Ciências*, realizada pelo Observatório da Educação do Grande ABC (OEGABC, 2015), com análise do Ideb e suas consequências no ensino de Ciências; *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências* (OEGABC, 2016), que, por meio de um estudo de caso, analisou a percepção de 49 professores de Ciências dessa região, uma das mais ricas do país, sobre a percepção acerca das políticas atuais de avaliação. É necessário reconhecer, no caso desses estudos, que se trata de uma amostra particularizada, de uma região, e que, portanto, não pode ser tomada como possuidora de representatividade nacional.

Como amostra nacional foram analisados os dados do PISA, na área de Ciências, das últimas quatro edições (2006, 2009, 2012 e 2015), que são os resultados publicados logo depois da criação do Ideb no Brasil, a principal política nacional em relação à responsabilização das escolas e dos professores com foco apenas nos testes de Português e Matemática.

Por fim, foram utilizados e avaliados os dados do PISA, dos três países, na área de Ciências, dos anos de 2006, 2009, 2012 e 2015, buscando ampliar a análise e os entendimentos sobre o fenômeno. Considera-se que as influências das avaliações em larga escala, com foco exclusivo em Linguagem e Matemática, apresentem também implicações nos resultados dos alunos de 15 anos.

RESULTADOS E DISCUSSÃO

Os dados, coletados em cada contexto de cada país, são apresentados inicialmente relacionados às suas realidades e particularidades próprias: Austrália, Brasil e Canadá.

CONTEXTO AUSTRALIANO

No contexto australiano, as evidências são provenientes da análise dos dados do TIMSS, a partir do ano de 1995. A Tabela 1 sintetiza os resultados:

TABELA 1 - Média da Austrália em Ciências no TIMSS

	1995	2003	2007	2011	2015
Quarto ano	521	521	527	516	524
Oitavo ano	514	527	515	516	512

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os dados indicaram, na série histórica (1995-2015), que os resultados obtidos pelos jovens, em Ciências, permaneceram praticamente no mesmo patamar ao longo dos anos, considerando que as médias das pontuações não foram estatisticamente diferentes de um período para outro, na referida série ($p > 0,05$), para os quartos e oitavos anos. Uma estagnação que coincidiu com a proliferação das políticas de avaliação com o foco substantivo em Linguagem e Matemática no país.

Em 2015, os resultados dos alunos do quarto ano mostraram que, enquanto a Austrália superou, em relação ao desempenho dos estudantes em Ciências no TIMSS, 17 países, ela foi superada por outras 17 nações. Da mesma forma, no oitavo ano, o país superou, em termos de diferenças estatisticamente significantes, 20 outros, mas foi superado por outros 14 países. Tal situação ilustra também o processo de estagnação em relação ao desempenho do ensino de Ciências.

Outros estudos locais indicaram (REID, 2009; LINGARD, 2010) que o forte investimento em políticas de avaliação em Linguagem e Matemática e de responsabilização, utilizando programas de mérito, tem, entre outras coisas, afetado e reduzido os currículos àquilo que tem sido cobrado nos testes nacionais, revelando o fenômeno de afinilamento curricular presente em outras partes do mundo (BAUER, 2013; MADAUS; RUSSELL; HIGGINS, 2009). Ao mesmo tempo, essas iniciativas têm interferido na prática pedagógica dos professores e, em alguns casos, indicado um processo de uniformização de ensino.

É necessário sinalizar que os investimentos nessas políticas, com foco em LM e de responsabilização, continuam em pleno vigor na Austrália e estão sendo ainda mais ampliadas, como pode ser apreciado em um documento recente publicado pelo governo de Nova Gales do Sul: *Literacy and Numeracy Strategy 2017-2020* (NEW SOUTH WALES, s.d.).

CONTEXTO BRASILEIRO

No Brasil, as evidências coletadas e apresentadas são provenientes das análises realizadas sobre os dados do PISA dos anos de 2009, 2012 e 2015, e outras originárias particularmente da região do Grande ABC Paulista.

BRASIL NO PISA

Os dados do PISA mostraram que o país estava em 63º lugar no *ranking* de Ciências, em 2015, quatro posições abaixo da atingida em 2012. Uma posição incômoda para a nação que é a nona economia do mundo. A Tabela 2 sintetiza os dados do PISA:

TABELA 2 – Médias do Brasil em Ciências no PISA

2006	2009	2012	2015
390	405	405	401

Fonte: Programme for International Student Assessment.

Os dados do PISA revelam, apesar de uma redução na pontuação média do ano de 2015, estagnação do país no desempenho dos alunos em Ciências, considerando que

[...] não foram encontradas evidências empíricas que apontem diferenças estatisticamente significativas entre o desempenho dos estudantes brasileiros em Ciências no PISA 2015 e nas três últimas edições da avaliação. (BRASIL, 2015, p. 16)

Não se pode assumir diretamente que a ausência de avanços nos dados seja reflexo das políticas de avaliação em larga escala e de responsabilização. No entanto, pode-se constatar que a estagnação começou, coincidentemente, após a criação da principal política para a verificação da qualidade da educação básica no Brasil, o Ideb, que utiliza dados

da Prova Brasil e do Saeb, ambos incidindo apenas sobre as áreas de Português e Matemática.

A REGIÃO DO GRANDE ABC

Na região do grande ABC, as evidências foram coletadas a partir de três pesquisas realizadas na área. O primeiro estudo (GARCIA; BIZZO, 2015) mostrou a percepção de um grupo de alunos de escolas de ensino fundamental I³ em relação às disciplinas mais relevantes no contexto escolar. Esses jovens acreditavam que as matérias mais importantes eram as de Português e Matemática e secundarizavam as Ciências.

3 Adotou-se essa nomenclatura “ensino fundamental I” para se referir aos anos iniciais do ensino fundamental: primeiro, segundo, terceiro, quarto e quintos anos.

Os pesquisadores mostraram que os jovens foram influenciados na construção desse imaginário pelos professores, escola e sistema de ensino, inseridos em um contexto marcado pela cultura do teste em Português e Matemática e de responsabilização. No primeiro caso, os docentes solicitavam que os jovens estudassem mais as matérias de PM para a realização dos simulados da escola, para a prova local e para a Prova Brasil ou Saresp. Alguns professores atribuíam notas para os alunos realizarem esses testes de forma séria. No segundo, várias estratégias eram utilizadas para que os estudantes tivessem altos resultados nas avaliações de larga escala: realizavam algum tipo de prova em PM duas ou três vezes ao ano, simulando a Prova Brasil, participavam de projetos de reforço relacionados a essas disciplinas no contraturno, entre outros. Salienta-se ainda que alguns diretores apresentavam estratégias bem definidas para que os alunos com baixo rendimento intelectual não realizassem as provas (alguns ficavam na biblioteca da escola). No último caso, em relação às secretarias de educação e suas políticas educacionais, eram realizadas provas municipais para averiguação das habilidades dos alunos em PM, e os resultados eram avaliados pelos especialistas da secretaria, pelos diretores e coordenadores das escolas e depois pelos professores e os jovens. Esses departamentos endereçavam maiores recursos (compra de materiais) para as áreas de Português e Matemática. Um grande movimento que induzia nos alunos a crença de que essas matérias eram as mais importantes no contexto escolar.

O segundo estudo, na mesma região (OEGABC, 2015), com escolas de ensino fundamental II, mostrou que elas valorizavam os resultados do Ideb ou do Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp), e concentravam esforços nas áreas de PM: os professores dessas disciplinas dispunham de mais tempo com a coordenação pedagógica para reuniões, participavam mais de projetos externos (olimpíada de Matemática, por exemplo) e possuíam maiores recursos para a compra de materiais (jogos, *softwares*, entre outros). A disciplina de Ciências ficava em um segundo plano e não era valorizada. Em alguns casos, ela era encarada como uma matéria de conhecimentos secundários. Tal situação ocorria porque uma alta posição no *ranking* do Ideb indicava que a escola seria muito mais reconhecida e valorizada pelos pais e pela rede de ensino. Em alguns contextos (Santo André), os professores recebiam um bônus salarial em virtude dos bons resultados alcançados. Trata-se do projeto de meritocracia (FREITAS, 2013).

O terceiro estudo foi realizado com 49 professores de Ciências do ensino fundamental II da mesma região: *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências* (OEGABC, 2016). A pesquisa tinha o objetivo de conhecer a percepção do professor sobre as políticas de avaliação em larga escala com foco em Português e Matemática e de responsabilização difundidas na região do grande ABC.

Os resultados mostraram uma realidade preocupante em relação ao ensino de Ciências. Em geral, faltavam recursos para equipar os laboratórios e materiais para a realização de experiências, a disciplina não era valorizada pelas redes de ensino e pela gestão escolar e o ensino era baseado praticamente no uso do livro didático. Por outro lado, os maiores recursos das secretarias de educação eram endereçados aos programas e projetos nas áreas de PM. As escolas ofereciam reforço escolar, olimpíadas, jogos, simulados focando nessas duas disciplinas. Em algumas cidades, só existiam professores eventuais em PM. Havia também mais tempo destinado à formação desses professores (cursos, congressos). Os cursos, palestras e encontros, em alguns municípios

(São Caetano do Sul), focavam mais os conteúdos dessas disciplinas. O tempo de reuniões de diretores, coordenadores e docentes dessas áreas era bem maior nas escolas (São Caetano do Sul). Nenhuma dessas estratégias descritas existiam para fomentar a aprendizagem em Ciências.

Um fato a se destacar é que alguns professores de Ciências faziam piadas sobre seus colegas de PM, sinalizando que eles trabalham muito mais, e no final o bônus era recebido por todos. Em geral, esses docentes sinalizaram que essas questões começaram a acontecer depois da criação e instauração do processo de avaliação em larga escala.

Uma análise documental das escolas desses 49 professores revelou que nos últimos cinco anos mais de 55 projetos associados às disciplinas de Português e Matemática foram criados nas instituições de ensino para melhorar o rendimento dos alunos nos testes padronizados (simulado da escola, prova local, Saesp, Prova Brasil). Não houve mais de dez projetos relacionados às Ciências.

CONTEXTO CANADENSE

No contexto canadense, o estudo de Dow (2015), com professores de Ciências, mostrou que os docentes consideravam que a disciplina não era prioridade nas políticas públicas educacionais recentes, nos programas e projetos dos sistemas de ensino e nas iniciativas e ações realizadas pelas escolas e pelos gestores escolares.

Para a maioria dos professores, a disciplina de Ciências era considerada como sendo de “segunda classe”, considerando que os docentes das áreas de Linguagem e Matemática possuíam os maiores recursos pedagógicos e materiais para as aulas (laboratório de redação, *software* para Matemática), mais horas de formação (cursos, encontros), maior número de projetos extraclasse (olimpíada de Matemática), mais tempo de ensino (número maior de aulas semanais) e mais apoio dos diretores. O foco excessivo em LM, visando a melhorar o rendimento do aluno, é fortemente incentivado nos documentos oficiais recentes publicados pela Educational Quality and Assessment Office (EQAQO, 2015), uma organização que avalia o quanto o sistema de ensino público de Ontário está

desenvolvendo as habilidades de leitura, escrita e matemática dos alunos. A instituição fornece informações confiáveis e úteis que são usadas para auxiliar a melhorar o desempenho dos jovens e garantir a responsabilidade dos conselhos escolares (EQAQ, 2016).

O estudo de Fazio e Karrow (2013), na mesma região, mostrou, por um lado, a falta de recursos para o ensino de Ciências, por outro, a desmotivação dos professores para o ensino. Paralelamente, esses pesquisadores revelaram uma redução do desempenho dos alunos no TIMSS nos anos de 2003, 2007 e 2011. A Tabela 3 sintetiza os dados dos quarto e oitavo anos:

TABELA 3 - Média do Canadá em Ciências no TIMSS

	2003	2007	2011	2015
Quarto ano	540	536	528	530
Oitavo ano	533	526	521	524

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os dados revelam uma tendência de queda nos resultados dos jovens na área de Ciências a partir de 2003, em ambos os anos (quarto e oitavo). Em 2011, o desempenho diminuiu, em média, 15 pontos em relação ao ano de 2003. Quando é realizada uma análise do Índice de Referência Internacional (IRI), os dados também apontam para uma diminuição no aproveitamento dos alunos: A Tabela 4 revela os dados do IRI:

TABELA 4 - Desempenho dos jovens de acordo com o Índice de Referência Internacional

NÍVEL DE DESEMPENHO		2003	2007	2011
Avançado	Quarto ano	13%	12%	9%
	Oitavo ano	7%	7%	6%
Alto	Quarto ano	47%	45%	40%
	Oitavo ano	41%	37%	35%
Intermediário	Quarto ano	81%	79%	77%
	Oitavo ano	81%	77%	76%
Baixo	Quarto ano	96%	95%	94%
	Oitavo ano	97%	96%	96%

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os resultados revelam que no IRI o percentual de alunos no nível avançado, alto e intermediário também vem diminuindo a partir de 2003. Esses dados indicam um declínio do rendimento dos jovens nos níveis, como mostra a Tabela 4.

Sinteticamente, os resultados mostram queda no aproveitamento dos jovens. Em 2011, as notas médias dos testes em Ciências (Tabela 3) eram significativamente mais baixas do que em 2003 para os alunos do quarto e oitavo anos. Essa queda do rendimento, no país, coincidiu com o crescimento das políticas de avaliação com o foco em LM, em que os principais recursos eram e são utilizados para melhorar as notas de avaliação em larga escala nessas áreas (VOLANTE; BEN JAAFAR, 2008).

O PISA NAS TRÊS REALIDADES ANALISADAS

⁴ Embora seja possível a comparação de mudanças no desempenho dos alunos ao longo do tempo em cada domínio do PISA (Leitura, Matemática e Ciências), há que se considerar que cada edição do exame tem um foco particular. Em 2006 e 2015, o foco recaiu sobre Ciências.

A análise histórica dos dados do PISA⁴ dos três países avaliados é reveladora de realidades distintas. Embora a ausência de melhora nos resultados exija uma análise mais detalhada, é importante notar o comportamento do desempenho dos alunos em Ciências no período compreendido entre 2006 e 2015. A Tabela 5 ilustra os resultados:

TABELA 5 – Média em Ciências no PISA, segundo o país

PAÍS	2006	2009	2012	2015
Austrália	527	527	521	510
Brasil	390	405	405	401
Canadá	534	529	525	528

Fonte: Programme for International Student Assessment (elaboração dos autores).

Os resultados apresentados na Tabela 5 revelam, no contexto brasileiro (BRASIL, 2015, p. 16) e também no canadense (O'GRADY et al., 2016, p. 29), uma estagnação nos dados de Ciências no período de 2006 a 2015, considerando que não existem diferenças estatisticamente significativas.

No caso australiano, que apresenta uma queda ao longo das últimas três avaliações, os dados, no entanto, apontam em outra direção, indicando que existem diferenças estatisticamente significativas no desempenho em Ciências no período entre 2006 e 2015 (THOMSON; BORTOLI; UNDERWOOD, 2015).

Embora Canadá e Austrália não tenham apresentado aumento no desempenho médio de seus estudantes em Ciências, afetado também pelas ALE que priorizam somente Língua Materna e Matemática, a situação do Brasil é claramente mais grave, uma vez que os aspectos de letramento científico contemplado no exame estão muito aquém dos resultados dos dois países, em média cerca de 20%, ou mais de 100 pontos numa escala de 0 a 700, em todas as edições do exame.

CONSIDERAÇÕES FINAIS

Este estudo reuniu e apresentou evidências de três realidades diferentes, demonstrando as consequências para o ensino de Ciências de um cenário marcado por políticas de avaliação em larga escala com foco exclusivo em Linguagem e Matemática e de responsabilização.

Entre as consequências, verificou-se que essas políticas estão induzindo à distribuição desigual dos recursos e, o que é mais preocupante, muitas vezes, dentro da mesma escola,

incluindo a questão do tempo de formação e da criação de programas e projetos educativos e pedagógicos, que estão muito mais centralizados nas disciplinas de Linguagem e Matemática.

Essas políticas estão desalinhando a questão da formação dos professores, à medida que os tempos para capacitação docente tem sido, significativamente, maiores para os docentes das áreas de Linguagem e Matemática, assim como os investimentos também têm sido mais amplos nessas áreas, deixando as outras disciplinas em um segundo plano.

Nas três realidades investigadas, uma das conclusões válidas relaciona-se à questão de que essas políticas com ênfase em LM e de responsabilização têm afetado negativamente a qualidade do ensino de Ciências. No contexto australiano, foi encontrado estagnação no desempenho dos alunos em Ciências no TIMSS, nos quartos e nonos anos, isso depois de 20 anos dessas políticas no país. Além disso, tem acontecido uma redução dos currículos e uma uniformização nas práticas pedagógicas dos professores, como sinalizado pelos pesquisadores. Há ainda a questão de que os investimentos nesses programas continuam avançando.

No Canadá e no Brasil, a disciplina de Ciências não tem sido prioridade para as autoridades e gestores, e também não tem sido valorizada pelas redes de ensino e pelos gestores escolares. Entre outras situações identificadas em ambas as realidades avaliadas, a disciplina foi considerada como de segunda classe. Há ainda a falta de recursos para as escolas e para os docentes (laboratórios e materiais) e a desmotivação dos professores para o trabalho de sala de aula.

No Canadá, houve uma redução no desempenho dos alunos no TIMSS que também se manifesta no Índice de Referência Internacional dos últimos 15 anos, coincidindo com os avanços das políticas de avaliação e de responsabilização. No Brasil, há uma estagnação no desempenho dos alunos em Ciências no PISA nas últimas edições (nos outros dois países, queda), coincidentemente após a criação da principal política de responsabilização (Ideb). Ao mesmo tempo, em uma das áreas mais ricas do país, a região do Grande ABC, essas políticas têm influenciado o imaginário, a

percepção dos alunos em relação à importância das disciplinas de Português e Matemática nas escolas e, nesse sentido, a disciplina de Ciências tem sido secundarizada.

Todas essas iniciativas e ações direcionadas com o foco em Linguagem e Matemática e essas políticas de prestação de contas têm sido endereçadas para melhorar os resultados dos alunos nos testes nas áreas de LM. No entanto, essas políticas têm gerado como consequência não intencional o descuido e o descaso com o ensino de Ciências, que tem sido negligenciado. Tal desvalorização e secundarização, considerando a relevância dessa disciplina na formação do cidadão em um contexto social que valoriza o conhecimento científico e as tecnologias, pode ampliar ainda mais o desinteresse dos alunos no ensino de Ciências, contribuindo para que os jovens sejam menos críticos, reflexivos, questionadores e transformadores da realidade em relação à crise ambiental, ao consumismo, à manipulação gênica, entre outros. Ao mesmo tempo, essas políticas podem distanciar ainda mais os jovens das carreiras científicas, o que já vem acontecendo em várias partes do mundo (AIKENHEAD, 2005; BERNARDO et al., 2008; BROK et al. 2006; GOUW; BIZZO, 2016).

Os resultados deste estudo podem ser utilizados pelas secretarias de educação e pelos gestores escolares para ampliar a compreensão sobre o fenômeno, alertar as autoridades políticas e educacionais sobre as consequências de tais políticas e fomentar o debate sobre a qualidade educacional.

REFERÊNCIAS

AIKENHEAD, Glen S. Research Into STS Science Education. *Educación Química*, Belo Horizonte, v. 16, p. 384-397, 2005.

ALAVARSE, Ocimar Munhoz; MACHADO, Cristiana; BRAVO, Maria Helena. Avaliações externas e qualidade na educação básica: articulações e tendências. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 54, p. 12-31, jan./abr. 2013.

ARELARO, Lisete Regina Gomes. Direitos sociais e política educacional: alguns ainda são mais iguais que outros. In: SILVA, S.; VIZIM, M. (Org.). *Políticas públicas: educação, tecnologias e pessoas com deficiências*. Campinas: Mercado de Letras, 2003. p. 13-36.

- BALL, Stephen J. Cidadania global, consumo e política educacional. In: SILVA, L. H. da. *A escola cidadã no contexto da globalização*. Petrópolis: Vozes, 1999. p. 127-137.
- BAUER, Adriana. Estudos sobre sistemas de avaliação educacional: um retrato em branco e preto. *Revista @mbienteeducação*, São Paulo, v. 5, p. 7-31, 2012.
- BAUER, Adriana. Limites, desafios e possibilidades das avaliações de sistemas educacionais: contribuições do ciclo de debates para as políticas de avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialava R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores*. Florianópolis: Insular, 2013. p. 281-294.
- BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela. Avaliações em larga escala: uma sistematização do debate. *Educação e Pesquisa*, São Paulo, v. 41, n. especial, p. 1367-1382, dez. 2015.
- BENAVOT, Aaron; TANNER, Erin. The growth of national learning assessments in the world, 1995-2006. Paris: Unesco, 2007. Background paper for the Education for All Global Monitoring Report 2008, Education for All by 2015: Will we make it.
- BERNARDO, Allan B. I. et al. Students' perceptions of science classes in the Philippines. *Asia Pacific Education Review*, Philippines, v. 9, n. 3, p. 285-295, ago. 2008.
- BONAMINO, Alicia C. de. Avaliação educacional no Brasil 25 anos depois: onde estamos? In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores*. Florianópolis: Insular, 2013. p. 43-60.
- BRASIL. Instituto Nacional de Estudos e Pesquisas. *Inclusão de Ciências no SAEB*: documento básico. Brasília, DF: Inep, ago. 2013.
- BRASIL. Diretoria de Avaliação da Educação Básica. *Brasil no PISA 2015*: sumário executivo. Brasília, DF, 2015. Disponível em: <http://download.inep.gov.br/acoes_internacionais/pisa/documentos/2016/pisa_brasil_2015_sumario_executivo.pdf>. Acesso em: 04 mar. 2016.
- BROK, Perry Den et al. Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation: An International Journal on Theory and Practice*, England & Wales, London, v. 12, n. 1, p. 3-25, 2006.
- BROOKE, Nigel P.; CUNHA, Maria Amália; FALEIROS, Matheus. *A avaliação externa como instrumento da gestão educacional nos estados*: relatório final. Belo Horizonte: Game/UFGM; Fundação Victor Civita, 2011. Disponível em: <http://www.educadores.diaadia.pr.gov.br/arquivos/File/pdf/avaliacao_externa_fvc.pdf>. Acesso em: 27 dez. 2012.

BRUNNER, José Joaquín. Límites de la lectura periodística de resultados educacionales. In: IAIES, Gustavo et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIPE, 2003. p. 67-84.

CANADÁ. Council of Ministers of Education. *About CMEC*. Toronto: CMEC, 2016a. Disponível em: <<http://www.cmec.ca/11/About/index.html>>. Acesso em: 2 dez. 2017.

CANADÁ. Council of Ministers of Education. *Pan-Canadian Assessment Program (PCAP)*. 2016b. Disponível em: <[http://www.cmec.ca/302/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-\(PCAP\)/PCAP-2013/index.html](http://www.cmec.ca/302/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-(PCAP)/PCAP-2013/index.html)>. Acesso em: 2 dez. 2017.

CASASSUS, Juan. Política y metáforas: un análisis de la evaluación estandarizada en el contexto de la política educativa. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 21-46.

CHAPMAN, David; SNYDER, Conrad. Can high stakes national testing improve instruction: re-examining conventional wisdom. *International Journal of Educational Development*, Amsterdam, v. 20, p. 457-474, 2000.

DORN, Sherman. *Accountability Frankenstein: understanding and taming the monster*. Charlotte, NC: Information Age Publishing, 2007.

DOW, Allison. *Exploring the current state of grades 4 to 8 science education in Ontario*. 2015. Dissertação (Mestrado em Educação) – Faculty of Education, Brock University, Ontario, 2015. Disponível em: <<http://dr.library.brocku.ca/handle/10464/6983?show=full>>. Acesso em: 12 abril 2016.

EDUCATIONAL QUALITY AND ASSESSMENT OFFICE. *Highlights of the provincial results*. Assessments of reading, writing and Mathematics, primary division (grades 1-3) and junior division (grades 4-6). English-Language Students, 2014-2015. Ontario: OQRE, 2015. Disponível em: <<http://www.eqao.com/en/assessments/results/communication-docs/provincial-report-highlights-elementary-2015.pdf>>. Acesso em: 12 dez. 2015.

EDUCATIONAL QUALITY AND ASSESSMENT OFFICE. *Highlights of the provincial results*. Assessments of reading, writing and Mathematics, primary division (grades 1-3) and junior division (grades 4-6). English-Language Students, 2015-2016. Ontario: OQRE, April 2016. Disponível em: <<http://www.eqao.com/en/assessments/results/communication-docs/provincial-report-highlights-elementary-2016.pdf>>. Acesso em: 15 mar. 2017.

ESTEBAN, Maria Teresa. Considerações sobre a política de avaliação da alfabetização: pensando a partir do cotidiano escolar. *Revista Brasileira de Educação*, Rio de Janeiro, v. 17, n. 51, p. 573-592, 2012.

EUROPEAN COMMISSION. Eurydice Network. *National testing of pupils in Europe: objectives, organisation and use of results*. Bruxelas: Eurydice Network, 2009.

FAZIO, Xavier; GARCIA, Paulo Sergio; PANIZZON, Debra Lee. Comparative analysis of elementary science teacher preparation in Australia, Brazil, and Canada. In: INTERNATIONAL ORGANIZATION FOR SCIENCE TECHNOLOGY EDUCATION SYMPOSIUM, 18., set. 2008, Esmirna, Turquia. *Proceedings of...* Ankara: Palme & Bookshops, 2008. v. 1, p. 737-745.

FAZIO, Xavier E.; KARROW, Doug. Science takes a back seat: an unintended consequence of prioritizing literacy and numeracy achievement. *Canada Education*, Ontario, v. 56, n. 1, jun. 2013.

FREITAS, Dirce Nei Teixeira de. *A avaliação da educação básica no Brasil*. Campinas: Autores Associados, 2007.

FREITAS, Luiz Carlos de. Caminhos da avaliação de sistemas educacionais no Brasil: o embate entre a cultura da auditoria e a cultura da avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 147-176.

GARCIA, Paulo Sérgio; BIZZO, Nelio. Educational policies and science education in Brazil: a case study. In: EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION, 11., 2015, Helsinki. *Proceedings...* Helsinki: European Science Education Research Association, 2015.

GARCIA, Paulo Sérgio; BIZZO, Nelio; FAZIO, Xavier. Challenges faced by science teachers in distance continuing education. In: INTERNATIONAL ORGANIZATION FOR SCIENCE TECHNOLOGY EDUCATION SYMPOSIUM, 15., 2012, Hammamet, Tunis. *Proceedings...* Hammamet, Tunis: 2012. v. 1, p. 23-33.

GARCIA, Paulo Sérgio; BIZZO, Nelio; FAZIO, Xavier. Desafios da formação contínua a distância para professores de ciências. *Revista Iberoamericana de Educacion a Distancia*, Ecuador, San Cayetano Alto, v. 17, n. 2, p. 37-57, 2014.

GARCIA, Paulo Sérgio; FAZIO, Xavier; PANIZZON, Debra. Formação Inicial de professores de ciências na Austrália, Brasil e Canadá: uma análise exploratória. *Ciência e Educação*, Bauru, v. 17, p. 1-19, 2011.

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

GOUW, Ana Maria; BIZZO, Nelio. A percepção dos jovens brasileiros sobre suas aulas de Ciências. *Educar em Revista*, Curitiba, n. 60, p. 277-292, abr./jun. 2016.

HAGOPIAN, Jesse (Ed.). *More than a score: the new uprising against high-stakes testing*. Chicago: Haymarket Books, 2014.

HORTA NETO, João L. Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o SAEB de 2005. *Revista Iberoamericana de Educación*, Madrid, v. 42, n. 5, p. 1-14, abr. 2007.

HORTA NETO, João L. As avaliações externas e seus efeitos sobre as políticas educacionais: uma análise comparada entre a União e os Estados de Minas Gerais e São Paulo. 2013. Tese (Doutorado em Política Social) – Instituto de Ciências Humanas, Departamento de Serviço Social, Universidade de Brasília, Brasília, 2013.

HYPÓLITO, Álvaro Moreira. A necessária meta-avaliação das políticas de avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 211-227.

IAIES, Gustavo. Evaluar las evaluaciones. In: IAIES, G. et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIEPE, 2003. p. 15-36.

INTERNATIONAL ASSOCIATION FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT. *Trends in International Mathematics and Science Study (TIMSS), 2011: Ontario Report*. Toronto: Educational Quality and Accountability Office, 2012. Disponível em: <http://www.eqao.com/pdf_e/12/TIMSS_Ontario_Report_2011.pdf>. Acesso em: 07 mar. 2015.

KLINGER, Don A.; DELUCA, Chistopher; MERCHANT, Stefan. Canada: the intersection of international achievement testing and educational policy development. In: VOLANTE, L. (Ed.). *The intersection of international achievement testing and educational policy: global perspectives on large-scale reform*. Routledge: Nova York, 2016. p. 140-159.

LIMA, Licínio. O paradigma da educação contábil. Políticas educativas e perspectivas gerencialistas no ensino superior em Portugal. *Revista Brasileira de Educação*, Rio de Janeiro, n. 4, p. 43-59, 1997.

LINGARD, Bob. Policy borrowing, policy learning: testing times in Australian schooling. *Critical Studies in Education*, London, v. 51, n. 2, p. 129-147, 2010.

MADAUS, George; RUSSELL, Michael; HIGGINS, Jennifer. *The paradoxes of high stakes testing: how they affect students, their parents, teachers, principals, schools, and society*. Charlotte: Information Age, 2009.

MONS, Nathalie. Theoretical and real effects of standardised assessment. Bruxelas: Eurydice Network, 2009.

NEW SOUTH WALES. Department of Education. *Literacy and numeracy strategy 2017-2020*. Sydney: New South Wales Government, [s.d.]. Disponível em: <<https://www.det.nsw.edu.au/media/downloads/about-us/literacy-numeracy/literacy-and-numeracy-strategy.pdf>>. Acesso em: 10 fev. 2017.

OBSERVATÓRIO DA EDUCAÇÃO DO GRANDE ABC. *Efeitos IDEB na região do grande ABC sobre o Ensino de Ciências*. São Caetano do Sul: Universidade Municipal de São Caetano do Sul, 2015.

OBSERVATÓRIO DA EDUCAÇÃO DO GRANDE ABC. *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências*. São Caetano do Sul: Universidade Municipal de São Caetano do Sul, 2016.

O'GRADY, Kathryn et al. *Measuring up: Canadian results of the OECD PISA study. The performance of Canada's youth in science, reading, and Mathematics. 2015 First Results for Canadians Aged 15.* Toronto: Council of Ministers of Education, 2016.

OLIVEIRA, Romualdo P. de. A utilização de indicadores de qualidade na unidade escolar ou porque o IDEB é insuficiente. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores.* Florianópolis: Insular, 2013. p. 87-100.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. *PISA 2015 assessment and analytical framework: Science, Reading, Mathematics and Financial Literacy.* Paris: OECD, 2016.

OZGA, Jenny. *Investigação sobre políticas educacionais: terreno de contestação.* Porto: Porto, 2000.

PHILLIPS, David. Learning from elsewhere in education: some perennial problems revisited with reference to British interest in Germany. *Comparative Education*, London, v. 36, n. 3, p. 297-307, 2000.

RAVITCH, Diane. *The death and life of the great American school system: how testing and choice are undermining education.* New York: Basic Books, 2010.

REID, Alan. *Rethinking national curriculum collaboration: towards an Australian curriculum.* Canberra: Department of Education, Science and Training, 2005.

REID, Alan. Is this really a revolution? A critical analysis of the Rudd government's national education agenda. *Curriculum Perspectives*, Berlin, v. 29, n. 3, p. 1-13, 2009.

SANTOS, Lucíola L. A avaliação em debate. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos.* Florianópolis: Insular, 2013. p. 229-245.

SMITH, William. *National testing policies, school practices, and student outcomes: an analysis using data from PISA 2009.* Paris: OECD, 2014. (Education Working Papers).

SONGER, Nancy Butler; RUIZ-PRIMO, Maria Araceli. Assessment and science education: Our essential new priority? *Journal of Research in Science Teaching*, New Jersey, v. 49, n. 6, p. 683-690, 2012.

SOUSA, Sandra Zákia. Avaliação externa e em larga escala no âmbito do Estado brasileiro: interface de experiências estaduais e municipais de avaliação da educação básica com iniciativas do governo federal. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores.* Florianópolis: Insular, 2013. p. 61-85.

SOUSA, Sandra Zákia; PIMENTA, Claudia Oliveira; MACHADO, Cristiane. Avaliação e gestão municipal da educação. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 53, p. 14-36, set./dez. 2012.

STECHEER, Brian M. Consequences of large-scale, high-stakes testing on school and classroom practice. In: HAMILTON, Laura; STECHER, Brian M.; KLEIN, Stephen P. (Ed.). *Making sense of test-based accountability in education*. Santa Monica: Rand, 2002. p. 79-100.

TEDESCO, Juan C. Prólogo. In: IAIES, Gustavo et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIPE, 2003. p. 11-14.

THOMSON, Sue; BORTOLI, Lisa de; UNDERWOOD, Catherine. *TIMSS 2015: a first look at Australia's results*. Melbourne: Australian Council for Educational Research, 2016.

VOLANTE, Louis; BEN JAAFAR, Sonia. Educational assessment in Canada. *Assessment in Education: Principles, Policy & Practice*, London, v. 15, n. 2, p. 201-210, 2008.

PAULO SÉRGIO GARCIA

Professor titular da Universidade Municipal de São Caetano do Sul (USCS). Coordenador do Observatório de Educação do Grande ABC, São Caetano do Sul, São Paulo, Brasil
paulo.garcia@uscs.edu.br

XAVIER FAZIO

Professor associado da Faculty of Education da Brock University, Ontário, Canadá
xfazio@brocku.ca

DEBRA PANIZZON

Professora associada da University of South Australia (UniSA), Adelaide, Austrália
debra.panizzon@unisa.edu.au

NELIO BIZZO

Professor titular da Faculdade de Educação da Universidade de São Paulo (FE/USP), São Paulo, São Paulo, Brasil
bizzo@usp.br

Recebido em: 17 JULHO 2017

Aprovado para publicação em: 10 OUTUBRO 2017

AVALIAÇÃO INSTITUCIONAL NA EDUCAÇÃO BÁSICA: OS DESAFIOS DA IMPLEMENTAÇÃO

**LUIZ SÍVERES
JOSÉ ROBERTO DE SOUZA SANTOS**

RESUMO

O objetivo deste artigo é o de analisar os fatores relacionados com a implementação de avaliações institucionais em redes de ensino e em escolas. O corpus do estudo foi constituído por artigos que discutem resultados de pesquisas acerca da implementação desse tipo de avaliação na educação básica publicados em periódicos nacionais. As evidências presentes nos estudos permitem afirmar que o fator que mais influencia a implementação da avaliação institucional é o desenvolvimento da gestão democrática das escolas. Por outro lado, o trabalho coletivo e a participação da comunidade escolar foram identificados como mudanças decorrentes de tal implementação. Conclui-se que a avaliação institucional é um processo condicionado a outras mudanças na escola e na rede. Não obstante, quando implementada, ela contribui para muitas dessas mudanças. Assim, se se pretende desenvolvê-la, o melhor a fazer é iniciá-la e pô-la em prática, aperfeiçoando-a enquanto é praticada.

PALAVRAS-CHAVE AVALIAÇÃO INSTITUCIONAL • EDUCAÇÃO BÁSICA • GESTÃO DEMOCRÁTICA • IMPLEMENTAÇÃO.

EVALUACIÓN INSTITUCIONAL EN LA EDUCACIÓN BÁSICA: LOS DESAFÍOS DE LA IMPLEMENTACIÓN

RESUMEN

El objetivo del presente artículo es analizar los factores relacionados con la implementación de evaluaciones institucionales en redes de enseñanza y en escuelas. El corpus del estudio se conformó con artículos que discuten resultados de investigaciones acerca de la implementación de este tipo de evaluación en la educación básica publicados en revistas nacionales. Las evidencias presentes en los estudios permiten afirmar que el factor que más influye en la implementación de la evaluación institucional es el desarrollo de la gestión democrática de las escuelas. Por otra parte, el trabajo colectivo y la participación de la comunidad escolar se identificaron como cambios resultantes de tal implementación. Se concluye que la evaluación institucional es un proceso condicionado a otros cambios en la escuela y en la red. Sin embargo, cuando implementada, ella contribuye para muchos de estos cambios. Así, si se pretende desarrollarla, lo mejor a hacer es iniciarla y ponerla en práctica, perfeccionándola mientras es practicada.

PALABRAS CLAVE EVALUACIÓN INSTITUCIONAL • EDUCACIÓN BÁSICA • GESTIÓN DEMOCRÁTICA • IMPLEMENTACIÓN.

INSTITUTIONAL EVALUATION IN BASIC EDUCATION: IMPLEMENTATION CHALLENGES

ABSTRACT

The purpose of this article is to analyze factors related to the implementation of institutional evaluation in education networks and schools. The study corpus consisted of articles that deal with results of research on the implementation of this type of evaluation in basic education, published in national journals. The evidence found in the studies shows that the most influential factor in the implementation of institutional evaluation is the development of democratic management in schools. On the other hand, collaborative work and participation of the school community were identified as changes resulting from such implementation. Institutional evaluation is, therefore, a process dependent on other changes in the school and on the network. However, when implemented, the evaluation contributes to many of these changes. Hence, if we want to develop institutional evaluation, the best way is to start it and improve it as it is put into practice.

KEYWORDS INSTITUTIONAL EVALUATION • BASIC EDUCATION • DEMOCRATIC MANAGEMENT • IMPLEMENTATION.

INTRODUÇÃO

Nos últimos anos, vem crescendo o interesse pela avaliação institucional da escola. Isso pode ser comprovado com o aumento do espaço que esse tipo de avaliação passou a ocupar nos eventos e nos veículos da área de avaliação educacional, bem como nos dispositivos legais e documentos oficiais que tratam da promoção ou da avaliação da qualidade educacional. Um espaço ainda restrito, é bem verdade, quando comparado com outros tipos de avaliação educacional, mas que, de qualquer modo, aponta para uma tendência de defender cada vez mais modelos de avaliação da escola cujo protagonismo seja assumido pela própria escola.

Isso se deve, em grande parte, a uma expectativa que vem sendo criada em torno da avaliação institucional de que ela pode concorrer fortemente para a melhoria da qualidade escolar, uma vez que tem o potencial de envolver a comunidade escolar na avaliação e na promoção da qualidade da escola e tem na função formativa sua razão de existir.

Mas, para que a avaliação institucional de fato se efetive nas escolas e para que ela possa realmente se configurar

como um tipo de avaliação formativa, deverá ser assumida pelas redes de ensino e pelas unidades escolares. Disso decorre que não basta ser uma intenção de teóricos ou de políticas públicas. Também não é suficiente que seja potencialmente boa. É na concretude das redes de ensino e das escolas e na ressignificação que seus atores dela fazem que estão seu potencial e suas limitações. Desse modo, é preciso que se compreendam os fatores que se relacionam com a sua implementação. É este o objetivo deste artigo: elucidar os fatores que afetam a implementação da avaliação institucional em escolas e redes de ensino. Dessa maneira, será possível atuar de um modo mais assertivo no planejamento de novas propostas de avaliação institucional e na correção de rumos daquelas que já estão sendo implementadas.

A maioria dos pesquisadores considera que o processo de mudança educacional compreende três grandes fases: início, implementação e continuidade ou institucionalização (FULLAN, 2009). O início diz respeito à decisão de se adotar ou continuar uma mudança e ao processo que conduz a essa decisão. A implementação, por sua vez, refere-se às primeiras experiências de colocar em prática uma ideia ou um programa. Já a institucionalização nada mais é que a continuação da mudança ou sua rotinização. É, desse modo, uma implementação que não é descontinuada. Essas fases, porém, não são lineares nem independentes; o que ocorre em uma reverbera nas demais.

Sem desconsiderar os desafios próprios que envolvem a adoção de uma mudança educacional por um órgão governamental ou uma instituição de ensino, o que a investigação tem demonstrado é que a decisão de iniciar ou dar prosseguimento a uma inovação não é suficiente para que a mudança ocorra. A mudança efetiva dependerá do que ocorrer na prática real; dependerá de sua implementação.

O presente texto está estruturado da seguinte forma. Na primeira parte, são discutidos os fatores relacionados com a implementação de uma mudança educacional, com base em Fullan (2009). A seguir, são explicitados os procedimentos metodológicos utilizados para o levantamento dos estudos que constituem o *corpus* do trabalho. Na última parte, os

fatores que afetam a implementação de processos de avaliação institucional em escolas ou redes de ensino são analisados e cotejados com aqueles consagrados na literatura especializada no processo de mudança educacional. São discutidas, ainda, as principais mudanças nos sujeitos envolvidos na implementação da avaliação institucional e nas instituições que a implementaram ou tentaram implementá-la, bem como as recomendações apresentadas pelos estudos.

OS DESAFIOS DA IMPLEMENTAÇÃO DE UMA MUDANÇA EDUCACIONAL

Das três fases da mudança educacional, a implementação é a variável principal, uma vez que diz respeito àquilo que as pessoas fazem ou deixam de fazer. Desse modo, mais importante que se concentrar em mudanças formais, como mudanças na legislação ou o desenvolvimento de bons projetos, interessa o que efetivamente ocorre na ação, o uso real (FULLAN, 2009). Nessa perspectiva, reforça Fullan (2009), como é por meio da implementação que os objetivos da mudança são alcançados, ela se torna crítica.

Isso dito, é preciso investigar os fatores que afetam a implementação de uma mudança educacional. Para Fullan (2009), as variáveis básicas são em número reduzido, nove no total. Os fatores com influência causal na implementação podem ser distribuídos em três categorias: as características da inovação ou do projeto de mudança, os papéis locais e os fatores externos. O autor adverte, antes, que a lista é simplificada e que os fatores não podem ser pensados como conjuntos isolados; eles formariam uma espécie de sistema de variáveis em constante interação.

No conjunto das características do projeto de mudança estão a necessidade, a clareza, a complexidade e a qualidade ou praticidade da inovação. A necessidade diz respeito à questão das prioridades e de determinar necessidades comuns entre os atores envolvidos na mudança. Isso significa que os envolvidos precisam compreender e compartilhar as finalidades da mudança, seu significado, sob pena de sua implementação ficar comprometida.

Para que a mudança de fato ocorra, portanto, não basta que ela seja apresentada (ou imposta) e que seu significado seja compreendido apenas por aqueles que a propõem. É preciso que os agentes da mudança estejam convencidos de sua necessidade, com tal intensidade que aceitem superar a insegurança que envolve mudar. Diante disso, Marris (1975, p. 166 apud FULLAN, 2009, p. 32)¹ problematiza,

1 MARRIS, P. *Loss and change*. New York: Anchor Press; Doubleday, 1975.

Ninguém resolve a crise de reintegração em nome de outra pessoa. Cada tentativa de pré-esvaziar o conflito, argumentar, protestar com um planejamento racional, somente será abortiva: por mais razoáveis que sejam as mudanças propostas, o processo de implementá-las ainda deve permitir que o impulso de rejeição se liberte. Quando aqueles que têm poder para manipular mudanças agem como se apenas tivessem que explicar, e quando suas explicações não são aceitas, desdenham da oposição como ignorância ou preconceito, eles expressam um profundo desprezo pelo significado das vidas das outras pessoas. Pois os reformistas já assimilaram essas mudanças para seus fins, e fizeram uma reformulação que lhes faz sentido, talvez por meses ou anos de análise e debates. Se eles negarem aos outros a chance de fazer o mesmo, eles os tratarão como marionetes penduradas pelos fios de suas próprias concepções.

Muitas das tentativas de mudança se dão desse modo, pela imposição, sem que os envolvidos tenham a oportunidade de captar o seu significado – talvez porque os propositores pressuponham que o fato de eles já terem compreendido a importância da mudança seria o suficiente para que os demais a implementassem. Apesar disso, Fullan (2009) destaca a necessidade de estar atento para que os indivíduos envolvidos na mudança tenham a oportunidade de atravessar suas “zonas de incerteza”, numa espécie de “conservadorismo dinâmico”, a fim de assimilarem a importância da mudança. Desse modo, se se deseja que a mudança efetivamente ocorra, deve-se dar atenção ao seu significado subjetivo.

A fim de aprofundar essa questão, Fullan (2009) propõe uma diferenciação entre os termos reestruturação e recultu-

ração. Enquanto a reestruturação, por ser fruto da decisão de alguns, pode ser feita por decreto e repetida muitas vezes, a reculturação tem a ver com a mudança de crenças, ideias e hábitos de um determinado grupo, por isso é muito mais problemática e não é fruto de uma determinação de outrem. Para a mudança educacional, mais importante que a reestruturação é a reculturação dos agentes da mudança, o que exige que captem o significado da mudança, de maneira a alterarem seu comportamento e suas crenças. Essa é a razão pela qual é tão difícil alcançar a mudança.

A perspectiva do autor é de que a crença e a compreensão são os princípios básicos da mudança; são “a base para alcançar uma reforma duradoura” (FULLAN, 2009, p. 44). Sem que se alcancem a base teórica da mudança, suas finalidades e seu sentido, a implementação será superficial e, certamente, não perdurará.

O problema para a implementação, então, não é apenas que os professores ‘aprendam a fazer’, mas que eles aprendam a base teórica [...] o conhecimento ausente sobre *por que* estão fazendo o que estão fazendo. (MCLAUGHLIN; MILTRA, 2000, p. 10 apud FULLAN, 2009, p. 44, grifo do original)²

2 MCLAUGHLIN, M; MILTRA, D.
Theory-based change and change-based theory: going deeper, going broader. Stanford, CA: Stanford University, 2000. Artigo inédito.

É nesse sentido que o significado da mudança para os envolvidos na implementação é tão decisivo para o seu sucesso. Por isso também que aqueles interessados na mudança precisam estar atentos a esse significado, inclusive com vistas a potencializar a apreensão do sentido que a mudança pretende ter. Chama-se a atenção, porém, para que a instituição caminhe na direção de um “significado compartilhado”. Ainda que a apreensão de um significado seja um ato individual, seu valor real para o alcance dos resultados está no significado compartilhado alcançado por todo o grupo (FULLAN, 2009). Esse alerta é fundamental quando se pensa na implementação de uma política/processo de avaliação institucional, para a qual o trabalho coletivo é decisivo, conforme será visto adiante.

Os demais fatores reunidos no conjunto das características da mudança podem ser assim resumidos. A clareza da

inovação proposta é um ponto crítico na implementação, pois se os atores escolares não tiverem clareza dos objetivos da mudança, dos meios para pô-la em ação, dos processos para avaliá-la, podem simplesmente não a implementar. Já a complexidade refere-se ao nível de exigência dos indivíduos para por em ação uma mudança. Não basta, contudo, que a mudança seja simples, que exija pouco esforço, ainda mais porque talvez mudanças demasiado simples não tragam ganhos consideráveis (FULLAN, 2009). Por fim, a qualidade e praticidade da mudança geralmente estão relacionadas com o planejamento adequado da implementação. Ocorre que, por vezes, os propositores da mudança concentram-se somente em sua adoção e pouca atenção é dada à implementação, comprometendo a qualidade do projeto, sobretudo pela carência de tempo para prepará-lo adequadamente. Desse modo, o planejamento da implementação pode ser decisivo para que a mudança se dê efetivamente.

As características locais que afetam a fase da implementação englobam a atuação dos órgãos centrais dos sistemas e redes de ensino, da comunidade, da direção escolar e dos professores.

O primeiro fator é o apoio dos gestores dos sistemas ou redes de ensino. Esse apoio é central para mudanças nas práticas do sistema ou da rede. Mas não se limita a apoio verbal; há que se demonstrar esse apoio com ações efetivas. O apoio do órgão central é ainda mais fundamental quando se trata de uma mudança mais ampla, que atinge a própria cultura organizacional da rede, como é o caso, por exemplo, de uma política de avaliação institucional.

Outro fator decisivo na implementação é a parceria com a comunidade, nomeadamente por intermédio do conselho escolar. De acordo com Fullan (2009), praticamente em todas as escolas em que se observaram melhoras sólidas, a parceria das unidades escolares com os pais foi notória.

Não menos relevante para se empreenderem as mudanças é o apoio ativo da direção escolar. Para que esse apoio seja efetivo, é importante que o diretor se esforce para compreender a perspectiva dos professores e os ajude a captar o significado da mudança (FULLAN, 2009). Um dos

modos de se fazer isso é por meio de um trabalho coletivo e dialogado.

Esse trabalho coletivo é ainda mais evidenciado quando se trata do papel dos professores na mudança. É que “a qualidade das relações de trabalho entre os professores está bastante relacionada com a implementação” (FULLAN, 2009, p. 96). O trabalho coletivo dos profissionais da escola está ancorado na ideia de uma “aprendizagem social”. Esse tipo de aprendizagem, diz o autor, é fundamental no processo de mudança. Nas escolas onde ela é desenvolvida, as chances de uma implementação de sucesso aumentam consideravelmente. A perspectiva proposta por Fullan (2009) é a de que as escolas se tornem “comunidades profissionais de aprendizagem”, o que pressupõe o trabalho conjunto não só dos professores, mas também entre estes e a equipe diretiva.

O conjunto dos fatores externos que afetam a implementação de mudanças educacionais abarca desde órgãos governamentais até universidades e fundações. Embora não haja garantia de que os projetos emanados dessas agências serão efetivamente implementados, não parecem restar dúvidas de que o tipo de relação estabelecida entre estas e as secretarias de educação e unidades escolares exerce impacto na implementação. A qualidade dessas relações é crucial para sustentar as mudanças.

Além dos fatores enumerados acima, um desafio adicional para a implementação de uma mudança é que maioria das mudanças educacionais fracassa (FULLAN, 2009). Com isso, os sujeitos que se envolvem nelas vão acumulando um histórico de fracasso. Como os significados de uma experiência influenciam experiências futuras, esses sujeitos se tornam cada vez mais céticos quanto a mudanças propostas, independentemente de seus méritos. Por outro lado, Fullan (2009) chama a atenção para a importância de se aprender à medida que a mudança vai sendo experimentada, de sorte que a ação é mais relevante para a mudança que planejamentos bem elaborados. Isso porque, para que mudem suas crenças (fundamental para uma mudança sustentada), as pessoas precisam de experiências que lhes permitam formular novos significados.

METODOLOGIA

O estudo em tela resultou da identificação e análise dos artigos sobre implementação de avaliações institucionais na educação básica, publicados em periódicos brasileiros e revisados por pares no período de 2006 a 2016.

Para o levantamento dos trabalhos recorreu-se a três repositórios: o Portal de Periódicos da Capes, o SciELO e o Portal Educ@, da Fundação Carlos Chagas. Estes últimos foram escolhidos por indexarem os principais periódicos que abordam a temática da avaliação educacional no Brasil.

A fim de obter resultados mais completos, foram elaborados três conjuntos de descritores, utilizados no levantamento bibliográfico. Um, com termos mais amplos, cujo objetivo era permitir a busca de trabalhos que tratassem do tema da avaliação institucional ou da autoavaliação institucional, especificamente. Outro, contendo termos que, quando cruzados com os do primeiro conjunto, permitissem refinar a busca para o nível de ensino de interesse do estudo, a educação básica. E um terceiro conjunto, com termos já combinados, com vistas a possibilitar buscas a partir de termos exatos, o que resultaria em uma pesquisa ainda mais refinada.

QUADRO 1 - Descritores para levantamento de artigos acerca da avaliação institucional na educação básica

DESCRITORES		
CONJUNTO A	CONJUNTO B	CONJUNTO C
<ul style="list-style-type: none">• Avaliação institucional• Autoavaliação institucional• Auto-avaliação institucional	<ul style="list-style-type: none">• Escola(s)• Escola(s) pública(s)• Ensino fundamental• Educação básica• Instituições escolares (singular e plural)	<ul style="list-style-type: none">• Autoavaliação de/das escolas (singular e plural)• Autoavaliação institucional de/das escolas (singular e plural)• Avaliação institucional de/das escolas (singular e plural)• Avaliação de instituições escolares• Avaliação interna de/da escola (singular e plural)• Avaliação institucional participativa

Fonte: Elaboração dos autores.

No Portal de Periódicos da Capes, utilizando-se todos os descritores acima (como termos exatos ou não), embora tenham retornado centenas de artigos publicados em

periódicos revisados por pares, apenas sete referiam-se à avaliação institucional na educação básica (BETINI, 2010; DALBEN, 2010; FREITAS, 2007; MIRANDA; RODRIGUES, 2010; SÁ, 2009; SORDI, 2010; SORDI; LÜDKE, 2009a).

Por sua vez, no SciELO e no portal Educ@, foram localizados os seguintes trabalhos: Sordi (2006), Freitas (2007), Sá (2009), Sordi e Lüdke (2009b), Afonso (2010), Ribeiro e Gusmão (2010), Silva (2010), Sordi (2010), Brandalise e Martins (2011), Sordi (2011), Mendes e Sordi (2012), Sordi (2012), Souza e Dittrich (2012), Pacheco (2014), Brandalise (2015), Dalben e Almeida (2015), Lima (2015) e Mendes *et al.* (2015). Quinze deles, novos.

O levantamento demonstrou que há um número bastante reduzido de trabalhos que abordam a avaliação institucional na educação básica publicados no Brasil (22 trabalhos), especialmente quando comparado com o número de trabalhos acerca desse tema na educação superior, o que comprova a desproporcionalidade da produção científica nos dois níveis de ensino.

Identificados os trabalhos que tratavam da temática da avaliação institucional na educação básica, foram selecionados somente aqueles que discutiam resultados de pesquisas sobre a implementação desse tipo de avaliação em escolas ou redes de ensino. Para isso, foram lidos todos os trabalhos, na íntegra. Após essa etapa, chegou-se à seguinte relação: Sordi (2006, 2010, 2011, 2012), Betini (2010), Dalben (2010), Ribeiro e Gusmão (2010), Silva (2010), Brandalise e Martins (2011), Mendes e Sordi (2012), Souza e Dittrich (2012), Brandalise (2015) e Mendes *et al.* (2015).

Uma vez selecionados, os trabalhos foram analisados, primeiramente, quanto aos seus objetos e problemas de pesquisa. Em seguida, foram tratados em conjunto, o que permitiu a análise dos principais potencializadores e desafios da implementação de políticas e propostas de avaliação institucional em escolas e redes de ensino, bem como das mudanças decorrentes do processo de avaliação evidenciadas pelas pesquisas e das recomendações presentes nos estudos, categorias de análise que emergiram do material examinado.

A IMPLEMENTAÇÃO DA AVALIAÇÃO INSTITUCIONAL NA EDUCAÇÃO BÁSICA

A maioria dos trabalhos analisados aborda a implementação da Avaliação Institucional Participativa (AIP) em escolas públicas de Campinas-SP (BETINI, 2010; DALBEN, 2010; MENDES; SORDI, 2012; SORDI, 2006, 2010, 2011, 2012) com enfoques variados – a gestão da política, as reuniões de negociação entre os gestores do nível central e os atores da escola, o funcionamento das Comissões Próprias de Avaliação (CPAs) e a implementação em escolas da rede municipal ou estadual de ensino. Já Ribeiro e Gusmão (2010) discutem a implementação dos Indicadores de Qualidade na Educação (Indique) em escolas e secretarias de educação. Silva (2010), por sua vez, analisa a participação dos pais em iniciativa de avaliação de uma secretaria municipal de educação em quatro escolas. O trabalho de Mendes *et al.* (2015) discute propostas que emergiram de grupos de estudo formados por pesquisadores de universidades públicas, profissionais das escolas e das regionais de ensino, no âmbito de um projeto de pesquisa desenvolvido em parceria por duas universidades brasileiras e uma portuguesa. Dois trabalhos abordam o programa de avaliação institucional da rede estadual de ensino do Paraná (BRANDALISE; MARTINS, 2011; BRANDALISE, 2015). E um artigo, o de Souza e Dittrich (2012), estuda a implementação de uma proposta de avaliação elaborada por iniciativa da própria escola, apoiada, em seguida, por profissionais de uma universidade pública.

Como se vê, as propostas a serem implementadas foram variadas. Contemplam desde a existência de comissões de avaliação, indicadores pré-determinados, questionários a serem respondidos pelos membros da comunidade escolar, presença ou não de assessoria externa, até medidas longitudinais. De modo geral, contudo, elas são bastante semelhantes quanto aos princípios que consagram, como participação, diálogo entre sujeitos e avaliações, trabalho coletivo, qualidade multidimensional e avaliação do processo.

Os diferentes objetos e problemas de pesquisa abordados nos estudos, que contemplam desde a avaliação institucional enquanto política de Estado, até experiências individuais e

isoladas; os olhares diversos acerca da implementação, seja mais voltado para a gestão central, seja para o interior das escolas; os diferentes sujeitos das pesquisas (gestores centrais, técnicos de secretarias de educação, diretores escolares, orientadores e coordenadores pedagógicos, professores, funcionários, estudantes e pais); os métodos utilizados (estudos de caso, estudos exploratórios, ou ainda de modo colaborativo com pesquisadores de universidades públicas); e os múltiplos instrumentos de pesquisa empregados, a despeito do número reduzido de trabalhos, formam um conjunto de evidências consistente acerca da implementação da avaliação institucional na educação básica.

FATORES QUE COLABORAM PARA A IMPLEMENTAÇÃO DA AVALIAÇÃO INSTITUCIONAL

Na escola, os aspectos que mais favoreceram a implementação da avaliação institucional foram a participação da comunidade escolar, a gestão democrática e o trabalho coletivo. Onde a gestão democrática se encontra mais desenvolvida, onde há compromisso com o trabalho coletivo e participação efetiva da comunidade escolar, as chances de uma proposta de avaliação institucional prosperar parecem aumentar consideravelmente; é o que aponta a maioria dos estudos.

Fora dela, o aspecto essencial foi o envolvimento e apoio da alta gestão, o que resulta em continuidade da política de avaliação institucional e alocação de verbas para efetivá-la. Além desses fatores, a assessoria externa e a existência de modelos disponíveis mostraram-se potentes.

O quadro a seguir resume os fatores que mais colaboraram para a implementação da avaliação institucional em escolas de educação básica e redes de ensino, de acordo com as pesquisas. Esses fatores foram reunidos em cinco eixos: cultura de participação, apoio do órgão central, modelos e apoio externos, cultura de avaliação na escola e condições de trabalho.

QUADRO 2 – Aspectos que favorecem a implementação da avaliação institucional em escolas ou redes de ensino

EIXO	PONTOS OBSERVADOS
Cultura de participação	<ul style="list-style-type: none"> • Compromisso com a participação • Gestão democrática da escola • Trabalho coletivo • Liderança da gestão • Grande adesão de alunos
Apoio do órgão central	<ul style="list-style-type: none"> • Envolvimento e apoio da alta gestão • Continuidade do processo • Alocação de verbas para o desenvolvimento do processo
Modelos e apoio externos	<ul style="list-style-type: none"> • Assessoria externa • Existência de uma proposta já desenhada pelo órgão central • Embasamento em modelos já existentes • Proposta de fácil entendimento e operacionalização
Cultura de avaliação na escola	<ul style="list-style-type: none"> • Auto-organização da escola • Institucionalização de equipe de avaliação • Confiança na avaliação • Formação continuada com base nas necessidades da escola • Articulação da avaliação institucional com o Projeto Político-Pedagógico (PPP) • Mapeamento dos problemas da instituição • Existência da autoavaliação antes de a rede implementá-la • Familiaridade com a investigação científica/formação em pesquisa • Adesão voluntária
Condições de trabalho	<ul style="list-style-type: none"> • Escolas de tempo integral • Dedicção exclusiva dos professores

Fonte: Elaboração dos autores com base nos artigos sobre implementação da avaliação institucional na educação básica, referenciados no corpo do texto.

O primeiro destaque a ser dado é que são muitos e variados os fatores que podem contribuir para o uso real da avaliação institucional nas escolas. Contemplam desde aspectos da cultura escolar, como trabalho coletivo, participação e confiança na avaliação, até o apoio externo, como alocação de verbas, assessoria externa e disponibilidade, clareza e praticidade da proposta.

Em segundo lugar, embora sejam muitos os fatores que potencializam a mudança, alguns deles parecem ser mais determinantes. É o caso da cultura de participação da escola e do apoio externo, do órgão central do sistema de ensino ou de outras instituições.

Tamanha é a importância de uma gestão democrática e do trabalho coletivo na escola para o desenvolvimento de uma prática de avaliação ancorada na participação e na negociação que Sordi (2012) chega a concluir que, para tanto, as “condições anteriores de democracia na gestão do PP [pro-

jeto pedagógico] da escola precisam estar asseguradas, ainda que não necessariamente ocorram isentas de contradição.” (SORDI, 2012, p. 494).

Juntamente com a gestão democrática e o trabalho coletivo, o apoio, a participação e a liderança da gestão da escola são fatores fundamentais para a implementação da avaliação institucional, como comprovam os estudos de Betini (2009, 2010) e Silva (2010), além de pesquisas realizadas fora do Brasil (CHAPMAN; SAMMONS, 2013).

Por sua vez, muitos dos estudos analisados concluem que o apoio e envolvimento da alta gestão foram determinantes para o sucesso da proposta de avaliação (MENDES; SORDI, 2012; RIBEIRO; GUSMÃO, 2010; SORDI, 2012). Além disso, o fato de haver uma proposta de avaliação disponível facilitou a implementação (BRANDALISE, 2015; BRANDALISE; MARTINS, 2011; MENDES; SORDI, 2012). Outro fator importante (e controverso) é a presença de assessoria externa. Embora tenha ficado demonstrado por algumas pesquisas que o assessor externo potencializou a implementação (MENDES et al., 2015; MENDES; SORDI, 2012; RIBEIRO; GUSMÃO, 2010), quando este se ausenta ou não está disponível, as escolas parecem ter dificuldade em conduzir, sozinhas, seu processo de avaliação.

Ainda que somente com base nos estudos aqui analisados não seja possível avaliar a atuação dos assessores externos, propostas de avaliação que tenham como uma de suas características a assessoria externa deveriam ser planejadas considerando que o apoio externo poderá cessar em algum momento. Assim, embora inicialmente o assessor externo possa apoiar a comunidade escolar no desenvolvimento da avaliação – contribuindo, sobretudo, com conhecimento especializado em avaliação e pesquisa – e embora possa, inclusive, assumir a coordenação do processo de avaliação, como proposto por Cousins e Earl (1992), uma de suas funções deveria ser contribuir com a formação da equipe de avaliação da escola. Desse modo, *à medida* que fossem desenvolvidas e aperfeiçoadas habilidades de pesquisa/avaliação na instituição, alguns membros da comunidade local poderiam assumir a coordenação de novos projetos de avaliação, conforme propõem os autores citados. No entanto, isso não significa

o abandono do assessor externo. Seu papel poderia evoluir para uma postura de apoio e consultoria no decorrer do tempo, estando disponível para consulta, caso seja necessário, como sugere Nevo (1997).

Por fim, é interessante notar que muitos dos aspectos que, segundo os estudos, afetam a implementação da avaliação institucional foram elencados por Fullan (2009) como fatores que interferem na fase da implementação de uma mudança educacional, a saber, a compreensão e compartilhamento das finalidades da mudança, a clareza e a complexidade da inovação proposta, o apoio do órgão central com ações efetivas, a parceria com a comunidade escolar, o apoio e o envolvimento da direção escolar, e o trabalho coletivo e dialogado, o que dá maior robustez aos dados aqui analisados.

PRINCIPAIS DESAFIOS OBSERVADOS NA IMPLEMENTAÇÃO DA AVALIAÇÃO INSTITUCIONAL

De acordo com os estudos, os fatores que mais dificultam a implementação da avaliação institucional dizem respeito a aspectos relacionados à não efetivação da gestão democrática nas escolas, materializada na pouca participação dos atores escolares, sobretudo, as famílias. Tendo em vista que a avaliação institucional deve ser participativa, conforme preconizam os estudos, sem efetiva participação, sua consecução fica prejudicada. Desse modo, o desafio da avaliação institucional é também o desafio da gestão democrática nas escolas. Não obstante, como assinala Dalben (2008), se é na prática da gestão democrática que se encontra a principal potencialidade da avaliação institucional, investir na avaliação participativa pode ser um caminho para a gestão democrática.

Além desse fator, no âmbito da escola, é possível destacar como pontos críticos para o desenvolvimento da avaliação institucional a falta de uma cultura de avaliação, a percepção da avaliação apenas como cumprimento de uma exigência burocrática, a resistência de alguns profissionais a qualquer tipo de avaliação e a falta de tempo e espaço para realização da avaliação.

Na esfera da rede, sobressaem-se a falta de uma cultura de avaliação, as mudanças na gestão central e a descontinuidade da política ou proposta de avaliação, em função de sua não institucionalização, ficando restrita como opção de governos e não como política de Estado.

O quadro a seguir resume os pontos críticos para o desenvolvimento de propostas de avaliação institucional evidenciados nas pesquisas. Estes foram reunidos em cinco eixos: cultura de participação, apoio do órgão central, cultura de avaliação, condições de trabalho e modelos e apoio externos.

QUADRO 3 – Aspectos que dificultam a implementação da avaliação institucional em escolas ou redes de ensino

EIXO	PONTOS OBSERVADOS
Cultura de participação	<ul style="list-style-type: none"> • Gestão centralizadora • Não efetivação da participação da comunidade escolar, sobretudo dos familiares • Não participação dos pais nas decisões • Pouco aproveitamento da presença dos pais • Baixa participação dos alunos • Participação restrita aos profissionais • Maior poder dos professores quando comparado aos demais segmentos da comunidade escolar • Falta de autoridade dos segmentos não profissionais
Apoio do órgão central	<ul style="list-style-type: none"> • Mudanças na gestão central • Descontinuidade da política ou proposta de avaliação (não institucionalização) • Equipe central reduzida • Falta de confiança nos gestores centrais • Falta de mecanismos de supervisão da política • Dificuldade de compreensão das regionais de ensino quanto ao seu papel na implementação da política • As escolas não serem ouvidas na propositura da política • Divulgação insuficiente da proposta • Falta de retorno dos resultados das avaliações • Recorrência dos problemas diagnosticados
Cultura de avaliação	<ul style="list-style-type: none"> • Falta de uma cultura de avaliação na rede e nas escolas • Avaliação como cumprimento de uma exigência burocrática • Resistências de alguns profissionais a qualquer tipo de avaliação • Falta de iniciativa da escola de se autoavaliar • Pouca participação dos diretores nas discussões pedagógicas • Dificuldade em estabelecer prioridades • Não referência ao PPP
Condições de trabalho	<ul style="list-style-type: none"> • Falta de tempo e espaço para realização da avaliação • Rotatividade docente • Excesso de burocracia na escola • Rotatividade da equipe gestora • Falta de professores
Modelos e apoio externos	<ul style="list-style-type: none"> • Necessidade de assessoria externa • Dificuldade na compreensão dos indicadores presentes no modelo proposto

Fonte: Elaboração dos autores com base nos artigos sobre implementação da avaliação institucional na educação básica, referenciados no corpo do texto.

Quando analisados os pontos observados no eixo cultura de participação, é notória e flagrante a pouca (ou a não) participação dos segmentos não docentes. Pais, alunos e outros profissionais da escola parecem pouco participar da gestão e avaliação da instituição, o que é um entrave em uma proposta de avaliação que, para lograr êxito, depende da participação plural.

Estudo realizado por Itamar Silva (SILVA, 2010) concluiu que alguns fatores, como falta de domínio da leitura/escrita dos familiares e falta de confiança dos gestores na disposição dos pais em participar, prejudicaram sua participação no processo de avaliação institucional realizado pelas escolas. Por outro lado, mesmo quando a participação se dá, fica restrita a apenas uma das etapas da avaliação, responder aos questionários enviados pela escola, numa espécie de pesquisa de satisfação. Não são envolvidos, por exemplo, nos processos decisórios.

Dalben (2010) também concluiu em seu estudo que os pais não participam das decisões na escola. Mas suas conclusões são ainda mais contundentes. Após pesquisa que investigou as reais possibilidades e limitações para a implementação de uma AIP em uma escola estadual de ensino fundamental, ele constatou que a participação não se efetivava no ambiente escolar e elencou as inúmeras barreiras para isso: perda da noção de escola como bem público; predominância dos profissionais da escola nos órgãos colegiados, concorrendo para a autoexclusão dos alunos, pais e funcionários; inadequação dos horários de funcionamento dos órgãos colegiados às possibilidades dos familiares; falta de envolvimento dos professores; ideia de que os pais não se interessam pela vida escolar dos filhos; restrição da participação dos pais à execução de tarefas, sem ter parte na tomada de decisões; posição de espera da comunidade escolar, já que não se vê como participante da implantação de possíveis alterações e, conseqüentemente, como responsável por elas, dentre outras.

Embora as famílias sejam o segmento que menos participa na gestão e avaliação da escola, consoante aos estudos, a limitada participação dos alunos também tem sido

documentada (COSTA, 2012; RIBEIRO; GUSMÃO, 2010; SORDI, 2012). O mesmo se pode dizer dos funcionários da escola (BETINI, 2010; SORDI, 2012).

Chama a atenção a questão de poder que envolve a relação dos profissionais docentes com os demais membros da comunidade escolar. Em primeiro lugar, o coletivo da escola é, por vezes, entendido como sendo composto apenas pela direção e os professores (BETINI, 2010). Em segundo lugar, uma vez que os professores compreendem melhor os procedimentos e questões que compõem os instrumentos de avaliação, isso lhes dá maior poder frente aos demais membros da comunidade escolar (RIBEIRO; GUSMÃO, 2010). Em terceiro lugar, o necessário compartilhamento do saber dos professores com os demais membros da comunidade escolar esbarra em alguns desafios, como a falta de hábito dos professores em compartilhar questões internas das escolas com os familiares e o preconceito daqueles com a participação da comunidade, o que resulta, inclusive, em resistência à avaliação, conforme concluem as pesquisas.

Em estudo dedicado a analisar as relações (de poder) dos profissionais da educação com a comunidade escolar, mediante processos de avaliação institucional, Sordi e Lüdke (2009a) consideram que essa relação é um dos desafios na avaliação da qualidade do projeto pedagógico da escola. Elas também concluíram que há uma dificuldade dos profissionais docentes em aceitarem a participação dos segmentos não profissionais e acolherem seus saberes próprios. “Muito há que avançar para que profissionais da educação se acostumem a ouvir o que falam os demais atores a respeito da escola que querem ajudar a construir” (SORDI; LÜDKE, 2009a, p. 219). Somam-se a isso a falta de autoridade desses segmentos para se pronunciarem acerca da qualidade da escola e o fato de não se sentirem no direito de avaliá-la, acrescentam as autoras.

Diante desse quadro, elas propõem que os profissionais da educação sejam os articuladores dos atores da escola em um processo de avaliação institucional. Mas sempre com o cuidado de estabelecerem um diálogo verdadeiro com todos os participantes da avaliação da qualidade da escola, o que

implica não deslegitimar seus saberes e experiências. Segundo as autoras,

Parece ser indiscutível que os profissionais da educação, por dever de ofício, exercerão com mais propriedade a articulação dos atores da escola. No entanto, ao fazer essa articulação, convém que não se comportem como especialistas que usam o diálogo como artifício de cooptação, sem nenhum esforço de entendimento e compromisso com a incorporação dos argumentos trazidos pelos segmentos tidos como coadjuvantes na definição da concepção de qualidade que a escola precisa possuir. (SORDI; LÜDKE, 2009a, p. 224)

Toda essa discussão a respeito das relações que são estabelecidas entre os membros da comunidade escolar e de sua participação efetiva na gestão da escola e na determinação de sua qualidade é, na verdade, parte do próprio debate sobre o tipo de gestão que impera nas unidades escolares. Se se trata de uma gestão mais centralizadora ou, pelo contrário, de uma gestão que se estrutura na participação plural de todos os atores da escola. Isso demonstra mais uma vez a relação umbilical que há entre a avaliação institucional (participativa) e a gestão democrática da escola.

Essa relação é tão substancial que Betini (2009), em um dos estudos pioneiros acerca da implementação da AIP em escolas públicas de Campinas-SP, conclui que esse tipo de gestão é determinante para o desenvolvimento da avaliação participativa. Textualmente,

Constata-se que sem estar preparado para discutir politicamente com a escola, sem “abrir” as tarefas da escola para a participação de todos, sem deixar que os objetivos e metas, assim como as estratégias, sejam determinadas e definidas coletivamente, sem que a direção se coloque como dirigente “político”, não há possibilidade de haver Avaliação Participativa, pois essa não depende apenas de instrumentação técnica. (BETINI, 2009, p. 324)

Costa (2012), em estudo sobre as formas de participação dos estudantes na determinação da qualidade da escola no

contexto da AIP, apresenta conclusão semelhante: a gestão democrática é uma condição que antecede ou auxilia a avaliação com participação. A diferença está em que a autora, apesar de reconhecer a centralidade da gestão democrática para a avaliação institucional participativa, não condiciona esta àquela. Ou seja, admite que, mesmo que a gestão democrática não esteja consolidada, pode-se iniciar o processo de avaliação participativa, talvez suspeitando que esta, de certo modo, também seja auxiliar da gestão democrática, e não apenas vice-versa.

Outro fator de destaque no conjunto daqueles que mais afetam a implementação da avaliação institucional é a cultura de avaliação, nas escolas e na rede. Conforme discutido acima, um dos dificultadores da mudança é que ela envolve transformações nas concepções e nos comportamentos das pessoas, uma reculturação (FULLAN, 2009). Isso significa que não basta as pessoas envolvidas na mudança aprenderem tecnicamente como pô-la em prática; mais importante é que entendam (e compartilhem) conceitualmente a finalidade da mudança, o que exigirá que modifiquem suas crenças. Se não alcançarem e compartilharem as concepções que sustentam a mudança, sua base teórica e suas finalidades, a implementação certamente será superficial. É por isso, provavelmente, que uma fraca cultura de avaliação acaba por ter um efeito de retração não apenas na avaliação institucional, mas em toda e qualquer iniciativa de avaliação. É que ela, a cultura de avaliação, está na base da mudança proposta, a avaliação institucional. É fundamental destacar que a reculturação aqui referida, ainda que se dê individualmente, deve ser pensada também em termos de aprendizagem organizacional (COUSINS; EARL, 1992), de maneira que uma cultura de avaliação possa ser compartilhada por todos os atores da escola. Isso sugere, mais uma vez, a importância de um trabalho coletivo e dialogado, inclusive como meio para a aprendizagem social de novas crenças, ideias e comportamentos.

O compartilhamento de uma cultura de avaliação para a melhoria da escola poderia ajudar a superar, por exemplo, um dos grandes desafios da avaliação institucional, a noção da avaliação apenas como burocracia, observada em alguns

estudos (BRANDALISE; MARTINS, 2011; MACEDO, 2009). Além disso, uma consistente cultura de avaliação nas escolas e na rede poderia ajudar a minimizar os efeitos provocados pelas mudanças na gestão central e pela descontinuidade da política ou proposta de avaliação institucional, dois dos maiores desafios para a implementação e continuidade da avaliação institucional. É que, havendo na rede de ensino uma reculturação, mesmo com mudanças de governo, a avaliação institucional poderia perdurar, já que teria sido incorporada à prática das escolas. Teria sido institucionalizada.

Fullan (2009) já havia alertado quanto ao desafio de institucionalizar uma mudança educacional. Diante disso, havia sugerido que a mudança fosse incorporada na estrutura da instituição; que o seu significado fosse assimilado por gestores e professores, de modo a gerar compromisso com ela; e que o planejamento da implementação contemplasse estratégias para a continuidade da mudança, sobretudo com vistas a incorporar novos atores.

Outro fator que merece relevo são as condições de trabalho disponíveis na escola, nomeadamente tempo e espaço para realização da avaliação institucional. As condições de trabalho, precárias em muitas das escolas públicas, sem dúvida são um grande desafio para fazer prosperar mudanças que exigem tempo e espaço adequados para sua consecução, como é o caso da avaliação institucional. Além da falta desses elementos, a rotatividade docente e da equipe gestora e a falta de professores inibem a implementação da avaliação institucional. Por outro lado, conforme apresentado no Quadro 2, escolas de tempo integral, com professores com dedicação exclusiva, favorecem sua realização.

Como a garantia dessas condições de trabalho depende em grande medida do poder público, o apoio do órgão central do sistema de ensino se mostra ainda mais decisivo para a implementação da avaliação institucional. Vale ressaltar, contudo, que a própria avaliação institucional, ainda que realizada inicialmente sem as condições ideais, pode ser um dos caminhos para negociar com o poder público a garantia das condições mínimas necessárias para o bom funcionamento da escola e o cumprimento de seu dever de ensinar com

qualidade, consoante Bondioli (2013) e Freitas (2005). Afinal, a falta de professores não é um desafio apenas para a realização de um processo de avaliação institucional; é crítico para o próprio funcionamento da escola como instituição responsável pelo ensino.

MUDANÇAS PROVOCADAS PELA AVALIAÇÃO INSTITUCIONAL

Os estudos identificaram vários ganhos propiciados pela implementação da avaliação institucional, tanto para as instituições quanto para os sujeitos envolvidos na mudança. Destacam-se, na instituição, a superação da falta de trabalho coletivo, os avanços na participação da comunidade escolar, o fortalecimento dos órgãos colegiados da escola, o aumento da capacidade de negociação da escola e a melhoria da formação continuada. Para os sujeitos, o ganho mais significativo diz respeito à mudança na concepção de avaliação.

O quadro a seguir resume os avanços e as aprendizagens evidenciados pelas pesquisas, relacionados com a implementação da avaliação institucional em escolas ou redes. Essas mudanças foram categorizadas em cinco eixos: trabalho coletivo, negociação, gestão democrática, concepções e usos da avaliação, e relações na escola.

QUADRO 4 - Mudanças relacionadas com a implementação da avaliação institucional em escolas ou redes de ensino

EIXO	ASPECTOS OBSERVADOS
Trabalho coletivo	<ul style="list-style-type: none"> • Superação da falta de trabalho coletivo • Identificação coletiva das prioridades
Negociação	<ul style="list-style-type: none"> • Negociação na escola • Negociação da escola junto ao órgão central • Diálogo da escola com o órgão central
Gestão democrática	<ul style="list-style-type: none"> • Maior participação na escola • Fortalecimento de colegiados como o Conselho Escolar e o grêmio estudantil • Fortalecimento dos atores da escola • Compartilhamento de poder • Fortalecimento da gestão democrática
Concepções e Usos da avaliação	<ul style="list-style-type: none"> • Alteração da concepção de avaliação • Sentido formativo da avaliação • Melhora na compreensão dos resultados das avaliações externas • Diagnóstico dos problemas • Elaboração ou revisão do PPP de forma participativa • Propiciação ou melhora da formação continuada • Tomada de decisão dos gestores
Relações na escola	<ul style="list-style-type: none"> • Melhora na relação professor-aluno • Melhora no clima institucional

Fonte: Elaboração dos autores com base nos artigos sobre implementação da avaliação institucional na educação básica, referenciados no corpo do texto.

É interessante notar que a avaliação institucional, a julgar pelos dados das pesquisas, ajuda a desenvolver aqueles aspectos determinantes para o seu sucesso: trabalho coletivo, participação da comunidade escolar, aprendizagem organizacional por meio da formação continuada e mudanças na concepção de avaliação, uma reculturação. Isso demonstra, por um lado, que a avaliação institucional, para ser realizada de modo participativo, exige outras mudanças na instituição e na rede de ensino. Mudanças que, de certo modo, condicionam a implementação. Por outro lado, esse tipo de avaliação contribui para que essas mudanças ocorram. Disso é possível concluir que, embora a implementação (e institucionalização) da avaliação institucional necessite de outras mudanças educacionais, para a implementação destas últimas, a avaliação institucional contribui fortemente. Nesse caso, há que se atentar para o alerta de Fullan (2009): a ação é mais relevante para as mudanças que planejamentos bem elaborados.

RECOMENDAÇÕES DOS ESTUDOS

Nem todos os autores apresentam de modo explícito recomendações com o propósito de que a avaliação institucional logre êxito nas escolas e/ou nas redes de ensino. Não obstante, foi possível organizar um conjunto bastante razoável de orientações que certamente são úteis na empreitada de experimentar processos avaliativos baseados na escola e que visem à sua qualidade global.

Dentre todas as recomendações presentes nos estudos, o destaque, como era de se esperar, está no investimento maciço na gestão democrática das escolas, para a qual, como tem-se insistido, a avaliação institucional, se realizada de modo participativo, contribui bastante.

O quadro a seguir resume as recomendações extraídas dos estudos sobre a implementação da avaliação institucional em escolas ou redes de ensino. Elas foram reunidas em seis eixos: participação e gestão democrática, formação, projeto pedagógico da escola, política de avaliação, estratégias e princípios.

QUADRO 5 - Recomendações para a implementação da avaliação institucional em escolas ou redes de ensino

EIXO	SUGESTÕES
Participação e Gestão democrática	<ul style="list-style-type: none"> • Aumentar a participação na escola (para isso, identificar o que a inibe) • Investir na gestão democrática (para isso, investir na própria avaliação participativa) • Garantir que a avaliação institucional esteja intimamente relacionada com a gestão democrática • Elaborar a proposta de avaliação de forma participativa • Incorporar os pais na tomada de decisões • Aproveitar os momentos em que os pais estão na escola para que eles a avaliem
Formação	<ul style="list-style-type: none"> • Formação efetiva da comunidade escolar para realizar a avaliação
Projeto pedagógico da escola	<ul style="list-style-type: none"> • Tornar o PPP uma referência coletiva da escola • Garantir que a avaliação institucional esteja relacionada com o PPP da escola
Política de avaliação	<ul style="list-style-type: none"> • Criação de condições objetivas pelo Estado • Apoio da Secretaria de Educação • Apoio político com vistas a fortalecer os atores locais • Atendimento às demandas vindas da escola • Municipalização dos processos de avaliação
Estratégias	<ul style="list-style-type: none"> • Diversificação de instrumentos de coleta de dados • Realização da avaliação durante o processo • Utilização das CPAs como estratégia no processo de avaliação institucional
Princípios	<ul style="list-style-type: none"> • Ter clareza do que se pretende com a avaliação institucional proposta • Articular os saberes diversos na definição da concepção de qualidade que a escola precisa possuir • Não abrir mão dos princípios que regem os modelos contra-hegemônicos da avaliação (a participação, por exemplo, é um princípio inegociável)

Fonte: Elaboração dos autores com base nos artigos sobre implementação da avaliação institucional na educação básica, referenciados no corpo do texto.

Como visto ao longo de todo o texto, há uma relação íntima entre avaliação institucional e gestão democrática da escola, de forma que esta foi um dos aspectos que mais favoreceram a implementação da avaliação institucional nas escolas, mas, igualmente, se não desenvolvida minimamente, um dos fatores que mais dificultam a empreitada. Curiosamente, contudo, seu fortalecimento parece ser uma das mudanças verificadas nas instituições que implementaram o processo de avaliação institucional. Por tudo isso, a promoção da gestão democrática é a recomendação mais frequente àquelas redes de ensino e unidades escolares que pretendem adotar ou implementar esse tipo de avaliação ou que já estão a implementá-la, sob pena de esse processo vir a perecer por faltar-lhe sustentação.

Com efeito, muitas das demais recomendações que compõem o quadro acima são, de certo modo, caminhos possíveis para que a gestão democrática vá se concretizando nas

escolas. Para que a avaliação ocorra de modo participativo, alguns exemplos de estratégias incluem incorporação das famílias e dos estudantes na tomada de decisões; elaboração coletiva do PPP, a fim de que este se torne uma referência coletiva da escola; e incorporação de diferentes sujeitos e saberes na definição da qualidade da escola. Além disso, os princípios que sustentam uma avaliação participativa, como participação plural, diálogo, negociação e defesa da escola pública e estatal, são princípios compartilhados também quando se defende a gestão democrática da escola pública.

Mas nem todas as estratégias estão relacionadas diretamente com a gestão democrática. Há outros desafios, próprios do processo de avaliação institucional, que precisam ser enfrentados, sobretudo aqueles de ordem mais técnica. Por isso, os autores recomendam algumas ações no âmbito das escolas que podem qualificar o processo avaliativo, como a formação da comunidade escolar para avaliar a escola, para a qual parceiros externos podem ser importantes; articulação da avaliação institucional com o projeto pedagógico da escola; definição clara dos objetivos da avaliação; diversificação dos instrumentos de pesquisa e avaliação, conforme os objetivos pretendidos; constituição de uma equipe de avaliação; etc. Há também questões políticas intervenientes do processo avaliativo, de sorte que alguns estudos recomendam à gestão central o apoio à realização da avaliação, por meio de ações concretas, com o intuito de garantir suas condições materiais e subjetivas mínimas.

CONSIDERAÇÕES FINAIS

Quando analisados os aspectos que afetaram a implementação da avaliação institucional nas escolas ou redes de ensino, favorecendo-a ou inibindo-a, chamou a atenção o tipo de gestão da escola e da política de avaliação e a cultura de avaliação das unidades escolares e das redes de ensino. A gestão democrática é o fator isolado que mais impacto teve no sucesso da implementação. Sua fragilidade, por outro lado, torna muito difícil a realização de uma avaliação participativa. O apoio da alta gestão, a descontinuidade da política e as concepções de

avaliação dominantes na escola são outros elementos que parecem afetar fortemente a mudança em questão.

Grande parte dos aspectos que, segundo as pesquisas, influenciam a implementação da avaliação institucional é da mesma natureza daqueles elencados por Fullan (2009) em seu estudo sobre mudança educacional. A maioria dos aspectos apontados pelas pesquisas poderia perfeitamente ser classificada como se tratando de características da proposta de mudança (proposta de avaliação) ou de papéis locais (atuação da gestão central, participação da comunidade escolar, tipo de gestão escolar, envolvimento da direção escolar, trabalho coletivo e relações de poder entre os membros da comunidade escolar).

Analisados todos os fatores até aqui discutidos, é fácil concluir que uma das razões para que a avaliação institucional seja tão complexa é que seu sucesso está relacionado com a implementação de outras mudanças educacionais, igualmente desafiadoras, como a gestão democrática, o trabalho coletivo, a cultura de avaliação voltada para a melhoria escolar, a horizontalização das relações de poder, dentre outras. Com isso, poder-se-ia inferir que, enquanto essas últimas mudanças não estivessem disponíveis, a avaliação institucional fatalmente fracassaria. Contudo, e isso parece ser a conclusão mais importante até aqui, os estudos também revelaram que a avaliação institucional, quando ocorre de modo participativo, provoca mudanças nos indivíduos e nas instituições. Exatamente aquelas mudanças que mais favorecem a implementação desse tipo de avaliação.

Como se vê, há no processo de avaliação institucional uma circularidade. Isso significa que reconhecer aqueles fatores que influenciam a implementação da avaliação institucional é fundamental, no sentido de buscar garantir as condições sem as quais a implementação poderá fracassar. Por outro lado, partindo do princípio de que a avaliação institucional, realizada de modo coletivo e tendo por objetivo a melhoria da escola e a construção de comunidades profissionais de aprendizagem, contribui para que algumas dessas condições emergam, não parece razoável aguardar que a mudança só seja experimentada quando todas as condições favoráveis estiverem disponíveis.

REFERÊNCIAS

- AFONSO, Almerindo Janela. Políticas educativas e autoavaliação da escola pública portuguesa: apontamentos de uma experiência. *Estudos em Avaliação Educacional*, São Paulo, v. 21, n. 46, p. 1-19, 2010. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/ae/article/view/2016/1976>>. Acesso em: 22 fev. 2017.
- BETINI, Geraldo. *Avaliação institucional em escolas públicas de ensino fundamental de Campinas*. 2009. Tese (Doutorado em Educação) – Universidade Estadual de Campinas, Campinas, SP, 2009.
- BETINI, Geraldo. Avaliação Institucional Participativa em escolas públicas de ensino fundamental. *Educação: Teoria e Prática*, Rio Claro, v. 20, n. 35, p. 117-132, jul./dez. 2010. Disponível em: <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/4089/3296>>. Acesso em: 22 fev. 2017.
- BONDIOLI, Anna (Org.). *O projeto pedagógico da creche e sua avaliação: a qualidade negociada*. 2. ed. Campinas: Autores Associados, 2013.
- BRANDALISE, Mary Ângela Teixeira. Avaliação institucional na escola pública: os (des)caminhos de uma política educacional. *Educar em Revista*, Curitiba, n. 1, p. 55-74, 2015. Edição especial. Disponível em: <<http://www.scielo.br/pdf/er/nspe1/1984-0411-er-1-spe-00055.pdf>>. Acesso em: 22 fev. 2017.
- BRANDALISE, Mary Ângela Teixeira; MARTINS, Clícia Bührrer. Programa de avaliação institucional da educação básica do Paraná: da produção à implementação da política na escola. *Estudos em Avaliação Educacional*, São Paulo, v. 22, n. 50, p. 1-21, set./dez. 2011. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/ae/article/view/1964/1940>>. Acesso em: 22 fev. 2017.
- CHAPMAN, Christopher; SAMMONS, Pamela. *School self-evaluation for school improvement: what works and why?* Slough: CfBT, 2013. Disponível em: <www.educationdevelopmenttrust.com>. Acesso em: 12 abr. 2017.
- COSTA, Maria Simone. *Avaliação institucional no ensino fundamental: a participação dos estudantes*. 2012. Tese (Doutorado em Educação) – Universidade Estadual de Campinas, Campinas, SP, 2012.
- COUSINS, Bradley J.; EARL, Lorna M. The Case for Participatory Evaluation. *Educational Evaluation and Policy Analysis*, Kalamazoo, Michigan, v. 14, n. 4, p. 397-418, 1992.
- DALBEN, Adilson. *Avaliação institucional participativa na educação básica: possibilidades, limitações e potencialidades*. 2008. Dissertação (Mestrado em Educação) – Universidade Estadual de Campinas, Campinas, SP, 2008.
- DALBEN, Adilson. Avaliação institucional na escola de educação básica: uma aproximação orientada pelos princípios da participação. *Educação: Teoria e Prática*, Rio Claro, v. 20, n. 35, p. 133-146, jul./dez. 2010. Disponível em: <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/4090/3297>>. Acesso em: 22 fev. 2017.

DALBEN, Adilson; ALMEIDA, Luana Costa. Para uma avaliação de larga escala multidimensional. *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 61, p. 12-28, jan./abr. 2015. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/ea/article/view/3140/2954>>. Acesso em: 22 fev. 2017.

FREITAS, Luiz Carlos. Qualidade negociada: avaliação e contra-regulação na escola pública. *Educação & Sociedade*, Campinas, v. 26, n. 92, p. 911-933, out. 2005.

FREITAS, Luiz Carlos. Eliminação adiada: o ocaso das classes populares no interior da escola e a ocultação da (má) qualidade do ensino. *Educação & Sociedade*, Campinas, v. 28, n. 100, p. 965-987, out. 2007. Disponível em: <<http://www.scielo.br/pdf/es/v28n100/a1628100.pdf>>. Acesso em: 22 fev. 2017.

FULLAN, Michael. *O significado da mudança educacional*. 4. ed. Porto Alegre: Artmed, 2009.

LIMA, Licínio. A avaliação institucional como instrumento de racionalização e o retorno à escola como organização formal. *Educação e Pesquisa*, São Paulo, v. 41, n. especial, p. 1339-1352, dez. 2015. Disponível em: <<http://www.scielo.br/pdf/ep/v41nspe/1517-9702-ep-41-spe-1339.pdf>>. Acesso em: 22 fev. 2017.

MACEDO, Mara. *Auto-avaliação institucional na educação básica: uma contribuição necessária para o aprimoramento das práticas pedagógicas*. 2009. Dissertação (Mestrado em Educação) – Universidade do Oeste Paulista, Presidente Prudente, SP, 2009.

MENDES, Geisa Vaz; SORDI, Mara Regina Lemes. Implementação de política de avaliação institucional no ensino fundamental: a percepção do gestor municipal. *Revista de Educação PUC-Campinas*, Campinas, v. 17, n. 2, p. 241-250, jul./dez. 2012. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/reveducacao/article/view/1001/978>>. Acesso em: 22 fev. 2017.

MENDES, Geisa Vaz et al. Autoavaliação como estratégia de resistência à avaliação externa ranqueadora. *Educação e Pesquisa*, São Paulo, v. 41, n. especial, p. 1283-1298, dez. 2015. Disponível em: <<http://www.scielo.br/pdf/ep/v41nspe/1517-9702-ep-41-spe-1283.pdf>>. Acesso em: 22 fev. 2017.

MIRANDA, Antonio Carlos; RODRIGUES, Sueli Carrijo. O uso da DEA como ferramenta alternativa da gestão escolar na avaliação institucional. *Educação: Teoria e Prática*, Rio Claro, v. 20, n. 35, p. 163-180, jul./dez. 2010. Disponível em: <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/3464/3302>>. Acesso em: 22 fev. 2017.

NEVO, David. *Evaluación basada en el centro: un diálogo para la mejora educativa*. Bilbao: Ediciones Mensajero, 1997.

PACHECO, José Augusto. Políticas de avaliação e qualidade da educação. Uma análise crítica no contexto da avaliação externa de escolas, em Portugal. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas; Sorocaba, v. 19, n. 2, p. 363-371, jul. 2014. Disponível em: <<http://www.scielo.br/pdf/aval/v19n2/a05v19n2.pdf>>. Acesso em: 22 fev. 2017.

RIBEIRO, Vanda Mendes; GUSMÃO, Joana Borges Buarque. Uma leitura dos usos dos indicadores da qualidade na educação. *Cadernos de Pesquisa*, São Paulo, v. 40, n. 141, p. 823-847, set./dez. 2010. Disponível em: <<http://www.scielo.br/pdf/cp/v40n141/v40n141a08.pdf>>. Acesso em: 22 fev. 2017.

SÁ, Virgínio. A (auto)avaliação das escolas: “virtudes” e “efeitos colaterais”. *Ensaio: Avaliação de Políticas Públicas em Educação*, Rio de Janeiro, v. 17, n. 62, p. 87-108, jan./mar. 2009. Disponível em: <<http://www.scielo.br/pdf/ensaio/v17n62/a05v1762.pdf>>. Acesso em: 22 fev. 2017.

SILVA, Itamar Mendes. Autoavaliação e gestão democrática na instituição escolar. *Ensaio: Avaliação de Políticas Públicas em Educação*, Rio de Janeiro, v. 18, n. 66, p. 49-64, jan./mar. 2010. Disponível em: <<http://www.scielo.br/pdf/ensaio/v18n66/a04v1866.pdf>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes. Avaliação institucional participativa: contradições emergentes a partir do exame da categoria espaço/tempo. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, v. 11, n. 4, p. 53-61, dez. 2006. Disponível em: <<http://periodicos.uniso.br/ojs/index.php/avaliacao/article/view/948>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes. Há espaços para a negociação em políticas de regulação da qualidade da escola pública? *Revista Educação: Teoria e Prática*, Rio Claro, v. 20, n. 35, p. 147-162, jul./dez. 2010. Disponível em: <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/3387/3301>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes. Comissão Própria de Avaliação (CPA): similaridades e dessemelhanças no uso da estratégia na educação superior e em escolas do ensino fundamental. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, v. 16, n. 3, p. 603-617, nov. 2011. Disponível em: <<http://www.scielo.br/pdf/aval/v16n3/v16n1a06.pdf>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes. Implicações ético-epistemológicas da negociação nos processos de avaliação institucional participativa. *Educação & Sociedade*, Campinas, v. 33, n. 119, p. 485-510, abr./jun. 2012. Disponível em: <<http://www.scielo.br/pdf/es/v33n119/a09v33n119.pdf>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes; LÜDKE, Menga. Os profissionais da educação e as suas relações com a comunidade escolar: aprendizagens no processo de qualificação da escola mediado pela avaliação institucional. *EccoS: Revista Científica*, São Paulo, v. 11, n. 1, p. 209-227, jan./jun. 2009a. Disponível em: <<http://www4.uninove.br/ojs/index.php/eccos/article/viewFile/1514/1273>>. Acesso em: 22 fev. 2017.

SORDI, Mara Regina Lemes; LÜDKE, Menga. Da avaliação da aprendizagem à avaliação institucional: aprendizagens necessárias. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas; Sorocaba, v. 14, n. 2, p. 313-336, jul. 2009b. Disponível em: <<http://periodicos.uniso.br/ojs/index.php/avaliacao/article/view/310/311>>. Acesso em: 22 fev. 2017.

SOUZA, Ângelo Ricardo; DITTRICH, Douglas Danilo. Avaliação na escola, avaliação da escola: análise de uma experiência escolar. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 52, p. 26-46, maio/ago. 2012. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/1927/3185>>. Acesso em: 22 fev. 2017.

LUIZ SÍVERES

Coordenador e professor do Programa de Pós-Graduação
em Educação da Universidade Católica de Brasília (UCB),
Brasília, Distrito Federal, Brasil
luiz.siveres@catolica.edu.br

JOSÉ ROBERTO DE SOUZA SANTOS

Doutor em Educação pela Universidade Católica de Brasília (UCB).
Pesquisador do Instituto Nacional de Estudos e Pesquisas
Educaionais Anísio Teixeira (Inep). Professor da Rede
Pública de Ensino do Distrito Federal, Brasília, Distrito
Federal, Brasil
joser.santos@inep.gov.br

Recebido em: 31 OUTUBRO 2017

Aprovado para publicação em: 30 JANEIRO 2018

PRÁTICAS LETIVAS DE SALA DE AULA DE MATEMÁTICA NOS ANOS INICIAIS

ISABEL CRISTINA RODRIGUES DE LUCENA

JOSETE LEAL DIAS

ANTÓNIO MANUEL ÁGUAS BORRALHO

RESUMO

Este artigo visa caracterizar as práticas letivas que ocorrem nas salas de aula dos anos iniciais, na área curricular de matemática, a partir de uma pesquisa em cooperação internacional. Por meio de uma metodologia de investigação qualitativa, em que a recolha de dados foi realizada a partir de observações de aulas e entrevistas assentes numa matriz de investigação, foram elaboradas sínteses conclusivas e interpretativas as quais denominamos de narrativas. A análise dessas narrativas mostrou alguns resultados: práticas avaliativas ainda centradas na classificação; participação efetiva dos alunos no processo de ensino e aprendizagem muito reduzida; ausência de integração no que diz respeito à relação entre os processos de ensino, avaliação e aprendizagem; dentre outras. Pelo estudo ainda é possível afirmar que há um desejo de se realizar uma avaliação para as aprendizagens, porém é necessário fortalecer o campo teórico sobre avaliação para que em aula se possam problematizar e transgredir práticas estabelecidas.

PALAVRAS-CHAVE PROCESSO DE ENSINO-APRENDIZAGEM • AVALIAÇÃO DAS APRENDIZAGENS • MATEMÁTICA • ANOS INICIAIS DO ENSINO FUNDAMENTAL.

PRÁCTICAS LECTIVAS DE MATEMÁTICAS EN AULAS DE PRIMÁRIA

RESUMEN

Este artículo tiene como objetivo estudiar prácticas lectivas en aulas de los primeros años de enseñanza, en el área de las matemáticas, en base de una investigación de cooperación internacional. A partir de una investigación cualitativa, donde la recogida de datos se fundamentó en observaciones en el aula y entrevistas basadas en una matriz de la investigación, se realizaron síntesis conclusivas y interpretativas a las que denominamos de narrativas. El análisis de estas narrativas produjo algunos resultados: las prácticas evaluativas todavía se centran en la clasificación; la participación real de los alumnos en el proceso de enseñanza y aprendizaje es muy reducida; la ausencia de integración en lo que se refiere a la relación entre los procesos de enseñanza, evaluación y aprendizaje, entre otros. Por el estudio sigue siendo posible afirmar que existe un interés en realizar una evaluación para las aprendizajes, no obstante, es necesario fortalecer el campo teórico de la evaluación de modo que, en aula, se puedan problematizar e transgredir las prácticas establecidas.

PALABRAS CLAVE PROCESO DE ENSEÑANZA-APRENDIZAJE • EVALUACIÓN DEL APRENDIZAJE • MATEMÁTICAS • AÑOS INICIALES DE LA EDUCACIÓN BÁSICA.

MATHEMATICS CLASSROOM PRACTICES IN THE ELEMENTARY SCHOOL

ABSTRACT

This article aims to characterize teaching practices that occur in the classrooms of the elementary school, in the area of mathematics, based on an international cooperative research. Conclusive and interpretative summaries were developed using a qualitative methodology in which data were gathered from classroom observations and interviews, grounded in an investigative matrix, which we called narratives. The analyses of these narratives showed the following results: assessment practices are still centered on classification; effective participation of students in the teaching and learning process is greatly reduced; lack of integration regarding the relationship among the processes of teaching, assessment and learning; among others. This study also shows that there is a desire to conduct an assessment for learning. However, it is necessary to strengthen the theoretical field on assessment so that, in the classroom, we can problematize and counter established practices.

KEYWORDS TEACHING AND LEARNING PROCESS • LEARNING ASSESSMENT • MATHEMATICS • EARLY YEARS OF ELEMENTARY SCHOOL.

INTRODUÇÃO

Avaliação das aprendizagens ou *avaliação para as aprendizagens*? De modo geral podemos dizer que a primeira tem sido tema recorrente no cotidiano escolar, especialmente quando se trata de diagnósticos por meio de provas em larga escala, como tem sido nos últimos anos via políticas estabelecidas em âmbitos internacional e nacional (no Brasil também em níveis estadual e municipal). Não muito distante, essa indicação ainda abarca a avaliação interna, ou seja, a de sala de aula. Quanto à segunda, mediante referenciais nacionais e internacionais, podemos considerá-la um desafio a ser alcançado, pois um ambiente formativo para uma cultura de sucesso em aulas, em especial de matemática, exige mudança de paradigma no que se refere à avaliação.

Ao mencionar o termo “avaliação para as aprendizagens”, o que estamos a significar? Que relações são possíveis entre aprendizagem-ensino-avaliação? Que práticas ocorrem no interior das salas de aulas para que haja a aprendizagem? Estes questionamentos compuseram o pano de fundo do Projeto Aera – Avaliação e Ensino na Educação Básica em

Portugal e no Brasil: Relações com as Aprendizagens (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Fundação para a Ciência e a Tecnologia – Capes/FCT – Universidade Federal do Pará/Universidade de Évora – UFPA/Uévoa – 2014 a 2016), que assume a avaliação como uma ação deliberadamente organizada para ajudar os alunos a aprender, pois visa acompanhar todo o processo de construção do conhecimento de forma temporal e circunstanciada. Com esta intencionalidade, o projeto foi realizado em cinco fases distintas e interdependentes: fase teórica e conceitual; análise documental (dispositivos legais sobre avaliação de aprendizagens); estudo extensivo (questionários para professores); estudo intensivo (*observação de aulas e entrevistas*); e interação social e reflexão (*entrevistas com professores e grupos de estudantes*).

As duas últimas fases ocorreram em salas de aulas localizadas no município de Belém – Pará, Brasil – e na cidade de Évora – Portugal. O estudo tinha como foco de pesquisa turmas de 4º e 5º anos do ensino fundamental, correspondentes aos 3º e 4º anos do primeiro ciclo de escolaridade de Portugal. Um dos objetivos foi analisar práticas de ensino e de avaliação dos professores e a participação dos alunos nos processos conducentes às suas aprendizagens.

Este artigo proporciona conhecimento para problematizar as práticas avaliativas em matemática dos anos iniciais em escolas públicas, como forma de trazer ao debate aspectos pontuais de vivências didáticas em que a avaliação é tratada como objeto. O desejo é dialogar com formadores e professores no sentido de subsidiar reflexões sobre a prática com fins de contribuir para o amadurecimento e a transgressão de práticas letivas em matemática que ainda explicitam indícios de uma avaliação classificatória, embora seja inequívoco, por parte da comunidade educativa em geral, o reconhecimento da avaliação formativa com maiores contributos para as aprendizagens.

Para este fim, apresentamos um quadro teórico-metodológico de referência com foco no método utilizado para a compreensão e análise dos dados construídos na investigação e alguns resultados com vista a responder aos propósitos deste artigo.

QUADRO TEÓRICO-METODOLÓGICO DE REFERÊNCIA

A avaliação ocupa todas as formas de atividades humana e, como termo polissêmico, tem assumido diferenciados paradigmas. Em meio a essas assertivas compreendemos a avaliação como oportunidade de aprendizagem que

[...] toma o seu caráter subjetivo como auxiliar na tomada de decisões, quer seja dos estudantes quer seja do professor, envolvendo por vezes alguma modificação no modo de estudar ou de organizar as ações, em uma perspectiva interpretativa e qualitativa do trabalho que realizam. (PIRES; BURIASCO, 2012, p. 15)

Essa tomada de decisão ocorre a partir do pressuposto da concepção de avaliação que o sujeito assume.

Com isso, é possível dizer que diferentes significados e concepções de avaliação podem gerar variadas práticas avaliativas. Portanto, compreender o que se quer dizer com o termo *avaliação de aprendizagens* passa por percepções existentes sobre ele. Avaliação como medida, avaliação como diagnóstico, avaliação como regulação são algumas terminologias que carregam significados sobre qual propósito e com que práticas é possível avaliar aprendizagens. No entanto, é admissível identificarmos pelo menos uma diferença de significados com alta relevância para as práticas cotidianas do trabalho docente: avaliar as aprendizagens não é o mesmo que avaliar *para* as aprendizagens.

Avaliar para que um estudante aprenda mais e melhor é eminentemente um trabalho interno das salas de aula, sob a mediação do professor e colaboração intensa entre os demais estudantes. Nesse tipo de avaliação, o ensino e a aprendizagem são constituídos conjuntamente. Não há um momento especial para que a avaliação ocorra, pois ela pertence ao processo, isto é, ao instante da aula (LEAHY et al., 2005). Nessa perspectiva, a avaliação passa a ser formativa porque tem o compromisso de melhorar as aprendizagens dos envolvidos – professores e alunos –, com vista a atender às necessidades dos alunos por meio da regulação do ensino. Assim, o aspecto classificatório e seletivo fica em segundo plano e apenas por obrigatoriedade de atendimento ao sistema

educacional poderá se traduzir em uma classificação. Em se tratando de uma avaliação para as aprendizagens – formativa –, os aspectos com fins formativos que mais interessam são: a natureza e diversificação das tarefas de ensino-aprendizagem; a qualidade e quantidade do *feedback* prestado aos estudantes; a clareza e partilha de critérios entre os alunos; bem como o uso constante de auto e heteroavaliação.

Segundo Ponte (2014), as tarefas evocam diferentes papéis para as aprendizagens, podendo se caracterizar como: exercícios e problemas, de natureza mais fechada, centrando os raciocínios para desenvolvimento de algoritmos e resultados únicos; explorações e investigações, de natureza mais aberta, provocando o desenvolvimento de raciocínios em situações complexas que contribuem para a formação de efetivas experiências com a criatividade matemática; e tarefas de natureza mais aberta ou mais fechada com variação para o sentimento de desafio que possam provocar nos estudantes e o tempo de realização/duração até sua conclusão.

É relevante destacar que a sequenciação das tarefas a serem propostas em aulas faz parte do planejamento docente, levando em conta a adequação curricular. Então,

[...] em função do papel e das ações do professor e dos alunos no trabalho, a partir das tarefas, são proporcionados ambientes de sala de aula, mais ou menos propícios às aprendizagens, à avaliação e ao ensino. (BORRALHO; LUCENA; BRITO, 2015, p. 26)

Stein e Smith (1998) classificam as tarefas matemáticas em relação ao nível de exigência cognitiva em: tarefas de nível cognitivo reduzido, caracterizando-se entre memorização e procedimentos sem conexões; e tarefas de nível cognitivo elevado, distinguindo-se entre procedimentos com conexões e fazendo matemática. As mesmas autoras ainda diferenciam as tarefas em três formas: como aparecem nos materiais curriculares; como são apresentadas pelo professor; e como são realizadas pelos alunos. Na concepção das autoras, por vezes, a natureza da tarefa muda quando passa de uma forma para outra, isto é, a tarefa que o professor propõe aos alunos, geralmente escrita e complementada

oralmente, nem sempre é idêntica àquelas presentes nos materiais curriculares (por exemplo, no livro didático). Do mesmo modo, por terem informações adicionais e diferentes interpretações, as tarefas que os alunos realizam muitas vezes não são as mesmas que o professor apresentou inicialmente. Todos estes aspectos são importantes, especialmente na implementação em sala de aula, onde efetivamente ocorrem as aprendizagens dos alunos.

Ainda em relação à avaliação para as aprendizagens, outro aspecto relevante é a distribuição de *feedback* aos estudantes. Como afirma Fernandes (2016, p. 4):

[*feedback*] O que é isso? É sempre que possível dar indicações claras a cada um dos alunos acerca do esforço que ele deve fazer para aprender, acerca do ponto em que se encontra em relação às aprendizagens que tem que desenvolver.

É possível identificar formas e tipos de *feedback*, que podem ser apresentados na forma oral ou escrita, com caráter corretivo ou descritivo (TUNSTALL; GIPPS, 1996), esperando-se que sejam sempre positivos.

É preciso considerar que o *feedback* faz parte da comunicação entre alunos e professores sendo essencial para o desenvolvimento do processo ensino-aprendizagem. Por meio daquilo que ouvem ou leem, os alunos podem identificar seus avanços e quais trajetórias devem percorrer para superar suas dificuldades. O professor, por sua vez, também poderá perceber as modificações que necessita fazer para proporcionar um melhor ensino, que atenda às necessidades de seus alunos. Portanto, o *feedback*, não pode ser configurado como uma simples informação, e sim uma comunicação pensada, estruturada de acordo com a situação de cada aluno, estabelecida em tempo hábil, a fim de proporcionar oportunidades de esforço do aluno para com o êxito na sua aprendizagem, a superação de dificuldades, percepção de possíveis erros. (SILVA et al., 2016)

Como visto, o *feedback* é forte aliado da avaliação formativa por ter a função de encorajar o estudante em seu

aprendizado, uma vez que a comunicação entre os sujeitos visa orientar os processos cognitivos e metacognitivos necessários para o domínio do saber culturalmente acumulado. Além desse aspecto, é possível destacar a importância do *feedback* para pontos como autoavaliação e promoção do senso de responsabilização do aluno com o desenvolvimento de suas tarefas com vistas à efetivação da aprendizagem.

Tarefas e *feedback* são fundamentais e, por isso, a avaliação formativa exige outro ponto relevante – a transparência. É importante para os envolvidos saber que

[...] os objetivos, as aprendizagens a desenvolver e todos os processos de avaliação devam ser claramente expressos e devem estar sempre disponíveis para quem a eles quiser ter acesso. (FERNANDES, 2008, p. 19)

Portanto, é necessário que os critérios de avaliação sejam apresentados e negociados de forma clara entre os envolvidos, e que haja uma participação ativa dos alunos no processo de avaliação, dando subsídios, tanto ao professor quanto aos próprios estudantes, para a análise dos trabalhos realizados pelos alunos. Mediante os critérios, claros e partilhados por todos, os “alunos devem ser capazes de perceber sem problemas a sua situação face às aprendizagens que têm que adquirir ou desenvolver” (FERNANDES, 2008, p. 19), ou seja, serem protagonistas em seus processos de aprendizagem.

Se estivermos tratando de uma avaliação para a formação de sujeitos, em especial, participativos e protagonistas de suas aprendizagens, a aula necessariamente precisa oportunizar ambiência exequível para realizações de auto e heteroavaliações. Assim como a avaliação possui seus indicadores – natureza e modalidades, funções, objeto, tempo-espço pedagógicos, a quem se destina, entre outros –, a autoavaliação e, por conseguinte, a heteroavaliação também precisam trazer seus indicadores de forma clara e o mais objetivamente possível para os alunos (MIDDLE STATES COMMISSION ON HIGHER EDUCATION, 2007). Tornam-se necessárias ações que visem ao desenvolvimento da autocrítica, da autonomia, da responsabilidade e da autorregulação.

Toda autoavaliação é um ato intencional e, por isso, a transparência torna-se fundamental, pois aprendizagem e autoavaliação regulada se desenvolvem e influenciam-se mutuamente. A este respeito, o estudo de Vieira (2013) indica que a apropriação dos critérios de avaliação é essencial para o desenvolvimento da capacidade de autorregulação/autoavaliação, pois

Com uma participação ativa na discussão e negociação dos termos e efeitos da avaliação, o aluno que se autoavalia revela um grande envolvimento pessoal no processo de avaliação, gerindo o seu desempenho e refletindo sobre ele, tornando-se assim regulador da sua própria aprendizagem. (VIEIRA, 2013, p. 28)

Os critérios e os registros de informações/resultados podem ser instrumentalizados para que o acompanhamento e a tomada de decisão sejam mais bem observados pelos alunos durante os processos de auto e heteroavaliação.

Por fim, tomamos como uma importante referência o estudo realizado por Black e Wiliam (2006), que, em síntese, destacou os seguintes pontos: as práticas sistemáticas de avaliação formativa melhoram significativamente as aprendizagens de todos os alunos; os alunos que mais se beneficiam de tais práticas são os que revelam mais dificuldades de aprendizagem; e os alunos que frequentam aulas em que a avaliação predominante é de natureza formativa obtêm melhores resultados em exames e provas de avaliação externa do que os que frequentam aulas em que a avaliação é essencialmente somativa. Portanto, há de se investigar quais e como ocorrem as práticas de avaliação no âmbito das salas de aulas diante dos referenciais da avaliação de natureza formativa, ou seja, para as aprendizagens dos estudantes.

Tal problemática foi investigada pelo projeto Aera, cujo objetivo central de investigação deriva da necessidade de se compreenderem as relações entre uma variedade de elementos que afetam o desenvolvimento da aprendizagem do aluno. Entre outros, destacamos: ensino e práticas de avaliação de professores; percepções dos professores sobre ensino, avaliação e aprendizagem; natureza das tarefas de avaliação

utilizadas nas salas de aula; frequência, distribuição e natureza do *feedback* utilizado; e participação dos alunos nos processos de ensino, avaliação e aprendizagem. Os focos de investigação foram norteados por duas matrizes originárias do projeto e que se relacionam: a matriz investigação (Quadro 1); e a matriz trianguladora de análise (Quadro 2). Por meio das matrizes, a fase do estudo intensivo (observação de aulas e entrevistas) permitiu descrever detalhadamente as ações e interações que corporizam as atividades de ensino, aprendizagem e avaliação, constituindo uma oportunidade única para a compreensão de uma variedade de relações entre os elementos já referidos, tomando a sala de aula, e não os alunos ou os professores, como unidade de análise.

No Quadro 1, observam-se os objetos (unidades de análise) que foram alvo da investigação, orientando toda a pesquisa no âmbito do Aera, em particular as observações e entrevistas.

QUADRO 1 - Matriz de investigação

OBJETOS	DIMENSÕES
Práticas de ensino	Organização e desenvolvimento do ensino
	Recursos, materiais e tarefas utilizados
	Dinâmicas de sala de aula (e.g., trabalho de grupo; trabalho em pares; trabalho individual; organização das discussões)
	Papel do professor
	Papel dos alunos
	Estrutura da aula
Práticas de avaliação	Integração/articulação entre os processos de ensino/avaliação/aprendizagem
	Utilizações da avaliação (e.g., para classificar, para orientar, para regular, para melhorar)
	Instrumentos de avaliação predominantes (e.g., testes, trabalhos escritos, questões orais, listas de verificação, tarefas de sala de aula)
	Natureza, frequência e distribuição de <i>feedback</i>
	Dinâmicas de avaliação (e.g., autoavaliação, heteroavaliação)
	Natureza da avaliação formativa (formal e informal)
	Natureza da avaliação somativa (formal e informal)
	Papel do professor
	Papel dos alunos
Participação dos alunos	Participação dos alunos (dinâmicas, frequência e natureza)
	Relação pedagógica com os professores

Fonte: Projeto Aera, 2015.

A recolha de dados, feita nas fases de estudo intensivo e de interação social, foi basilar à composição de narrativas e suas respectivas análises, considerando os seguintes objetos de análise: *práticas de ensino, práticas de avaliação e participação dos alunos*. Por fim, foram organizadas sínteses interpretativas e conclusivas. O Quadro 2 apresenta de forma diagramada essa organização:

QUADRO 2 - Matriz trianguladora de análise



Fonte: Projeto Aera, 2015.

Por meio desses parâmetros, o projeto Aera tem realizado ações para descrever, analisar e interpretar práticas de ensino e de avaliação de docentes da educação básica. A seguir destacamos situações que nos indicam as relações entre ensino, aprendizagem e avaliação como forma de explicitar práticas avaliativas como contribuição à reflexão sobre a temática.

ALGUNS RESULTADOS

Os resultados foram organizados tomando os objetos ensino, avaliação e participação dos alunos de forma articulada

e referem-se à análise e interpretação dos dados provenientes das fases de estudos intensivos e interação social e reflexão. A partir da matriz de investigação (Quadro 1), é possível perceber os objetos e dimensões de análise que nos possibilitaram levantar um rico e variado conjunto de informações, de modo a tomar como significativas as relações que emergem entre os seus componentes. Apresentamos os resultados em três vertentes – práticas de ensino, práticas avaliativas e participação dos alunos –, pois é na materialidade do ensino que de fato se articulam os objetos da didática. É a partir dessas vertentes que traremos o movimento ocorrido nas aulas no tocante à articulação entre ensino, aprendizagem e avaliação.

Partindo dessa materialidade, podemos dizer que, em geral, nos contextos brasileiro e português, as práticas de ensino, geralmente, estiveram centradas no professor. A socialização do conhecimento ocorreu repetidas vezes por meio da explanação oral do objeto de ensino, seguida da aplicação de conhecimento por parte dos alunos, em tarefas do tipo exercícios em folhas de papel, no caderno e, algumas vezes, no livro didático ou no quadro de escrever.

Em meio a essa tradição áulica, é possível indicar vivências pedagógicas como contexto social, uma vez que as tarefas eram realizadas a partir do vivido, em especial, no espaço escolar ribeirinho. Mesmo com essa ressalva, ainda se pode concluir que o tipo de tarefa mais utilizada foi a que se centra na feitura de exercícios, muitas vezes indicados pelos professores como de perspectiva formativa. Apresentadas de forma rotineira e com pouca variação para com os sentimentos de desafios a provocar nos alunos, as tarefas eram resolvidas individualmente pelos alunos. Vale ressaltar que o *feedback* esteve voltado ao atendimento individual, porém em tempo hábil.

Este tipo de constructo de aula também é observado pelos alunos, pois, segundo eles:

A professora sempre segue o mesmo jeito [se referem ao roteiro da aula] que acontece na sequência seguinte: copia no quadro, faz a explicação, solicita a resolução e verifica de forma individual em sua mesa. (EABM42)¹

¹ EABMxy: Entrevista Aluno x, país Brasil, aula de Matemática, referente ao Professor y.

Ela começa assim, ela escreve no quadro e aí ela responde umas [questões], depois quando chega ao final, ela passa umas mais difíceis para ver se a gente aprendeu. (EABM71)

A professora sempre dá aula usando o quadro e o pincel. (EABM42)

Poucas práticas assistidas tiveram presente o desenvolvimento de aulas que envolviam tarefas com materiais manipuláveis ou com dinâmicas em que os alunos precisavam formar grupos e discutir sobre possíveis resoluções. Principalmente no contexto brasileiro, em metade destas salas de aulas, o espaço físico e a mobília foram obstáculos para o alcance dos objetivos postos, pois o professor apresentava dificuldades em deslocar-se pela sala, os alunos tinham dificuldades em ouvir as orientações do professor e os debates entre eles, o material manipulável (das poucas situações ocorridas) não cabia adequadamente nas carteiras, a organização da sala levava mais tempo que o necessário e a aula terminava sem a conclusão da tarefa.

Há uma percepção geral entre os professores de que a repetição de tarefas similares no estilo de exercício, bem como a centralidade do professor como o principal explanador de conteúdos e de exemplos, é a estratégia mais segura para a garantia de aprendizagens dos alunos, como podemos ver nas manifestações dos docentes: “Eu bato [repite] bastante no exemplo para ver se fica alguma coisa na memória desses meninos” (EPBM1);² “Utilizo muito pouco os materiais concretos [...] vou dando exemplos, [...] faço exercícios” (EPPM2).³

Mesmo de maneira diminuta, foi possível identificar práticas diferenciadas desta percepção geral, o que, conseqüentemente, concorre para um tipo de avaliação para as aprendizagens. Alguns relatos destacam isso:

Eu utilizo garrafas, materiais deles como remo, pipa. É um planejamento bem flexível aos objetos que os estudantes conhecem e tem disponível na escola. (EPBM6)

Então para começar qualquer conteúdo, primeiro os ouço [os alunos], depois vou pesquisar a história, como ele [o conteúdo] se

2 EPBMx: Entrevista Professor x, país Brasil, aula de matemática.

3 EPPMx: Entrevista Professor x, país Portugal, aula de matemática.

aproxima e se aplica à nossa realidade, realidade deles [dos alunos], da escola, da comunidade e como a partir de um conteúdo podemos conhecer outros, porque cada conteúdo puxa outros e têm uma ramificação, uma conexão com outros conteúdos e isso a gente faz através de vídeos, músicas, teatro. (EPBM6)

Gosto mais de colocar os alunos a trabalhar em grupo, porque acho que os ajuda a desenvolver melhor as suas aprendizagens, principalmente aos alunos com mais dificuldades de aprendizagem. (EAPM16)⁴

4 EAPMxy: Entrevista Aluno x, país Portugal, aula de matemática, referente ao Professor y.

Gosto da aula de matemática. A gente brinca, canta e aprende. (EABM62)

Eu gosto da aula de matemática, nem parece aula. (EABM61)

As aulas abalizadas pelo o que os estudantes já conheciam do conteúdo foram mais envolventes e participadas. As tarefas apresentavam-se mais desafiadoras e permitiam discussão coletiva. A avaliação acontecia, também e com mais ênfase, por meio das tarefas de ensino e não somente em momentos especiais de testes individuais. Nesse tipo de ambiência observamos maior disponibilidade dos alunos para perguntar e responder dúvidas emanadas entre eles ou de questionamentos feitos pelo professor, bem como maior interação entre pares a fim de finalizar uma dada tarefa.

Contudo, é possível sintetizar que, na maioria das aulas, houve um estilo de ensino que seguia as seguintes ações: o professor inicia a aula copiando na lousa o conteúdo a ser estudado ou lendo o livro didático, resolve na lousa um modelo de exercício para os alunos perceberem a lógica da resolução e, em seguida, propõe tarefas análogas às que havia resolvido e corrige os cadernos individual ou coletivamente. Os professores mencionaram planejar e organizar o ensino de modo prévio, sendo que as tarefas com potencial para ensinar, aprender e avaliar apareceram em menor quantidade.

Seguramente, podemos dizer que as práticas de avaliação de maior expressão nas escolas davam ênfase ao uso de

testes individuais, com foco na obtenção de informações sobre a diferença entre o que foi ensinado e o que o aluno responde corretamente neste instrumento, finalizando o processo com a atribuição de notas e classificações.

Mediante o estudo é possível assinalar que, na maioria das práticas avaliativas, a dinâmica foi deliberadamente unilateral, cabendo ao professor realizá-la. Os critérios em geral eram subtendidos pelos instrumentos selecionados (testes individuais, resolução de listas de exercícios e atitudes relacionadas ao comportamento em sala de aula). Embora tivéssemos identificado variedades de instrumentos em algumas dinâmicas avaliativas, estes eram centrados em resultados (produto final) individuais, sem possibilidade para revisões e sem o envolvimento coletivo para a autoavaliação ou heteroavaliação. Vejamos alguns relatos:

Eu observo, eu faço atividade individual, eu faço teste, eu faço prova, tem um período que eles fazem prova mesmo, até as atividades no caderno eu verifico. Eu tenho um mapa que eu coloco com todas as atividades que eu passo para eles [...] e aí eu conto como bônus [as atividades realizadas] [...] então eu procuro variar um pouco na avaliação. (EPBM7)

Eu os deixo à vontade. Como a turma se conhece há algum tempo, eles se organizam em grupos entre eles. Eles pedem para se organizarem. Agora, quando o trabalho é avaliativo, eu exijo que seja individual. (EPBM6)

O foco em resultados aparece muito atrelado às preocupações dos professores com os testes demandados por avaliações externas. Parece haver uma sujeição das avaliações do contexto escolar para com as provas externas, no sentido que aquelas precisam repetir “modelos” postos por estas a fim de promover maior segurança para com os resultados dos diagnósticos (de desempenho dos estudantes) realizados pelos órgãos competentes (Ministério da Educação/Secretarias de Educação).

Na maioria dos casos, foi observada uma certa preocupação em elaborar tarefas avaliativas que estejam em consonância com o ensino realizado, ou seja, cuidado em não

apresentar como tarefa de avaliação algo que ainda não era conhecido pelo estudante, como um “desafio” que surgiu pela primeira vez em uma tarefa eminentemente de avaliação. No entanto, o ensino nas escolas belenenses aparece regulado pela matriz de referência de provas externas e não por um currículo que aborde conteúdos necessários às melhores aprendizagens daqueles alunos. Diz uma professora:

Agora essa [prova] que a gente faz em sala, eu passo uma parecida, eu não passo igual, até mesmo antes da avaliação. Se eu já formulei a prova, eu passo uma parecida, mudando só os números e alguns dados, para quando eles chegarem na prova, eles “ah, já fiz isso aqui. Não é tão desconhecido pra mim”, mas em alguns momentos sim, é essa prova que eu faço [corrige] com eles. Mas eu tenho corrigido mais a prova municipal [prova externa vinda da Secretaria Municipal de Educação]. Essa maior que veio da prefeitura e um pré-teste que eles [Secretaria] fazem. Aí eu corrijo com eles [alunos] em sala, converso com eles sobre a nota “quem tirou tal, precisa estudar mais. Quem tirou tal está bom, mas precisa estudar”. (EPBM1)

Também houve quem transgredisse a rota comum de guiar as avaliações do contexto escolar por instrumentos análogos aos elaborados pelas avaliações externas.

Tudo aquilo que fazemos [fichas de trabalho] é corrigido na sala por eles [alunos]. Eles veem as coisas que erraram ou que não estavam totalmente certas [...]. Eles fazem [a correção com caneta], não é marcar certos nem errados, fazem a correção e eu vou vendo o que é que cada fez bem ou fez menos bem, portanto eu sei o que cada um está a fazer. [...] Partilhamos o mesmo, eles sabem e eu sei. (EPPM2)

Os alunos manifestavam-se mencionando experiências de avaliação que para eles eram significativas:

Eu não gosto de fazer nenhuma prova. Gosto das atividades em sala de aula quando o professor confecciona [os objetos de apoio ao ensino] como as garrafas, jornais, pipas, fitas e as aulas com o violão. (EABM62)

É possível dizer que a avaliação precisa ser problematizada como prática constituinte das relações ensino e aprendizagem e, para tal, cabe ao professor indicar o que ensinar e os critérios de sucesso a alcançar. Estas ações são fundamentais para uma prática avaliativa que vise à promoção da aprendizagem. É necessário compreender que a avaliação tem forma e conteúdo próprios e que conhecê-los torna-se fundamental para práticas avaliativas deliberadamente compreendidas em seus fins.

A respeito da participação dos alunos, pode-se dizer que, embora compreendida como importante pelos professores, apareceu de maneira não estruturada e geralmente apoiada por perguntas feitas pelo professor. A participação em geral ficou reservada aos que se mostravam potencialmente mais capazes em relação ao conteúdo ensinado.

Em algumas práticas pedagógicas observadas havia distribuição de tarefas para grupo de estudantes e, também, de forma individual. As discussões, quando ocorriam, eram tratadas coletivamente. Assim, as “questões formuladas” e o “*feedback* distribuído” não faziam parte de um planejamento e de uma organização do ensino de forma deliberada, de modo a provocar participação ativa do aluno. Ambos os aspectos surgiam na sala de aula espontaneamente no decorrer das aulas.

Era comum uma expectativa estar voltada para resultados corretos e pouco direcionada aos processos de discussão coletiva que poderiam estar sendo produzidos. “Só participamos da aula e respondemos de vez em quando, quando a professora faz perguntas” (EABM71), disse um aluno. Assim, em termos de participação, é possível dizer que os alunos centraram suas dúvidas na figura do professor e tiveram reforço dele para isso: “A gente faz tudo e ela coloca a resposta no quadro para a gente ver se acertou. Se não acertamos, ela explica de novo como fazer” (EABM3), mencionou outro aluno. A participação coletiva dos estudantes, em contextos de heteroavaliação, por exemplo, algumas vezes foi estimulada pelos professores, muito embora a expectativa se voltasse para o resultado correto como principal foco de atenção. Mesmo com a presença destes contextos estimulados pelo

professor, a heteroavaliação não era considerada constituinte para a avaliação de aprendizagens dos alunos.

De modo induzido, via trabalho em grupo, ou de modo espontâneo, os alunos tendem a se auxiliarem na realização das tarefas. “Uma coisa que eu não sei e o amigo sabe, aí ele vem e me ajuda, e se eu sei eu o ajudo” (EABM71). Porém, em momentos específicos de avaliação por meio de testes, por exemplo, esse tipo de comportamento não era aproveitado para o compartilhamento de aprendizagens. Em nossas observações não houve caso em que a avaliação se valesse de testes em pares, por exemplo, apoiados na comunicação entre os estudantes e nas aprendizagens conjuntas, ou ainda avaliações em grupos em que houvesse necessidade de auxílio entre eles para que o grupo pudesse alcançar êxito. Geralmente, as tarefas em grupos também ficavam centradas na entrega de um resultado e não no processamento do trabalho requerido.

CONCLUSÕES

Compreendemos que as práticas letivas de sala de aula de matemática nos anos iniciais, nos contextos brasileiro e português, apontam para um desejo de mudança, uma vez que os professores indicam conhecer a necessidade de uma avaliação para a aprendizagem; no entanto, articular esta pretensão de modo a implementá-la ainda precisa de uma sólida formação que possibilite tal articulação. Pensar uma avaliação que concorra para a aprendizagem sugere um ensino problematizador e dinâmico. A organização do ensino parece ser um aspecto fundamental para que dimensões como a participação dos alunos e a natureza e modalidades da avaliação sejam mobilizadas no sentido de favorecer *feedback* qualitativo para todos os envolvidos no processo. Potencializar a organização das tarefas é essencial para articular o ensino, a aprendizagem e a avaliação. Podemos dizer que a avaliação orientada para ajudar os alunos a aprender existiu de forma muito tênue, não deliberada e, muitas vezes, confundida com a distribuição de *feedback* que, geralmente, não foi variado, gerando pouca interlocução entre os envolvidos.

O *feedback*, essencial na comunicação entre professores e estudantes, precisa atingir o propósito de dar indicações claras aos estudantes sobre onde estão e qual o esforço necessário para melhorar em termos de suas próprias aprendizagens. Perguntas do tipo “entenderam?” e “alguma dúvida?”, feitas pelos professores em sala de aula, bem como indicações tipo “vocês precisam estudar mais” e “vocês não estão estudando direito”, não são suficientes para que os estudantes compreendam sobre as aprendizagens que lhes são inerentes. São necessárias indicações específicas, claras e positivas para que os estudantes sejam capazes de identificar e superar dificuldades e compreender seus avanços durante a realização das tarefas. Com isso, o professor também pode regular seu ensino fazendo reforços ou modificações no planejamento e na elaboração das tarefas propostas aos estudantes com vistas à melhoria das aprendizagens. O papel de avaliador de suas próprias aprendizagens ou das aprendizagens de pares esteve quase que inexistente nas práticas observadas e nos relatos dos alunos e professores. Ainda há certo receio quanto ao uso de auto e heteroavaliações, por conta do entendimento culturalmente construído de que avaliação é denominação de classificações/outorga de certificações em concorrência com outros e, portanto, a *permissão* para autoavaliações seria imprudência ou permissividade do professor. A incorporação da avaliação de si e dos pares, por parte dos alunos, depende da orientação do professor e esse parece ainda não ter clareza e segurança de que práticas como esta colaboram com as aprendizagens.

Percebemos a necessidade de investir tanto em pesquisas como em formação docente sobre o domínio da avaliação numa perspectiva formativa e, portanto, a serviço de quem aprende e de quem ensina. Mas não só. É necessário também considerar a problemática em questão na relação com outros espectros que compõem a realidade educacional. A precariedade, muitas vezes estabelecidas nos contextos das escolas públicas, embota as possibilidades de existência da avaliação formativa e, conseqüentemente, melhores resultados nas aprendizagens de estudantes e do ensino.

Para além do exposto, é certo indicar que as práticas letivas de matemática nos anos iniciais, embora apresentem

um estilo que conserva a tradição da avaliação como uma marca social e cultural para a classificação, contam com a vigilância, o esforço e a contribuição de professores que, mesmo não articulando – ainda – as vertentes e dimensões aqui discutidas a respeito da integração da avaliação, do ensino e da aprendizagem de forma majoritária, apontam para o fortalecimento de que avaliar para as aprendizagens é um desejo e, fundamentalmente, uma possibilidade para uma escola melhor.

REFERÊNCIAS

- BLACK, Paul; WILLIAM, Dylan. The reliability of assessments. In: GARDNER, John (Ed.). *Assessment and learning*. London: Sage, 2006. p. 81-100.
- BORRALHO, António; LUCENA, Isabel; BRITO, Maria Augusta. *Avaliar para melhorar as aprendizagens*. Belém: SBEM, 2015.
- FERNANDES, Domingos. *Avaliar para aprender: fundamentos, práticas e políticas*. São Paulo: Unesp, 2008.
- FERNANDES, Domingos. Professor da Universidade de Lisboa defende o foco na aprendizagem. *O Liberal*, Belém-PA, 17 jun. 2016. *Atualidades, Caderno Cidades*, p. 8.
- LEAHY, Siobhan et al. Classroom assessment: minute by minute, day by day. *Educational Leadership*, Alexandria, VA, v. 63, n. 3, p. 19-24, 2005.
- MIDDLE STATES COMMISSION ON HIGHER EDUCATION. *Student learning assessment: options and resources*. 2. ed. Philadelphia, PA: MSCHE, 2007. p. 13-30.
- PIRES, Magna N. Marin; BURIASCO, Regina. L. Corio de. Prova em fases: instrumento para aprender. In: SEMINÁRIO INTERNACIONAL DE PESQUISA EM EDUCAÇÃO MATEMÁTICA, 5., 2012, Petrópolis. *Anais...* Petrópolis: SBEM, 2012. p. 1-17.
- PONTE, João P. Tarefas no ensino e na aprendizagem da Matemática. In: PONTE, João P. (Org.). *Práticas profissionais dos professores de matemática*. Lisboa: Universidade de Lisboa/Instituto de Educação, 2014. p. 13-30.
- SILVA, Nazaré et al. Avaliação formativa: o feedback como instrumento potencializador da aprendizagem em matemática. In: ENCONTRO NACIONAL DE EDUCAÇÃO MATEMÁTICA, 12., 2016, São Paulo. *Anais...* São Paulo: SBEM, 2016. p. 1-12.
- STEIN, Mary Kay; SMITH, Margaret. Mathematical tasks as framework for reflection: from research to practice. *Mathematics Teaching in the Middle School*, v. 3, n. 4, p. 268-275, 1998.

TUNSTALL, Pat; GIPPS, Caroline. How does your teacher help you to make your work better? Children's understanding of formative assessment. *The Curriculum Journal*, London, v. 7, n. 2, p. 185-203, 1996.

VIEIRA, Isabel. *A autoavaliação como instrumento de regulação da aprendizagem*. 2013. 161 f. Dissertação (Mestrado em Supervisão Pedagógica) – Departamento de Educação e Ensino a Distância, Universidade Aberta, Lisboa, 2013.

ISABEL CRISTINA RODRIGUES DE LUCENA

Professora do Instituto de Educação Matemática e Científica da Universidade Federal do Pará (UFPA), Belém, Pará, Brasil
ilucena19@gmail.com

JOSETE LEAL DIAS

Professora da Escola de Aplicação da Universidade Federal do Pará (UFPA), Belém, Pará, Brasil
lealdiasjosete@gmail.com

ANTÓNIO MANUEL ÁGUAS BORRALHO

Professor do Centro de Investigação em Educação e Psicologia da Universidade de Évora, Évora, Portugal
amab@uevora.pt

Recebido em: 15 NOVEMBRO 2017

Aprovado para publicação em: 11 JANEIRO 2018

INDICADORES EDUCACIONAIS DE MATRÍCULAS DE ALUNOS COM DEFICIÊNCIA NO BRASIL (1974-2014)

**ANDRESSA SANTOS REBELO
MÔNICA DE CARVALHO MAGALHÃES KASSAR**

RESUMO

O uso de dados estatísticos foi por algum tempo preterido por parte da literatura educacional. Ciente desse movimento, este trabalho propõe-se a analisar a utilização de indicadores educacionais para conhecimento da realidade desse campo. São apresentados e analisados indicadores educacionais referentes a matrículas de alunos da Educação Especial no Brasil entre 1974 e 2014. Os dados são estudados sob uma perspectiva crítica, em interlocução com a literatura da área. O texto questiona a forma como tais informações têm sido usadas, em especial pelo Governo Federal, como demonstrativos dos avanços da política de inclusão escolar. Conclui-se que tais indicadores necessitam ser articulados a uma leitura que considere os impactos das ações dos governos, além de ter como interlocução dados e pesquisas que propiciem diferentes leituras.

PALAVRAS-CHAVE EDUCAÇÃO ESPECIAL • INDICADORES EDUCACIONAIS • POLÍTICAS EDUCACIONAIS • INDICADORES SOCIAIS.

INDICADORES EDUCACIONALES DE MATRÍCULAS DE ESTUDIANTES CON DISCAPACIDAD EN BRASIL (1974-2014)

RESUMEN

El uso de los datos estadísticos fue por algún tiempo desconsiderado por parte de la literatura educacional. Consciente de ese movimiento, este trabajo se propone a analizar la utilización de indicadores educacionales para el conocimiento de la realidad de este campo. Son presentados y analizados los indicadores educacionales de matrículas de alumnos de Educación Especial en Brasil entre los años 1974 y 2014. Se estudian los datos desde una perspectiva crítica, manteniendo interlocución con la literatura del área. El texto cuestiona la forma como tales informaciones han sido utilizadas, especialmente por el Gobierno Federal, como muestras de los avances de la política de inclusión escolar. Se concluye que hay que articular tales indicadores a una lectura que considere los impactos de las acciones de los gobiernos y, además, que tengan interlocución con datos e investigaciones académicas para que sean posibles distintas miradas.

PALABRAS CLAVE EDUCACIÓN ESPECIAL • INDICADORES EDUCACIONALES • POLÍTICAS EDUCACIONALES • INDICADORES SOCIALES.

EDUCATION INDICATORS RELATED TO THE ENROLLMENT OF PUPILS WITH DISABILITIES IN BRAZIL (1974-2014)

ABSTRACT

The use of statistical data has been underestimated by part of educational literature for some time. Being aware of this, the present paper aims to discuss the use of educational indicators to understand the situation of this field. The text presents and analyzes education indicators regarding the enrollment of Special Needs pupils in Brazil between 1974 and 2014. The data are analyzed from a critical perspective, in connection with the area literature. The study challenges the way in which the Federal Government has been using such data, in order to demonstrate advances in school inclusion policy. The conclusion is that such indicators need to be developed from a perspective that considers the impacts of government action, and to be connected with other data and academic research.

KEYWORDS SPECIAL EDUCATION • EDUCATIONAL INDICATORS • EDUCATIONAL POLICIES • SOCIAL INDICATORS.

INTRODUÇÃO

Até os anos 1970, julgou-se que a pesquisa em avaliação prescindia de tomada de posição, de valores, que era objetiva, técnica e distante da política (ROSEMBERG, 2013). Isso se deve à existência de uma dualidade entre teoria e empiria, tradição na pesquisa em educação no país, influenciada pela pesquisa social americana e europeia. Ambas as formas de pensamento impactaram as pesquisas no Brasil, mas não foram, segundo a avaliação de Goergen (1985, p. 3-4), objeto de estudos epistemológicos mais sistemáticos e continuados:

Talvez seja esta uma das razões da existência paralela e isolada das duas formas de investigação: uma, voltada para os textos, aspectos históricos, fundamentos, de caráter teórico-filosófico e que não dá muita atenção aos fatos reais; outra, dedicada às pesquisas empíricas, obcecada, muitas vezes, pelo afã de juntar dados e desatenta aos elementos explicativos que se escondem por detrás dos fatos.

Posicionamentos exclusivos como esse foram amplamente criticados no campo da pesquisa em educação (FRANCO; GOLDBERG, 1976; FRANCO, 1984). Também marcou presença a crítica da até então ausência de pesquisas com o objetivo de avaliar programas ou projetos educacionais (GOUVEIA, 1976). A partir dos anos 1980, a pesquisa em educação sofreu uma inflexão, vinculada ao esforço de se conjugar teoria-empíria. Assim, embora sujeita a críticas, a pesquisa educacional tem procurado avançar nesse sentido (GAMBOA, 1994).

Desse histórico, tem origem a defesa de que a pesquisa em políticas públicas deve assumir a não neutralidade, independentemente dos procedimentos metodológicos que use. Ao se adotar uma perspectiva política, a pesquisa avaliativa estabelece e declara suas lealdades em relação à sua função, seus objetivos, métodos, procedimentos, ética e divulgação de resultados (ROSEMBERG, 2013).

Cabe revelar algumas características das tradições da pesquisa em política pública. Conforme Souza (2007), na Europa, a área de política pública surgiu no século XX como um desdobramento dos trabalhos baseados em teorias explicativas sobre o papel do Estado e de uma das mais importantes instituições do Estado, o governo, produtor por excelência de políticas públicas. Nos Estados Unidos, a área ingressou no mundo acadêmico sem estabelecer relações com as bases teóricas sobre o papel do Estado, passando diretamente para a ênfase nos estudos sobre a ação dos governos.

Rosemberg (2001) faz a apreciação de que, na avaliação de políticas públicas, o conhecimento acadêmico deve se desvencilhar da visão economicista predominante em pesquisas sociais, carregada da aura de neutralidade científica, inquestionável, já que objetiva e orienta opções absolutas de prioridades e estratégias em políticas públicas. Sob essa ótica, decisões políticas são vestidas do qualificativo de ‘decisão técnica’ competente, diminuindo o poder dos demais atores sociais. Esse posicionamento pretensamente objetivo/imparcial vem sustentando um olhar para a política educacional, de modo a fundamentar procedimentos de avaliação e de responsabilização docente (SHIROMA; EVANGELISTA, 2011; VIDAL; VIEIRA, 2011). Tais ações estão fortemente

embasadas em índices divulgados publicamente, o que impacta a compreensão da realidade educacional (BROOKE, 2013). Ainda em relação ao uso de dados, Jannuzzi (2014a) chama a atenção para o fato de que, desde a implantação de um programa governamental, as demandas por informação e conhecimento têm se voltado à avaliação de seus resultados e impactos.

Se, por um lado, esses números e índices têm sido empregados para justificar ações e programas por diferentes governos, por outro, a pesquisa em educação ainda tem se utilizado pouco de dados quantitativos. Por esses e outros argumentos, pesquisas acadêmicas que abordem a metodologia adotada pelas avaliações oficiais ou que usem os dados disponíveis colaboram para que as universidades adquiram independência em relação ao vasto material que circula como verdade (FREITAS, 2013). A apropriação das estatísticas públicas pela comunidade acadêmica, portanto, é estratégica. Na educação propriamente, esse movimento se faz indispensável, tendo em vista a amplitude das políticas nesse setor, que atingem grande parcela da população, principalmente a partir da expansão da educação básica no Brasil no decorrer do século XX.

Na primeira metade do século XX, Lourenço Filho (1947, p. 70) já ressaltava a relevância das estatísticas escolares para atender à realidade educacional. Entendia que a despreocupação com os problemas do ensino primário, até essa época, “como obra nacional”, devia-se, na maior parte, à falta de levantamentos estatísticos periódicos que viessem a atestar o andamento ‘excessivamente vagaroso’ do desenvolvimento geral dos sistemas escolares estaduais.

Em relação ao uso da estatística na pesquisa em Educação, Sass (2012, p. 131) observa, na atualidade, reações diferentes:

As reações divergentes suscitadas pelas ciências que oscilam entre a desconfiança absoluta e a fé cega, compreensíveis de se manifestarem pelo senso comum e preocupantes à medida que grassam em meios acadêmicos, adquirem expressões singulares quando referidas às ciências particulares, como é o caso da Estatística. Na

educação é frequente a crítica, muitas vezes procedente, de que as estatísticas invadem a educação, ocultando os problemas educacionais reais, na mesma proporção que educadores e políticos clamam por elas, em outras circunstâncias, sob o argumento da falta de informações consistentes que orientem as proposições e programas educacionais.

Amplamente utilizada no início das pesquisas educacionais, a metodologia quantitativa se fazia presente em centros de pesquisa (LOURENÇO FILHO, 1964; GATTI, 2002). Passou a ser desacreditada a partir dos anos 1980 na pesquisa em Educação, sob o argumento de se tratar de um *método sociológico positivista* (SANTOS FILHO; GAMBOA, 2009). Se, por um lado, essa mudança de posição permitiu um grande desenvolvimento das metodologias qualitativas e o reconhecimento de sua legitimidade, por outro, ajudou a corroborar a retirada da Estatística da pesquisa em Educação (FERRARO, 2012), apesar do trabalho de recuperação desses estudos (FERRARO, 1979; GATTI, 1982) desde então.

O Brasil possui, atualmente, um conjunto notável de dados. Embora a sua coleta seja, a princípio, direcionada ao que o poder público pretende avaliar como resultado de determinada ação e existam problemas na metodologia do processo e no modo pelo qual são coligidos (JANNUZZI, 2009; PEREIRA; MAZZON; ROSENBERG, 1998; SASS, 2013), as informações podem ser examinadas pela academia como elementos para a avaliação de políticas públicas. Dentre esse conjunto de dados, há os obtidos através do Censo Escolar da Educação Básica, que fornecem elementos para a compreensão de aspectos da materialização das ações governamentais no campo da educação. Dentre os dados educacionais, estão os referentes a uma população que tem sido identificada como alunos da 'Educação Especial'. Este estudo enfoca, justamente, indicadores educacionais referentes às matrículas desses alunos no Brasil de 1974, data do primeiro levantamento dedicado à identificação de alunos "excepcionais" (termo utilizado à época), até 2014, ano em que se encontram os dados mais recentes sobre essa população¹, com o propósito de analisar a

¹ Os dados de 2015 foram disponibilizados após a finalização deste artigo.

utilização de indicadores educacionais para o conhecimento da realidade educacional.

Procurou-se examinar as características das políticas públicas de educação implantadas no período, destinadas a esse grupo, e a produção acadêmica relativa ao campo de investigação.

A pesquisa recorreu aos dados organizados pelo Ministério da Educação e Cultura (MEC) nos anos 1970, aos levantamentos estatísticos empreendidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e aos microdados obtidos pelo Censo Escolar da Educação Básica (MEC/Inep). Para a leitura e o tratamento estatístico dos microdados, utilizou-se o *software IBM (International Business Machines) SPSS Statistics*.

INDICADORES EDUCACIONAIS DE MATRÍCULAS DE ALUNOS DA ‘EDUCAÇÃO ESPECIAL’

Diversos organismos que produzem estatísticas, assim como entidades que coletam dados administrativos, realizam inquéritos censitários ou amostrais, reunindo informações sobre um conjunto de unidades de observação. Em geral, os microdados são organizados em bases de dados em que cada linha, ou registro, contém informações (sistematizadas pelas diversas variáveis) sobre cada unidade de observação (MARTINS; MAURITTI; COSTA, 2013). O termo ‘micro’ indica que os dados representam características das unidades mínimas do levantamento – por exemplo: pessoas, empresas, ônibus –, e não os dados agregados em níveis superiores – tais como: municípios, setores econômicos, frotas de transporte urbano. Dados agregados não podem ser cruzados ou reprocessados; o processo de agregação é irreversível para quem não dispõe dos dados originais (SCHRADER, 2002).

Os indicadores sociais, sob uma perspectiva programática, constituem um instrumento operacional para o monitoramento da realidade social, para fins de formulação e reformulação das políticas públicas (JANNUZZI, 2009). Pelo fato de a educação ser uma área de intervenção das políticas públicas e versar sobre a realidade social de grande parte da população,

entende-se que todo indicador educacional seja um indicador social, pois leva à avaliação de um dos aspectos de suas condições de vida (JANNUZZI, 2005; JANNUZZI et al., 2009).

Dessa forma, toma-se por conceito de *indicador social*:

[...] uma medida em geral quantitativa dotada de significado social substantivo, usado para substituir, quantificar ou operacionalizar um conceito social abstrato, de interesse teórico (para pesquisa acadêmica) ou programático (para formulação de políticas). É um recurso metodológico, empiricamente referido, que informa algo sobre um aspecto da realidade social ou sobre mudanças que estão se processando na mesma. (JANNUZZI, 2009, p. 15)

Partindo desse conceito, para o estudo pretendido, delimitou-se por temática social de interesse a *escolarização do grupo de alunos classificado como da “Educação Especial”*. Essa população vem recebendo diferentes denominações nos registros estatísticos do Governo Federal: “alunos excepcionais”, nos documentos dos anos 1970; “alunos portadores de deficiências, condutas típicas e altas habilidades”, entre os anos 1980 e 1990; “pessoa portadora de necessidades especiais”, no início dos anos 1990; “alunos com necessidades educacionais especiais”, entre o final dos anos 1990 e o início do século XXI; e, atualmente, “aluno com deficiência, transtornos globais do desenvolvimento, altas habilidades/superdotação”, seguindo orientação do documento Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva (BRASIL, 2008a). Tendo em vista essas alterações, ao longo do texto, quando se faz menção a ‘alunos da Educação Especial’, refere-se aos alunos público-alvo da educação especial em cada momento histórico.

As dimensões operacionais de interesse (indicador social) são o atendimento escolar dessa população, segundo *tipo de atendimento* – escola especial, classe especial, classe regular com apoio de sala de recursos, classe regular sem apoio, salas de recursos, entre outros – e *matrículas* – dependência administrativa, estabelecimentos e turmas.

No período de 1974 a 2006, utilizaram-se como fonte privilegiada as sinopses estatísticas existentes; de 2007 a

2014, os microdados do Censo Escolar (MEC/Inep), divulgados pelo Ministério da Educação. Embora o banco de microdados do Inep esteja disponível desde 1995, entre esse ano e 2006, a ‘escola’ era a unidade de coleta. Assim, por meio do código da escola, concedia-se acesso às matrículas, obtendo as características dos alunos de escolas específicas. A partir de 2007, a unidade de coleta passa a ser o ‘aluno’, derivando daí a dificuldade de um estudo incidir sobre uma série histórica que perpassasse esses dois períodos.

Ressaltam-se outros aspectos que marcam a data. A partir de 2007, a política educacional proposta a essa população tem duas metas claramente definidas: sua matrícula em classes comuns do ensino regular e o oferecimento de atendimento educacional especializado em Salas de Recursos Multifuncionais, no contraturno escolar. Foram identificados, como materialização dessa política, a criação do Programa de Implantação de Salas de Recursos Multifuncionais (BRASIL, 2007a) e o direcionamento de recursos para esse fim, através da dupla matrícula de alunos da Educação Especial na escola, que recebem atendimento educacional especializado, proporcionado pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb) (BRASIL, 2007b).

Sobre o atendimento educacional especializado oferecido em Salas de Recursos Multifuncionais, os dados passaram a ser compilados e disponibilizados apenas a partir de 2009.

O ano de 1995 é marcante para o sistema de informação estatística educacional no Brasil. Nessa data, foi reformulada a Diretoria de Informações e Estatísticas Educacionais (Sec) do Inep, até então responsável por organizar os dados censitários educacionais, e criado o Sistema Integrado de Informações Educacionais (SIEd). Até janeiro de 1995, o último Censo da Educação Básica concluído e divulgado era o de 1989. A participação em projetos internacionais para a produção de indicadores educacionais e a reestruturação do sistema de informações e estatísticas educacionais são consequências também de uma mudança da forma de fazer a política educacional. A partir de 1995, a informação deixou de ser coadjuvante para assumir um papel central como instrumento

para a formulação e a implementação das políticas educacionais. O Plano Diretor da Reforma do Aparelho do Estado (BRASIL, 1995) atingiu o Ministério da Educação, que precisava de um sistema de informação eficaz e confiável para a avaliação adequada do sistema educacional brasileiro, a fim de tornar os dados coletados relevantes para as administrações nas instâncias escola, município e estado (ALMEIDA; WOLYNEC, 1998; DINIZ, 1999).

Essas alterações indicam a racionalização da organização do Estado, emanando decisões qualificadas como “técnicas e competentes”, aspecto já abordado no início deste artigo, e sua penetração nos diferentes níveis administrativos. Sobre esse movimento, Jannuzzi (2014b, p. 12) acrescenta que o fortalecimento da agenda produtora de informação estatística e da estruturação dos sistemas de monitoramento e avaliação na administração pública conduz ao questionamento sobre o “mito da falta de racionalidade técnica”, na decisão política do país, e à ausência de avaliação dentro do próprio governo.

A historicidade de um indicador é a propriedade de se dispor de séries históricas extensas e comparáveis, de modo a cotejar o valor presente com situações do passado, inferir tendências e avaliar efeitos de eventuais políticas sociais implementadas. O ideal é que as cifras passadas sejam compatíveis entre si sob o ponto de vista conceitual e com confiabilidade similar. Contudo, mesmo em situações em que isso não seja possível, as tendências passadas podem dar informações úteis para a interpretação das cifras atuais dos indicadores. Poucas vezes se dispõe de séries históricas plenamente compatíveis de indicadores para a escala geográfica ou o grupo social de interesse (JANNUZZI, 2009).

Levando em conta essas dificuldades, mas também o potencial do estudo de dados longitudinais para a pesquisa educacional (BONAMINO; OLIVEIRA, 2013), foram analisadas, nesta pesquisa, as variáveis com continuidade nesses períodos, no que diz respeito aos atendimentos educacionais especializados. Admitiram-se os limites e mesmo os riscos de recorrer a fontes de naturezas distintas.

O interesse foi o de investigar o movimento de matrículas dos alunos da Educação Especial no país, a partir do

primeiro levantamento estatístico nacional sobre Educação Especial, publicado pelo Centro Nacional de Educação Especial (Cenesp) (BRASIL, 1975).

Durante o regime militar (1964-1985), a Secretaria Geral do Ministério da Educação e Cultura (MEC) criou um grupo-tarefa destinado a realizar a implantação do Cenesp e elaborar a sua estrutura e regimento interno (BRASIL, 1973a, 1973b), como consequência de uma preocupação já sinalizada em documentos educacionais anteriores (BRASIL, 1957, 1958, 1960). O Cenesp foi idealizado “com a finalidade de promover, em todo o território nacional, a expansão e melhoria do atendimento aos excepcionais” (BRASIL, 1973c, art. 1º). No ano de 1986, momento de redemocratização do país, foi fundada a Secretaria de Educação Especial no Ministério da Educação (BRASIL, 1986), com a extinção do Cenesp. A partir dos anos 1990, os pressupostos de integração/inclusão mostraram-se mais incisivos sobre as políticas sociais, quando se sinalizou a opção pela escola comum como lócus preferencial de atendimento educacional especializado (BRASIL, 1988, 1996, 2007a), não sem o enfrentamento e o fortalecimento das lutas empenhadas pelas instituições privadas de caráter filantrópico para que não perdessem espaço e influência entre essas políticas (BRASIL, 1988, 1996, 2007b). Em 2011, a Secretaria de Educação Especial foi extinta e a incumbência por suas ações foi repassada à Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão (Secadi) (BRASIL, 2011a). Dentro da sua estrutura organizacional, foi implementada a Diretoria de Políticas de Educação Especial.

Com o propósito de abranger o período indicado, elaborou-se a tabela a seguir, que permite a comparação do número de matrículas na Educação Especial em classes comuns, especiais e escolas exclusivas (escolas especiais) e o número de matrículas da população em geral em classes comuns do ensino regular de 1974 a 2014:

TABELA 1 – Distribuição do número de matrículas de alunos da Educação Especial e a relação com o número de matrículas da população em geral. Brasil, 1974 a 2014

ANO	MATRÍCULAS DA EDUCAÇÃO ESPECIAL						TOTAL (A)	MATRÍCULAS DA POPULAÇÃO EM GERAL TOTAL (B)	% (A: B)
	CLASSES COMUNS		CLASSES ESPECIAIS		ESCOLAS EXCLUSIVAS				
	%	N	%	N	%	N	TOTAL (B)	(A: B)	
1974	-	-	38,3%	36.986	33,6%	32.399	96.413	18.573.193	0,51
1988	-	-	43,6%	78.322	56,4%	101.201	179.523	24.769.736	0,72
1996	31,1%	90.379*	68,9%**		201.142**		291.521	44.584.650	0,65
1998	13,1%	43.923	24,6%	83.201	62,3%	210.202	337.326	43.928.519	0,76
2000	21,4%	81.695	21,3%	81.400	57,3%	219.120	382.215	49.294.849	0,77
2002	24,8%	110.704	17,4%	78.353	57,8%	259.544	448.601	54.716.609	0,81
2007	46,8%	306.136	11,9%	77.728	41,4%	270.742	654.606	53.028.928	1,23
2008	54,0%	375.775	10,7%	74.481	35,3%	245.443	695.699	53.232.868	1,30
2009	60,5%	387.031	8,4%	53.430	31,1%	199.257	639.718	52.580.452	1,21
2010	68,9%	484.332	6,6%	46.255	24,5%	172.016	702.603	51.549.889	1,36
2011	74,2%	558.423	5,0%	37.497	20,8%	156.385	752.305	50.972.619	1,47
2012	75,7%	620.777	3,8%	31.168	20,5%	168.488	820.433	50.545.050	1,62
2013	76,9%	648.921	3,6%	30.453	19,4%	163.968	843.342	50.042.448	1,68
2014	78,8%	698.768	3,0%	27.004	18,2%	161.043	886.815	49.771.371	1,78

Fonte: MEC (conforme indicações abaixo, respectivas referências deste artigo e notas explicativas).

a) Dados de 1974: Brasil (1975, 1984, 1985) e NEPP (1988); b) Dados de 1988: Brasil (1991a, 1991b) e NEPP (1988); c) Dados de 1996: Brasil (1997); d) Dados de 1998: Brasil (1998a, 1998b, 1998c); e) Dados de 2000 e 2002: Brasil (2000, 2002); f) Dados de 2007 a 2014: MEC/Inep/Diretoria de Estatísticas Educacionais/Censo Escolar da Educação Básica (microdados) e Brasil (2010a).

Notas: Foram encontrados os números de matrícula utilizados em documentos do Observatório do Plano Nacional de Educação que tem como fonte: MEC/Inep/Censo Escolar, 2007 a 2013. Os documentos foram elaborados por Todos pela Educação. Dados estatísticos, informes e sinopses foram publicados em anos posteriores à coleta dos dados. Por exemplo, os dados referentes a 1974 estão publicados em Brasil (1975). Os dados das matrículas gerais de 1974 e 1988 não foram localizados (da direita para esquerda, segunda coluna, segunda e terceira linhas). Assim, nessas duas células foram considerados os dados de matrículas gerais dos anos de 1973 e 1985, publicados em NEPP (1988), que tem como fonte MEC/SEEC.

*Inclui as matrículas em outros tipos de atendimento. **Nos dados referentes ao ano de 1996, os números de matrículas em "classes especiais" e "escolas especiais" não estão desagregados.

Elaboração própria, com base nas fontes apresentadas.

Os documentos que apresentam os dados de 1974 e 1988 não traziam anotações das matrículas de alunos em classes comuns, embora se saiba que existiam, pois havia o registro de alunos em salas de recursos e elas, em princípio, atendiam alunos matriculados em salas de aulas comuns. Esses lançamentos passaram a existir em 1996. No entanto, nesse ano, não há informações sobre o número isolado de matrículas em classes comuns (com apoio especializado), já que os dados originais não possibilitam desagregar essas matrículas dos outros tipos de atendimento (sala de recursos, oficinas pedagógicas, educação precoce, etc.). No ano de 1974, o total

percentual de classes especiais e escolas exclusivas (escolas especiais) não corresponde a 100% (38,3% + 33,6% não totalizam 100%). Do seu complemento, não se apreende o quanto correspondia apenas às matrículas em classes comuns, pois havia outras formas de atendimento. Em 1988, não se localizou o registro de matrículas em classes comuns; no ano de 1996, não foi possível desagregar os dados de classes especiais e escolas exclusivas, modalidade Educação Especial.

A partir de 1998, observa-se um aumento contínuo da proporção de matrículas de alunos da Educação Especial nas classes comuns, enquanto o número de matrículas em classes especiais e escolas exclusivas sofre oscilações até apresentar a menor proporção ao final de 2014. No ano de 2008, verificou-se inflexão do número de matrículas em espaços exclusivos e inclusivos. Até 2007, o total de matrículas em espaços 'segregados' (classes especiais + escolas exclusivas) era superior ao de alunos matriculados em classes comuns, os alunos 'incluídos' (348.470 matrículas 'exclusivas' e 306.136 matrículas 'inclusivas'), passando as matrículas em classes comuns a 54% em 2008. Em 2014, foram contabilizadas 188.047 matrículas em espaços exclusivos e 698.768 matrículas em classes comuns. Esse número corresponde a 78,8% do total das matrículas de alunos com deficiência na educação básica (886.815).

No geral, é interessante notar que, em 40 anos, o número de matrículas de alunos da Educação Especial aumentou nove vezes, enquanto o número de matrículas da população em geral na Educação Básica aumentou apenas 2,67 vezes. Por outro lado, a proporção das primeiras em relação ao total de matrículas na educação básica não chegou a 2% dos registros (última coluna, à direita). Cabe ainda ressaltar que o número de matrículas não designa, necessariamente, o número de alunos. Desde 2007, com a implantação do Fundeb, alunos que frequentam salas de recursos recebem dupla matrícula, como forma de incentivo às redes educacionais a aderirem à política implantada.

Ainda quanto à comparação entre o número de matrículas da Educação Especial e o total geral de matrículas, segundo os dados de 1998 a 2006, Meletti e Bueno (2010) criticam

a pequena proporção de matrículas de Educação Especial. De acordo com os autores, tendo em conta as estimativas de incidência de pessoas com deficiência da Organização Mundial de Saúde (OMS) (10% da população) e do Instituto Brasileiro de Geografia e Estatística (23,9% no Censo Demográfico de 2010) (IBGE, 2010), adotadas oficialmente no Brasil, uma parcela muito pequena dessa população recebe algum tipo de escolarização no país (MELETTI; BUENO, 2013). Em todos os casos, não há como afirmar se existe subnotificação ou excesso de notificação de pessoas com deficiência, na escola e fora dela.

Além dos problemas mencionados por diferentes autores em relação à metodologia do processo de coleta, ao modo pelo qual os dados são coligidos (JANNUZZI, 2009; SASS, 2012) e às questões específicas da Educação Especial (MELETTI; BUENO, 2010, 2013), sublinha-se que o número de anotações de matrículas do público-alvo da Educação Especial é direcionado pelas mudanças conceituais na determinação dessa população, dispostas pela legislação educacional do período, algo que está distante de ser trivial ou secundário na análise dessa temática.

Como exemplos, nos anos 1970, a identificação desses alunos direcionou-se pela Lei Educacional 5.692/71, que entendia serem merecedores de “tratamento especial” na educação os alunos que apresentassem “deficiências físicas ou mentais”, os que se encontrassem “em atraso significativo quanto à idade regular de matrícula” e os “superdotados”, como determinava o artigo 9º (BRASIL, 1971). Trabalhos acadêmicos sobre a época (PASCHOALICK, 1981; FERREIRA, 1993; KASSAR, 1994) denunciaram a transformação de alunos “atrasados” em relação à série escolar esperada em “alunos com deficiência mental educável”, por sua matrícula compulsória em classes especiais.

Nos anos 2000, o público da Educação Especial eram os alunos com necessidades educacionais especiais, alguns dos quais não tinham, necessariamente, deficiência por causa orgânica, conforme a definição da Resolução CNE/CEB n. 2/2001 (BRASIL, 2001). Posteriormente, a Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva,

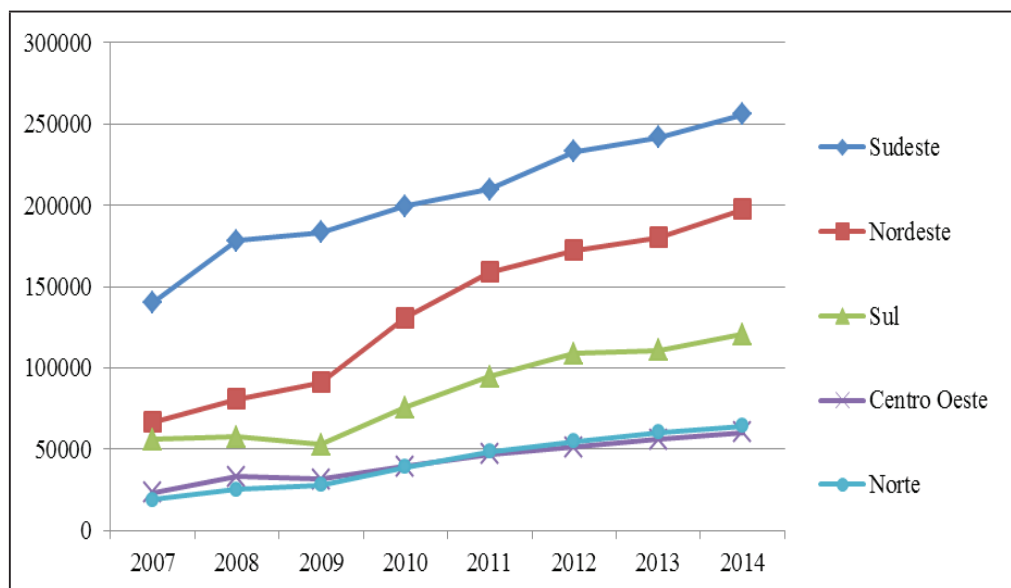
divulgada no ano de 2008, restringiu essa população a “alunos com deficiência, transtornos globais do desenvolvimento e altas habilidades ou superdotação” (BRASIL, 2008a, p. 6). É interessante observar que, apesar da restrição, não houve diminuição das matrículas. Ao contrário, ocorreu um contínuo crescimento.

Além de problemas que envolvem o diagnóstico de alguns casos, como avaliam Kassar (1994), Mendes (1995), Anache (1997), entre outros autores, há ainda outras questões. Um estudo elaborado por Dias (2014) aponta que a probabilidade de acontecerem erros no lançamento das informações é alta, não só por limitações dos técnicos das escolas (que usualmente são responsáveis por essa anotação), mas também por problemas relativos ao questionário do Censo e às possibilidades de registro. A autora constata que a escola relaciona diversas condições (dificuldade de aprendizagem, diabetes, problemas psiquiátricos) como deficiência, o que não só altera o resultado do Censo como também rotula o aluno como sendo da Educação Especial.

No tocante às instruções para preenchimento do Censo Escolar, ao verificar os Cadernos de Instruções (BRASIL, 2008b, 2009a, 2010b, 2011c, 2012a, 2013a) – documentos que servem de guia para preenchimento do sistema Educacenso –, Leijoto (2015) alerta que algumas categorias de deficiências estão listadas na identificação em alguns anos e em outros “desaparecem”, o que induz a erros nos relatórios dos casos. Esses aspectos indicam a imprecisão do significado do número de matrículas visto na Tabela 1.

O gráfico a seguir exhibe o total de matrículas de alunos da Educação Especial no ensino comum no Brasil segundo a região, a partir da publicação do Programa de Implantação de Salas de Recursos Multifuncionais e do incentivo de recursos proporcionado pelo Fundeb, no ano de 2007:

GRÁFICO 1 - Número de matrículas de alunos da Educação Especial em classes comuns, segundo a região do Brasil, 2007 a 2014



Fonte: MEC/Inep/Censo Escolar/Sinopse Estatística da Educação Básica. Brasil, 2007 a 2014 (elaboração própria).

O número de matrículas de alunos da Educação Especial aumentou em todas as regiões no período de 2007 a 2014. Considerando o foco da política educacional de universalização da educação no país, esse aumento é promissor, todavia insuficiente para conhecer as condições da escolarização desses alunos no sistema escolar como permanência, participação e desempenho (LAPLANE, 2015). Diferentemente de sucesso, trabalhos acadêmicos têm evidenciado limitações do funcionamento da escola quanto ao desenvolvimento acadêmico dessa população, com a identificação dos seguintes problemas: concentração de matrículas nas primeiras séries do ensino fundamental (MELETTI; RIBEIRO, 2014), pouco acesso ao conhecimento sistematizado (CARVALHO, 2013; PLETSCH, 2013) e pouco acesso ao atendimento educacional especializado (REBELO, 2012).

Os dados da Tabela 1 podem ser visualizados em contraste com as informações prestadas pelo documento Plano Nacional de Educação PNE 2014-2024: Linha de Base (BRASIL, 2015), elaborado pela Diretoria de Estudos Educacionais para

acompanhamento das metas e estratégias do Plano Nacional de Educação. O documento informa que, segundo o Censo Demográfico de 2010, de um total de 1.819.712 crianças e adolescentes de 4 a 17 anos que não conseguiam de modo algum ou tinham grande dificuldade para enxergar, ouvir, caminhar ou subir degraus, ou ainda possuíam alguma deficiência mental/intelectual permanente que limitava suas atividades habituais, 1.560.784 (85,8%) frequentavam escolas ou creches e 14,2% encontravam-se fora da escola. Se o documento consigna o número de 1.560.784 desses alunos na escola, o que explicaria apenas 886.815 matrículas no Censo Escolar de 2014 (BRASIL, 2007c, 2008c, 2009b, 2010c, 2011d, 2012b, 2013b, 2014c)?

² Disponível em: <<http://tabnet.datasus.gov.br/>>. Acesso em: 30 jul. 2016.

No que tange à incidência de deficiência no país, há ainda a Pesquisa Nacional de Saúde (IBGE)² de 2013, que assevera que a proporção de brasileiros com alguma deficiência é de 6,2%, enquanto o Censo Demográfico de 2010 traz um percentual de 23,9%. Diante de bases de dados com fontes e instrumentos de coletas diferentes que levam a resultados também diferentes, somente é possível constatar, quanto às matrículas de alunos da Educação Especial na educação básica, um aumento do número em classes comuns e uma diminuição em classes especiais e escolas exclusivas, mas não existe clareza quanto ao alcance da sua escolarização.

Apesar da implantação da política de educação inclusiva, ainda há formas de atendimento segregado, característica anteriormente constatada por alguns autores (PRIETO, 2010; KASSAR; MELETTI, 2012). A literatura (JANNUZZI, 1985; BUENO, 1991) mostra que a maior parte das escolas especializadas se encontra sob a administração privada (principalmente filantrópica). Para compreender a participação desse setor a partir do momento de inflexão das matrículas, separamos os dados em classes comuns, especiais e escolas exclusivas, conforme dependência administrativa, no período de 2007 a 2014, nas próximas três tabelas:

TABELA 2 - Distribuição do número de matrículas de alunos da Educação Especial em classes comuns, por dependência administrativa. Brasil, 2007 a 2014

ANO	TOTAL	PÚBLICA		PRIVADA	
		N	%	N	%
2007	306.136	285.923	93,4	20.213	6,6
2008	375.775	352.638	93,8	23.137	6,2
2009	387.031	365.796	94,5	21.235	5,5
2010	484.332	457.236	94,4	27.096	5,6
2011	558.423	525.812	94,2	32.611	5,8
2012	620.777	583.619	94,0	37.158	6,0
2013	648.921	609.839	94,0	39.082	6,0
2014	698.768	655.375	93,8	43.393	6,2

Fonte: MEC/Inep/Censo Escolar da Educação Básica (elaboração própria).

TABELA 3 - Distribuição do número de matrículas de alunos da Educação Especial em classes especiais, por dependência administrativa. Brasil, 2007 a 2014

ANO	TOTAL	PÚBLICA		PRIVADA	
		N	%	N	%
2007	76.765	-	-	-	-
2008	73.969	-	-	-	-
2009	53.430	50.269	94,0	3.161	6,0
2010	46.255	42.208	91,2	4.047	8,8
2011	37.497	33.684	89,8	3.813	10,2
2012	31.168	28.025	89,9	3.143	10,1
2013	30.453	26.538	87,1	3.915	12,9
2014	27.004	24.009	88,9	2.995	11,1

Fonte: MEC/Inep/Censo Escolar da Educação Básica. Matrículas em classes especiais do ensino regular e/ou da educação de jovens e adultos. Os dados de 2007 e 2008 não estavam disponíveis por completo nas fontes consultadas (elaboração própria).

TABELA 4 - Matrículas de alunos com deficiência nas escolas exclusivamente especiais, por dependência administrativa. Brasil, 2007 a 2014

ANO	TOTAL	PÚBLICA		PRIVADA	
		N	%	N	%
2007	270.742	57.216	21,1	213.526	78,9
2008	245.443	50.448	20,6	194.995	79,4
2009	199.257	38.862	19,5	160.395	80,5
2010	172.016	33.176	19,3	138.840	80,7
2011	156.385	29.400	18,8	126.985	81,2
2012	168.488	30.200	17,9	138.288	82,1
2013	163.968	28.089	17,1	135.879	82,9
2014	161.043	27.736	17,2	133.307	82,8

Fonte: MEC/Inep/Censo Escolar da Educação Básica (elaboração própria).

No período, mais de 90% das matrículas em classes comuns pertenciam ao setor público (Tabela 2). Nas classes especiais, o setor público também concentrava o maior número de matrículas. As classes especiais tiveram diminuição de matrículas ao longo dos anos, tanto no setor público quanto no privado (Tabela 3). As matrículas em escolas exclusivamente especiais confirmam as informações da literatura, que ressalta a força do setor privado nessa forma de atendimento (Tabela 4).

Para compreender os números da Educação Especial em termos da política educacional atual, cabe verificar a existência do atendimento educacional especializado.

MATRÍCULAS EM ATENDIMENTO EDUCACIONAL ESPECIALIZADO NO BRASIL

O Programa Implantação de Salas de Recursos Multifuncionais, conduzido pelo MEC/Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão, estabelecido por meio da Portaria Ministerial n. 13/2007 (BRASIL, 2007a), passou a integrar o Plano de Desenvolvimento da Educação (PDE) e o Plano Nacional dos Direitos da Pessoa com Deficiência – Viver sem Limite. O programa instituiu a forma e o lócus atuais (a política atual) do atendimento educacional especializado em implantação no país. No contexto da Política Nacional de Educação Especial na Perspectiva da

Educação Inclusiva, o programa objetiva: apoiar a organização da Educação Especial na perspectiva da educação inclusiva; assegurar o pleno acesso dos estudantes público-alvo da Educação Especial no ensino regular em igualdade de condições com os demais estudantes; disponibilizar recursos pedagógicos e de acessibilidade às escolas regulares da rede pública de ensino; e promover o desenvolvimento profissional e a participação da comunidade escolar (BRASIL, 2012c).

Aos gestores dos sistemas de ensino compete a definição da implantação das salas de recursos multifuncionais, o planejamento da oferta do Atendimento Educacional Especializado (AEE) e a indicação das escolas a serem inseridas, de acordo com as demandas da rede. Destaca-se que há a opção clara da escola pública como *locus* do programa: “a escola indicada deve ser da *rede pública de ensino regular*, conforme o Censo Escolar MEC/Inep (*escola comum*)” (BRASIL, 2012c, p. 10, grifo nosso), mesmo com a aprovação do Decreto n. 7.611/2011 (BRASIL, 2011b), que garantiu o AEE em instituições privado-assistenciais de caráter filantrópico.

Até o ano de 2014, segundo a versão preliminar do Relatório Educação para Todos 2000-2015, foram implantadas 42 mil salas de recursos multifuncionais com equipamentos, mobiliários, materiais pedagógicos e recursos de acessibilidade, para o atendimento educacional especializado, em 49% das escolas públicas de ensino regular com matrícula de estudantes da Educação Especial, localizadas em 93% dos municípios brasileiros. Segundo o Ministério da Educação, foram ofertados cursos de formação continuada para professores que atuam no atendimento educacional especializado dessas escolas, apoio financeiro para a acessibilidade arquitetônica nos prédios escolares e a aquisição de tecnologia assistiva (BRASIL, 2014b). Apresenta-se, a seguir, a comparação entre o número de matrículas em Atendimento Educacional Especializado no país e o número de matrículas no ensino regular de alunos da Educação Especial e a cobertura do programa. Como informado, os dados foram disponibilizados apenas a partir de 2009.

TABELA 5 – Atendimento educacional especializado complementar (AEE). Brasil, 2009 a 2014

ANO	NÚMERO DE MATRÍCULAS			% (B:A)	% (C:B)
	EDUCAÇÃO BÁSICA - TOTAL GERAL [A]	EDUCAÇÃO ESPECIAL EM CLASSES COMUNS - ALUNOS INCLUIDOS [B]	ATENDIMENTO EDUC. ESPEC. COMPLEMENTAR (AEE) [C]		
2009	52.580.452	387.031	129.697	0,74	33,5
2010	51.549.889	484.332	176.706	0,94	36,5
2011	50.972.619	558.423	231.740	1,09	41,5
2012	50.545.050	620.777	248.878	1,22	40,1
2013	50.042.448	648.921	290.799	1,29	44,8
2014	49.771.371	698.768	306.746	1,40	43,9
Varição (%)	(-) 5,34	80,5	136,5	0,75	10,4

Fonte: MEC/Inep/Censo Escolar da Educação Básica. Brasil, 2009 a 2014 (elaboração própria).

Vale lembrar que a Tabela 1, no início deste artigo, em sua última coluna, compara as matrículas da Educação Especial em todos os serviços e modalidades (escolas especializadas, classes especiais e classes comuns/ensino regular) em relação ao total de matrículas em geral. Já nessa Tabela 5, a quinta coluna exibe a comparação das matrículas de alunos da Educação Especial no ensino regular (classes comuns apenas) em relação ao total de matrículas na educação básica. Por esse motivo, as porcentagens nas duas tabelas são diferentes.

Na Tabela 5, constata-se que no período de 2009 a 2014 o número total de matrículas na educação básica diminuiu 5,34%, enquanto o número total de matrículas de alunos da Educação Especial nas classes comuns aumentou 80,5%. O número total de matrículas em AEE sofreu um aumento de 136,5%, mas a cobertura do programa aumentou apenas 10,4%. Apesar do grande número de estabelecimentos com salas de recursos que oferecem AEE, o alcance do programa não chegou a atender 50% dos alunos público-alvo da política, matriculados no ensino comum, o que sugere que grande parcela desses alunos pode não estar recebendo atendimento educacional especializado. A análise desses dados indica que a proposta, a princípio hegemônica, não chega a conquistar essa premência na execução da política na escola.

CONSIDERAÇÕES FINAIS

No início do texto, enfatizou-se a reserva da área da educação em relação a pesquisas que tratam de dados quantitativos e buscou-se analisar a utilização de indicadores educacionais para o conhecimento da realidade educacional. O foco versou sobre as matrículas de alunos na Educação Especial no Brasil entre 1974 e 2014, com a intenção de que fossem compreendidas, considerando aspectos da política educacional no período. Enfatizaram-se questões referentes à política atual, que preconiza a matrícula de todos os alunos na educação comum e o oferecimento de atendimento educacional especializado no contraturno, quando necessário.

Apesar das limitações evidenciadas na literatura (JANNUZZI, 2009; SASS, 2012; MELETTI; BUENO, 2010, 2013), algumas das quais citadas neste texto, os dados oficiais são utilizados pelos governos para propor e justificar ações públicas, sob uma perspectiva pragmática, que tem se sobreposto a uma avaliação crítica, afetando as condições de vida de grande parte da população (POPKEWITZ; LINDBLAD, 2001).

O aumento das matrículas nas classes comuns tem sido celebrado pelo Ministério da Educação como indicador do sucesso da política de inclusão escolar,

[...] mostrando a efetivação da educação inclusiva e o empenho das redes de ensino em envidar esforços para organizar uma política pública universal e acessível às pessoas com deficiência. (BRASIL, 2014a, p. 25)

A afirmação de sucesso da política de educação inclusiva remete à adoção, pelo Governo Federal, de um diagnóstico de certa forma superficial e parcial dos dados existentes. Pelo modo como os dados aparecem nos relatórios oficiais, não se verifica, por exemplo, que, em 40 anos, a despeito da implantação da política de educação inclusiva e do aumento do número de matrículas de alunos da Educação Especial, a sua proporção em relação ao total de matrículas na educação básica não chega a 2% (pois, em 40 anos, também cresceu a taxa de matrícula geral da população). Considerando que o número de matrículas não designa necessariamente o número de alunos nas escolas, devido às duplas matrículas, de

fato não se sabe ao certo qual foi o avanço da escolarização da população-alvo da Educação Especial no período.

Para a investigação sobre a situação de escolaridade dessa população, devem ser apreciados, ainda, estudos como os nomeados neste texto, que questionam a utilização apenas de indicadores de matrícula como demonstrativos dos avanços da política, sem incluir outros indicadores e diferentes informações que favoreçam a averiguação dos processos de permanência desses alunos na escola, bem como informações sobre as condições materiais e objetivas em que vem ocorrendo a sua escolarização. Diferentes pesquisas no campo da educação (EMER, 2011; MICHELS; GARCIA, 2010) mostram que a matrícula desses alunos na escola não garante, necessariamente, a realização do atendimento adequado (JORDÃO, 2013; MELO, 2008) e que o atendimento educacional especializado nas salas de recursos multifuncionais (que constitui o âmago da política atual de atendimento educacional especializado) também não tem respondido a todas as características e necessidades acadêmicas dos alunos (ALBUQUERQUE, 2008; ARARUNA, 2013; ARNAL, 2007).

Os dados estatísticos propiciam diferentes leituras. Se, por um lado, o término de um programa, a sua descontinuidade ou o fechamento de um órgão significam qualidades para a gestão pública gerencial, por outro, afetam as condições de vida de grande parcela da população, que depende exclusivamente das políticas públicas. Aqui é preciso lembrar o desafio expresso por Freitas (2013) sobre o papel de análise independente dos pesquisadores em relação ao material que circula como verdade. A academia, em posse de indicadores sociais e de um conjunto bastante relevante de pesquisas nas mais diferentes áreas (em política educacional, sobre o cotidiano de sala de aula, em didática colaborativa, entre outras), muitas das quais expostas neste artigo, pode compreender tais números com um olhar que transcenda uma perspectiva parcial/superficial, avaliando-os no contexto histórico e político em que são construídos.

REFERÊNCIAS

ALBUQUERQUE, Rosana Aparecida. *Educação e inclusão escolar: a prática pedagógica na sala de recursos de 5ª a 8ª séries*. 2008. 144 f. Dissertação (Mestrado em Educação) – Universidade Estadual de Maringá, Maringá, 2008.

ALMEIDA, Ivan Castro de; WOLYNEC, Elisa. A produção de indicadores educacionais no Brasil e a comparação internacional. *Revista Brasileira de Estudos Pedagógicos*, Brasília, DF, v. 79, n. 193, p. 148-160, set./dez. 1998.

ANACHE, Alexandra Ayach. *Diagnóstico ou inquirição: estudo sobre o uso do diagnóstico psicológico na escola*. 1997. Tese (Doutorado em Psicologia) – Universidade de São Paulo, São Paulo, 1997.

ARARUNA, Maria Rejane. *Dos desafios às possibilidades: a prática pedagógica do professor do atendimento educacional especializado com o aluno que apresenta deficiência intelectual*. 2013. 281 f. Dissertação (Mestrado em Educação Brasileira) – Universidade Federal do Ceará, Fortaleza, 2013.

ARNAL, Leila de Souza Peres. *Educação escolar inclusiva: a prática pedagógica nas salas de recursos*. 2007. 133 f. Dissertação (Mestrado em Educação) – Universidade Estadual de Maringá, Maringá, 2007.

BONAMINO, Alicia Maria Catalano; OLIVEIRA, Lúcia Helena Gazólis de. Estudos longitudinais e pesquisa na educação básica. *Linhas Críticas*, Brasília, DF, v. 19, n. 38, p. 33-50, jan./abr. 2013.

BRASIL. *Decreto n. 42.728, de 3 de dezembro de 1957*. Institui a Campanha para a Educação do Surdo Brasileiro. Brasília, DF: Presidência da República, 1957.

BRASIL. *Decreto n. 44.236, de 1 de agosto de 1958*. Institui a Campanha Nacional de Educação e Reabilitação dos Deficitários Visuais. Brasília, DF: Presidência da República, 1958.

BRASIL. *Decreto n. 48.961, de 22 de setembro de 1960*. Institui a Campanha Nacional de Educação e Reabilitação de Deficientes Mentais. Brasília, DF: Presidência da República, 1960.

BRASIL. *Lei n. 5.692, de 11 de agosto de 1971*. Fixa diretrizes e bases para o ensino de 1º e 2º graus, e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 1971.

BRASIL. Ministério da Educação e Cultura. *Portaria n. 215, de 20 de agosto de 1973*. Brasília, DF: MEC, 1973a.

BRASIL. Ministério da Educação e Cultura. *Portaria n. 233, de 4 de setembro de 1973*. Brasília, DF: MEC, 1973b.

BRASIL. Presidência da República. *Decreto n. 72.425, de 3 de julho de 1973*. Cria o Centro Nacional de Educação Especial (Cenesp), e dá outras providências. Brasília, DF, 1973c.

BRASIL. Ministério da Educação e Cultura. Centro Nacional de Educação Especial. Serviço de Estatística da Educação e Cultura. *Educação Especial: dados estatísticos – 1974*. Brasília, DF: SEEC/DDD, 1975. v. 1.

BRASIL. Ministério da Educação e Cultura. Serviço de Estatística da Educação e Cultura. Centro Nacional de Educação Especial. *Educação Especial no Brasil: síntese estatística*. Brasília, DF: SEEC, 1984.

BRASIL. Ministério da Educação. Serviço de Estatística da Educação e Cultura. Centro Nacional de Educação Especial. *Sinopse estatística da educação especial 1981*. Brasília, DF: SEEC, 1985. v. 3.

BRASIL. *Decreto n. 93.613, de 21 de novembro de 1986*. Extingue órgãos do Ministério da Educação, e dá outras providências. Brasília, DF: Presidência da República, 1986.

BRASIL. *Constituição da República Federativa do Brasil (1988)*. Brasília, DF: Senado Federal, 1988.

BRASIL. Ministério da Educação. Coordenação de Informações e Planejamento. *Sinopse Estatística da Educação Especial 1988: ensino regular*. Brasília, DF: MEC/CIP, 1991a. v. 1.

BRASIL. Ministério da Educação. Coordenação de Informações e Planejamento. *Sinopse Estatística da Educação Especial 1988: instituições especializadas*. Brasília, DF: MEC/CIP, 1991b. v. 2.

BRASIL. Presidência da República. Câmara da Reforma do Estado. *Plano Diretor da Reforma do Aparelho do Estado*. Brasília, DF, 1995.

BRASIL. Ministério da Educação. *Lei n. 9.394, de 20 de dezembro de 1996*. Estabelece as diretrizes e bases da educação nacional. Brasília, DF, 1996.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística, 1996: Brasil, regiões e unidades da Federação*. Brasília, DF: Inep, 1997.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação especial MEC/Inep – 1998*. Brasília, DF: MEC/Inep, 1998a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Informe estatístico da educação básica: evolução recente das estatísticas da educação básica no Brasil – 1998*. Brasília, DF: Inep, 1998b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação básica MEC/Inep – 1998*. Brasília, DF: MEC/Inep, 1998c.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação básica MEC/Inep – 2000*. Brasília, DF: MEC/Inep, 2000.

BRASIL. Conselho Nacional de Educação. Câmara de Educação Básica. *Resolução n. 2, de 11 de setembro de 2001*. Diretrizes nacionais para educação especial na educação básica. Brasília, DF, 2001.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação básica MEC/Inep – 2002*. Brasília, DF: MEC/Inep, 2002.

BRASIL. Ministério da Educação. *Portaria Normativa n. 13, de 24 de abril de 2007*. Dispõe sobre a criação do “Programa de Implantação de Salas de Recursos Multifuncionais”. Brasília, DF, 2007a.

BRASIL. Presidência da República. *Lei n. 11.494/2007, de 20 de junho de 2007*. Regulamenta o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação – Fundeb – e dá outras providências. Brasília, DF, 2007b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2007*. Brasília, DF: MEC/Inep, 2007c.

BRASIL. Ministério da Educação. Secretaria de Educação Especial. *Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva*. Brasília, DF: SEESP, 2008a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2008: cadernos de instruções*. Brasília, DF: DEED/CGCE, 2008b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2008*. Brasília, DF: MEC/Inep, 2008c.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2009: cadernos de instruções*. Brasília, DF: DEED/CGCE, 2009a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2009*. Brasília, DF: MEC/Inep, 2009b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação especial MEC/Inep – 2010*. Brasília, DF: MEC/Inep, 2010a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2010: cadernos de instruções*. Brasília, DF: DEED/CGCE, 2010b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2010*. Brasília, DF: MEC/Inep, 2010c.

BRASIL. Presidência da República. *Decreto n. 7.480, de 16 de maio de 2011*. Aprova a Estrutura Regimental e o Quadro Demonstrativo dos Cargos em Comissão do Grupo Direção e Assessoramento Superiores – DAS – e das Funções Gratificadas do Ministério da Educação e dispõe sobre remanejamento de cargos em comissão. Brasília, DF, 2011a.

BRASIL. Presidência da República. *Decreto n. 7.611, de 17 de novembro de 2011*. Dispõe sobre a educação especial, o atendimento educacional especializado, e dá outras providências. Brasília, DF, 2011b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2011: cadernos de instruções*. Brasília, DF: DEED/CGCE, 2011c.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2011*. Brasília, DF: MEC/Inep, 2011d.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2012: cadernos de instruções*. Brasília, DF: DEED/CGCE, 2012a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2012*. Brasília, DF: MEC/Inep, 2012b.

BRASIL. Ministério da Educação. Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão. *Documento orientador do Programa Implantação de Salas de Recursos Multifuncionais*. Brasília, DF: Secadi, 2012c.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2013: caderno de instruções*. Brasília, DF: DEED/CGCE, 2013a.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2013*. Brasília, DF: MEC/Inep, 2013b.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo escolar da educação básica 2013: resumo técnico*. Brasília, DF: Inep, 2014a.

BRASIL. Ministério da Educação. *Relatório Educação para Todos 2000-2015: versão preliminar*. jun. 2014. Brasília, DF: MEC, 2014b.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do censo escolar da educação básica 2014*. Brasília, DF: MEC/Inep, 2014c.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Plano Nacional de Educação PNE 2014-2024: linha de base*. Brasília, DF: Inep, 2015.

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

BUENO, José Geraldo Silveira. *Educação especial brasileira: a integração/segregação do aluno diferente*. 1991. 214 f. Tese (Doutorado em Educação) – Pontifícia Universidade Católica de São Paulo, São Paulo, 1991.

CARVALHO, Maria de Fátima. O aluno com deficiência intelectual na escola: ensino, aprendizagem e desenvolvimento humano. In: MELETTI, Sílvia Márcia Ferreira; KASSAR, Mônica de Carvalho Magalhães (Org.). *Escolarização de alunos com deficiências: desafios e possibilidades*. Campinas, SP: Mercado de Letras, 2013. p. 203-241.

- DIAS, Patrícia Jovino de Oliveira. *O registro de dados sobre alunos com necessidades educacionais especiais no censo escolar da educação básica na rede municipal de ensino de Londrina/PR*. 2014. 134 f. Dissertação (Mestrado em Educação) – Universidade Estadual de Londrina, Londrina, 2014.
- DINIZ, Ednar Maria Vieira. O censo escolar. *Revista Brasileira de Estudos Pedagógicos*, Brasília, DF, v. 80, n. 194, p. 156-163, jan./abr. 1999.
- EMER, Simone de Oliveira. *Inclusão escolar: formação docente para o uso das TICs aplicada como tecnologia assistiva na sala de recurso multifuncional e a sala de aula*. 2011. 163 f. Dissertação (Mestrado em Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.
- FERRARO, Alceu Ravanello. Utilização das estatísticas educacionais dos censos demográficos e dos registros escolares: uma tipologia de análises. *Educação e Realidade*, Porto Alegre, v. 4, n. 2, p. 253-266, 1979.
- FERRARO, Alceu Ravanello. Quantidade e qualidade na pesquisa em educação, na perspectiva da dialética marxista. *Pro-Posições*, Campinas, v. 23, n. 1, p. 129-146, jan./abr. 2012.
- FERREIRA, Júlio Romero. *A exclusão da diferença*. Piracicaba, SP: Ed. Unimep, 1993.
- FRANCO, Maria Laura P. Barbosa. Pesquisa educacional: algumas reflexões. *Cadernos de Pesquisa*, São Paulo, n. 51, p. 84-87, nov. 1984.
- FRANCO, Maria Laura P. Barbosa; GOLDBERG, Maria Amélia Azevedo. Prioridades em pesquisa educacional: prós e contras. *Cadernos de Pesquisa*, São Paulo, n. 16, p. 74-80, nov. 1976.
- FREITAS, Luiz Carlos de. Políticas de responsabilização: entre a falta de evidência e a ética. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 148, p. 348-365, jan./abr. 2013.
- GAMBOA, Silvio Ancízar Sanchez. A dialética na pesquisa em educação: elementos de contexto. In: FAZENDA, Ivani (Org.). *Metodologia da pesquisa educacional*. 3. ed. São Paulo: Cortez, 1994. p. 91-115.
- GATTI, Bernardete Angelina (Org.). Alternativas metodológicas para a pesquisa educacional: conhecimento e realidade. *Cadernos de Pesquisa*, São Paulo, v. 40, p. 4-6, fev. 1982.
- GATTI, Bernardete Angelina. *A construção da pesquisa em educação no Brasil*. Brasília, DF: Plano, 2002.
- GOERGEN, Pedro. A pesquisa educacional no Brasil: dificuldades, avanços e perspectivas. *Em Aberto*, Brasília, DF, ano 5, n. 31, p. 1-17, jul./set. 1985.
- GOUVEIA, Aparecida Joly. A pesquisa educacional no Brasil: de 1970 para cá. *Cadernos de Pesquisa*, São Paulo, n. 19, p. 75-79, dez. 1976.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo Demográfico de 2010*. Rio de Janeiro: IBGE, 2010. Disponível em: <<https://www.ibge.gov.br/>>. Acesso em: 07 jun. 2017.

INTERNATIONAL BUSINESS MACHINES. *Software IBM (International Business Machines) SPSS Statistics*. Versão 17.

JANNUZZI, Gilberta Sampaio de Martino. *A luta pela educação do deficiente mental no Brasil*. São Paulo: Cortez; Autores Associados, 1985.

JANNUZZI, Paulo de Martino. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil. *Revista do Serviço Público*, Brasília, DF, v. 56, n. 2, p. 137-159, 2005.

JANNUZZI, Paulo de Martino. *Indicadores sociais no Brasil: conceitos, fontes de dados e aplicações*. 4. ed. Campinas, SP: Alínea, 2009.

JANNUZZI, Paulo de Martino. Avaliação de programas sociais: conceitos e referências de quem a realiza. *Estudos em Avaliação Educacional*, São Paulo, v. 25, n. 58, p. 22-42, maio/ago. 2014a.

JANNUZZI, Paulo de Martino. A produção de informação e conhecimento para aprimoramento das políticas e programas de desenvolvimento social: princípios, conceitos e caracterização das pesquisas de avaliação realizadas pela SAGI/MDS de 2011 a 2014. In: JANNUZZI, Paulo de Martino; QUIROGA, Júnia (Org.). *Síntese das pesquisas de avaliação de programas sociais do MDS 2011-2014*. Brasília, DF: Ministério do Desenvolvimento Social e Combate à Fome, SAGI, 2014b. p. 12-35. (Cadernos de Estudos: Desenvolvimento Social em Debate, n. 16).

JANNUZZI, Paulo de Martino et al. Estruturação de sistemas de monitoramento e especificação de pesquisas de avaliação: os problemas dos programas públicos são. In: ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA (Org.). *Reflexões para a Ibero-América: avaliação de programas sociais*. 1. ed. Brasília, DF: Enap, 2009. p. 101-138. (Caderno EIAPP).

JORDÃO, Suelen Garay Figueiredo. *Política de inclusão escolar na ótica de professores das salas de recursos multifuncionais*. 2013. 132 f. Dissertação (Mestrado em Educação) – Universidade do Vale do Itajaí, Itajaí, 2013.

KASSAR, Mônica de Carvalho Magalhães. Diagnosticar a deficiência: sim ou não? *Revista Brasileira de Educação Especial*, Marília, v. 1, n. 2, p. 85-92, 1994.

KASSAR, Mônica de Carvalho Magalhães; MELETTI, Silvia Márcia Ferreira. Análise de possíveis impactos do Programa Educação Inclusiva: direito à diversidade. *Ciências Humanas e Sociais em Revista*, Rio de Janeiro, v. 34, n. 12, p. 49-63, jan./jun. 2012.

LAPLANE, Adriana Lia Frizman de. O que os dados do censo escolar revelam sobre as barreiras à inclusão escolar? *Educação e Fronteiras*, Dourados, v. 5, n. 13, p. 21-41, maio/ago. 2015.

LEIJOTO, Camila Pereira. *Registro do público da educação especial no censo escolar: deficiência intelectual em destaque*. 2015. 71 f. Dissertação (Mestrado em Educação) – Universidade Federal de Mato Grosso do Sul, Corumbá, 2015.

LOURENÇO FILHO, Manuel Bergström. Estatística e educação. *Revista Brasileira de Estudos Pedagógicos*, Rio de Janeiro, v. 11, n. 31, p. 60-73, nov./dez. 1947.

LOURENÇO FILHO, Manuel Bergström. Antecedentes e primeiros tempos do Inep. *Revista Brasileira de Estudos Pedagógicos*, Rio de Janeiro, v. 42, n. 95, p. 8-17, jul./set. 1964.

MARTINS, Susana da Cruz; MAURITTI, Rosário; COSTA, António Firmino da. Acesso a bases de microdados: aplicações e impactos na pesquisa em ciências sociais. *Mediações*, Londrina, PR, v. 18, n. 1, p. 66-82, jan./jun. 2013.

MELETTI, Sílvia 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: Educação no Brasil: o balanço de uma década, 33., 2010, Caxambu. *Anais... Caxambu: ANPEd*, 2010. v. 1, p. 1-17.

MELETTI, Sílvia Márcia Ferreira; BUENO, José Geraldo Silveira (Org.). *Políticas públicas, escolarização de alunos com deficiência e a pesquisa educacional*. 1. ed. Araraquara, SP: Junqueira & Marin, 2013.

MELETTI, Sílvia Márcia Ferreira; RIBEIRO, Karen. Indicadores educacionais sobre a educação especial no Brasil. *Cadernos CEDES*, Campinas, SP, v. 34, n. 93, p. 175-189, maio 2014.

MELO, Hilce Aguiar. *O acesso curricular para alunos com deficiência intelectual na rede regular de ensino: a prática pedagógica na sala de recursos como eixo para análise*. 2008. 171 f. Dissertação (Mestrado em Educação) – Universidade Federal do Maranhão, São Luís, 2008.

MENDES, Enicéia. *Deficiência mental: a construção científica de um conceito e a realidade educacional*. 1995. Tese (Doutorado em Psicologia) – Universidade de São Paulo, São Paulo, 1995.

MICHELS, Maria Helena; GARCIA, Rosalba Maria Cardoso. A organização curricular na articulação entre serviço especializado e classe comum: um modelo inclusivo? In: KASSAR, Mônica de Carvalho Magalhães (Org.). *Diálogos com a diversidade: desafios da formação de educadores na contemporaneidade*. Campinas, SP: Mercado das Letras, 2010. p. 211-231.

NÚCLEO DE ESTUDOS DE POLÍTICAS PÚBLICAS. *Brasil 1986: relatório sobre a situação social do país*. Campinas, SP: Unicamp/NEPP, 1988.

PASCHOALICK, Wanda. *Análise do processo de encaminhamento de crianças das classes especiais para deficientes mentais, desenvolvido nas escolas de 1º grau da delegacia de ensino de Marília*. 1981. Dissertação (Mestrado em Educação) – Pontifícia Universidade Católica de São Paulo, São Paulo, 1981.

PEREIRA, Carlos Alberto Bragança; MAZZON, José Afonso; ROSENBERG, Lia. Censo Escolar: controle de qualidade. *Revista Brasileira de Estudos Pedagógicos*, Brasília, DF, v. 79, n. 191, p. 74-82, jan./abr. 1998.

PLETSCH, Márcia Denise. A escolarização do aluno com deficiência intelectual... apesar do diagnóstico. In: MELETTI, Sílvia Márcia Ferreira; KASSAR, Mônica de Carvalho Magalhães (Org.). *Escolarização de alunos com deficiências: desafios e possibilidades*. Campinas, SP: Mercado de Letras, 2013.

POPKEWITZ, Tom; LINDBLAD, Sverker. Estatísticas educacionais como um sistema de razão: relações entre governo da educação e inclusão e exclusão sociais. *Educação & Sociedade*, Campinas, SP, v. 22, n. 75, p. 111-148, ago. 2001.

PRIETO, Rosângela Gavioli. Políticas de inclusão escolar no Brasil: sobre novos/velhos significados para a educação especial. In: MENDES, Enicéia; ALMEIDA, Maria Amélia (Org.). *Das margens ao centro: perspectivas para as políticas e práticas educacionais no contexto da educação especial inclusiva*. Araraquara, SP: Junqueira & Marin, 2010. p. 61-78.

REBELO, Andressa Santos. *Os impactos da política de atendimento educacional especializado: análise dos indicadores educacionais de matrículas de alunos com deficiência*. 2012. Dissertação (Mestrado em Educação) – Universidade Federal de Mato Grosso do Sul, Corumbá, 2012.

ROSEMBERG, Fúlvia. A avaliação de programas, indicadores e projetos em educação infantil. *Revista Brasileira de Educação*, Rio de Janeiro, n. 16, p. 19-26, jan./abr. 2001.

ROSEMBERG, Fúlvia. Políticas de educação infantil e avaliação. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 148, p. 44-75, jan./abr. 2013.

SANTOS FILHO, José Camilo; GAMBOA, Silvio Sánchez (Org.). *Pesquisa educacional: quantidade-qualidade*. 7. ed. São Paulo: Cortez, 2009.

SASS, Odair. Sobre os conceitos de censo e amostragem em educação no Brasil. *Estatística e Sociedade*, Porto Alegre, v. 2, p. 128-141, nov. 2012.

SASS, Odair. Estatística e estatísticas escolares no Brasil: a emergência dos serviços públicos de Estatística, sob a óptica de Rui Barbosa. In: MELETTI, Sílvia Márcia Ferreira; BUENO, José Geraldo Silveira (Org.). *Políticas públicas, escolarização de alunos com deficiência e a pesquisa educacional*. 1. ed. Araraquara, SP: Junqueira & Marin, 2013. p. 39-52.

SCHRADER, Achim. *Métodos de pesquisa social empírica e indicadores sociais*. Porto Alegre: Editora UFRGS, 2002.

SHIROMA, Eneida Oto; EVANGELISTA, Olinda. Avaliação e responsabilização pelos resultados: atualizações nas formas de gestão de professores. *Perspectiva*, Florianópolis, v. 29, n. 1, p. 127-160, out. 2011.

SOUZA, Celina. Estado da arte da pesquisa em políticas públicas. In: HOCHMAN, Gilberto; ARRETCHE, Marta; MARQUES, Eduardo (Org.). *Políticas públicas no Brasil*. Rio de Janeiro: Fiocruz, 2007. p. 65-86.

VIDAL, Eloisa Maia; VIEIRA, Sofia Lerche. Gestão educacional e resultados no Ideb: um estudo de caso em dez municípios cearenses. *Estudos em Avaliação Educacional*, São Paulo, v. 22, n. 50, p. 419-434, set./dez. 2011.

ANDRESSA SANTOS REBELO

Professora Adjunta da Universidade Federal de Mato Grosso do Sul (UFMS), Corumbá, Mato Grosso do Sul

andressarbl@gmail.com

<http://orcid.org/0000-0003-1873-5622>

MÔNICA DE CARVALHO MAGALHÃES KASSAR

Professora Associada da Universidade Federal de Mato Grosso do Sul (UFMS), Corumbá, Mato Grosso do Sul

monica.kassar@gmail.com

Recebido em: 15 JUNHO 2016

Aprovado para publicação em: 12 JUNHO 2017

INSTRUÇÕES A COLABORADORES

Estudos em Avaliação Educacional é um periódico quadrimestral que publica trabalhos inéditos, em versão *on-line* e impressa, relacionados à temática de avaliação educacional, apresentados sob a forma de artigos, relatos de pesquisa, ensaios teórico-metodológicos e resenhas, que devem ser submetidos exclusivamente ao periódico. Excepcionalmente, serão aceitos trabalhos de revistas estrangeiras (com a indicação da fonte), os quais terão de passar pelo mesmo processo de avaliação dos artigos inéditos. O autor deverá apresentar a autorização da revista em que seu artigo tenha sido originalmente publicado.

São publicados artigos em português e em espanhol. Trabalhos submetidos nos demais idiomas são traduzidos para o português.

Avaliação dos originais

Os originais recebidos são previamente avaliados pelo Comitê Editorial; aqueles que não se enquadram no escopo da revista são devolvidos aos autores; os que estão adequados aos critérios do periódico são encaminhados a dois especialistas da área – avaliadores externos, membros do Conselho Editorial ou pesquisadores da Fundação Carlos Chagas – com título de doutor e filiados a instituições distintas das quais se vinculam os autores. Os especialistas realizam a avaliação de acordo com os seguintes critérios: relevância do tema para a área de estudo; consistência teórico-metodológica; pertinência e atualidade da bibliografia; clareza na exposição do objeto de estudo e na argumentação; rigor na utilização dos conceitos; pertinência, clareza e fundamentação da metodologia estatística empregada. Nesse processo de avaliação, os nomes dos avaliadores e dos autores permanecem em sigilo. No último número de cada volume, é publicado o nome de todos os pareceristas que colaboraram com a revista ao longo do ano.

Os autores recebem comunicação relativa aos pareceres emitidos, informando se seu texto foi aceito, rejeitado ou se necessita de correções para publicação. O Comitê Editorial reserva-se o direito de recusar o artigo ao qual foram solicitadas correções, caso estas não sejam atendidas a contento.

Publicação dos originais

Se a matéria for aceita para publicação, a revista permite-se introduzir pequenas alterações formais no texto, respeitando o estilo e a opinião dos autores. Os trabalhos não poderão ser publicados em qualquer outra forma antes de decorridos seis meses de sua publicação em *Estudos em Avaliação Educacional*. Artigos de um mesmo autor só serão publicados com intervalo de, pelo menos, seis meses.

Autoria

Solicitam-se dos autores: nome completo (sem abreviações); vínculo institucional ou última ocupação profissional; cidade, estado e país da instituição; titulação

(graduado, especialista, mestre/mestrando, doutor/doutorando), endereço, telefone, celular e *e-mail* de contato. Pede-se, ainda, que o autor indique como seu nome deve constar da publicação e qual *e-mail* deve ser apresentado. Os autores de textos publicados receberão 1 (um) exemplar do número em que suas matérias forem veiculadas.

Os direitos autorais dos trabalhos publicados pertencem à revista *Estudos em Avaliação Educacional*, que adota a licença Creative Commons BY-NC do tipo "Atribuição Não Comercial". A versão *on-line* é de acesso aberto e gratuito. Os pontos de vista expressos nos textos são de responsabilidade dos autores.

Estudos em Avaliação Educacional não cobra taxa de submissão ou de editoração de artigos (articles processing charges – APC).

Para publicação do material em outros veículos, entrar em contato com: eae@fcc.org.br.

Apresentação dos originais

Para submeter um artigo à *Estudos em Avaliação Educacional*, é necessário fazer o cadastro como autor no Portal de Periódicos da Fundação Carlos Chagas: <<http://publicacoes.fcc.org.br/ojs/>>, criando *login* e senha de acesso. O autor deve acessar o sistema, clicar no link "Nova submissão" e preencher as informações necessárias nos cinco passos de submissão, a saber:

Passo 1. Iniciar submissão: preenchimento dos campos: Seção; Idioma da submissão; Condições para submissão (todos os itens deste campo devem ser selecionados); Declaração de direito autoral; e Comentários para o editor (caso houver necessidade).

Passo 2. Transferência do manuscrito: envio do artigo, em PDF, para avaliação. Dessa versão do artigo não devem constar quaisquer referências que permitam identificar a autoria do trabalho. Assim, é necessário excluir do artigo, além do nome dos autores, nomes das instituições e de projetos, ou mesmo alguma menção em fonte de tabela/gráfico/quadro/figura e em referência bibliográfica.

Passo 3. Metadados da Submissão (Indexação): inclusão dos dados dos autores, do título e do resumo do trabalho apresentado, de contribuidores e agências de fomento (se for o caso) e das referências do texto.

Passo 4. Transferência de documentos suplementares: envio do artigo, em Microsoft Word, com todas as informações de autoria. Além disso, nesta etapa da submissão, devem ser enviados tabelas, gráficos, quadros e figuras, em arquivos separados editáveis.

Passo 5. Confirmação da submissão: nesta etapa, o autor deve clicar em "Concluir submissão".

Para acompanhar o *status* da submissão, o autor deve acessar o sistema → menu → acesso → *login* e senha. Indica-se que todos os autores estejam cadastrados no sistema.

Todos os documentos enviados à *Estudos em Avaliação Educacional* devem ter o seguinte formato obrigatório: 3 cm de margem superior, 3 cm de margem inferior, 3 cm de margem esquerda e 2 cm de margem direita; parágrafo 1,25; com espaçamento de 1,5 entre as linhas; sem espaço (anterior ou posterior) entre os parágrafos, páginas enumeradas (após a folha de rosto, na margem inferior à direita), fonte em Times New Roman, no corpo 12.

Artigos, relatos de pesquisa e ensaios teórico-metodológicos: não podem ultrapassar 25 páginas (incluídos os anexos).

Resenhas: devem apresentar extensão máxima de 6 páginas, contendo referência da obra resenhada, título e análise crítica que traga contribuições ao campo de estudo do trabalho em questão.

• **Títulos e subtítulos** devem ter, no máximo, 11 palavras (incluindo artigos, preposições, conjunções, etc.). Toda matéria, à exceção de resenhas, precisa vir acompanhada de resumo em português contendo no máximo 11 linhas, com o objetivo, o método, os resultados e as conclusões do artigo. O espaçamento deve ser simples entre as linhas e sem espaço entre parágrafos, sem conter siglas nem referências, trazendo, em seu início, o título do trabalho. Ao final do resumo, indicar quatro palavras-chave (descritores) do conteúdo do texto, utilizando como referência o Thesaurus do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) (<http://inep.gov.br/thesaurus-brasileiro-da-educacao>).

• **Citações, remissões, notas e siglas** devem obedecer às regras da ABNT (NBR 10520, 2002). As **citações diretas** (textuais), com até três linhas, devem ser incorporadas ao texto, entre aspas, sendo necessário indicar o sobrenome do autor, ano e número da página. Ex.: Em função desses indicadores, “chegou-se à organização de cinco grandes grupos de escolas denominadas azul, verde, amarelo, laranja e vermelho” (SÃO PAULO, 2001, p. 55).

• **Citações com mais de três linhas** deverão ser apresentadas sem aspas, com recuo de 1,25 cm a partir da margem esquerda, com espaçamento simples entre as linhas e sem espaço entre parágrafos, em fonte Times New Roman e corpo 10, sendo necessário indicar o sobrenome do autor, ano e página. Ex.:

[...] a sofisticação técnica da avaliação nacional - que hoje ocupa a atenção da cúpula decisória e de seus assessores - apresenta-se como entrave para a compreensão; tanto pelos atores dos sistemas e escolas como pela população em geral, do processo avaliativo realizado. (FREITAS, 2004, p. 685)

• Na **citação de citação** deve ser empregada a expressão latina “apud” (citado por) para identificar a fonte que foi efetivamente consultada, a qual deve ter a referência completa no rodapé; e na lista de Referências incluir apenas a obra consultada (CARONE et al., 2003). Ex.: Para Watson (apud CARONE et al., 2003) [...].

• As **remissões bibliográficas indiretas** são incorporadas ao texto entre parênteses (ano). Ex.: Segundo João Barroso (2006), todos...

• As **notas explicativas** devem ser evitadas e utilizadas apenas quando for estritamente necessário, preferencialmente sem ultrapassar três linhas. Devem figurar sempre no rodapé da página, numeradas sequencialmente.

• As **siglas** devem ser desdobradas quando mencionadas à primeira vez no artigo. Ex.: Exame Nacional do Ensino Médio (Enem).

• A **metodologia estatística** empregada deve ser descrita com detalhamento que permita que os resultados possam ser replicados por um leitor com experiência no assunto que tenha acesso aos dados originais. Por outro lado, deve-se evitar o uso de linguagem excessivamente técnica, visando a apresentar a descrição metodológica com clareza suficiente para que um leitor não especializado no assunto também possa compreendê-la. Essas recomendações se traduzem, principalmente, em medidas como:

- ✓ fornecer detalhes sobre as unidades experimentais, aleatorização e procedimentos amostrais;
- ✓ descrever e/ou proporcionar acesso aos instrumentos de coleta da pesquisa;
- ✓ indicar todas as técnicas descritivas e inferenciais utilizadas, fornecendo referências bibliográficas sempre que necessário;
- ✓ apresentar os resultados quantitativos com suas respectivas medidas de incerteza (intervalos de confiança, *p-values*, etc.);
- ✓ adotar um nível de precisão na apresentação dos dados e resultados (por exemplo, número de casas decimais) padronizada e coerente com as possibilidades da pesquisa;
- ✓ utilizar recursos gráficos e tabelas sempre que esses elementos puderem facilitar a organização e síntese dos dados e a compreensão dos resultados;
- ✓ discutir a razoabilidade dos resultados obtidos e relatar possíveis limitações dos métodos empregados;
- ✓ definir os termos técnicos, os símbolos matemáticos e as siglas presentes no artigo;
- ✓ especificar os *softwares* estatísticos utilizados.

• **Tabelas, gráficos, quadros e figuras** (assim como os **títulos** e as **fontes**) devem ser apresentados no corpo do texto, e não em caixas de texto, alinhados à esquerda, em sua página correspondente, numerados com algarismos arábicos, com títulos (posicionados acima, em corpo 12) padronizados quanto ao formato e termos utilizados. Abaixo destes, sem estarem em caixas de texto, deve, obrigatoriamente, ser indicada a fonte dos dados (remetida às referências bibliográficas), com autoria e ano, inclusive se for de elaboração própria dos autores, em corpo 10, alinhada à esquerda, espaço 1,5 entre linhas. As tabelas, os gráficos e os quadros devem se apresentar em *software* compatível com o ambiente Windows, de preferência em Excel, e as figuras (ilustrações, imagens, mapas, fotos, etc.), em arquivo com alta resolução (300 dpi), todos produzidos em preto e branco, em tamanho máximo de 10 cm de largura.

• **Referências de cunho bibliográfico** devem vir ao final do texto, por ordem alfabética de sobrenome do autor, e, sempre que possível, fazer constar por extenso o prenome dos autores. Os títulos das obras devem ser apresentados em itálico. Quando houver dois ou três autores, separa-se o primeiro autor e os demais por ponto e vírgula; ultrapassando três autores, faz-se a entrada pelo autor principal (referenciado no texto) e substituem-se os outros pela expressão "et al.". Sua apresentação deve seguir as normas da ABNT (NBR 6023, 2002). A exatidão das referências e a correta citação no texto são de responsabilidade do(s) autor(es) dos textos, sendo uma exigência para a publicação do trabalho.

Seguem alguns exemplos de referência:

Livros com dois ou três autores

GUINCHAT, Claire; MENU, Michel. *Introdução geral às ciências e técnicas da informação e documentação*. 2. ed. corrig. e aum. Brasília: IBICT, 1994.

Capítulo de livro - mesmo autor

GADOTTI, Moacir. A paixão de conhecer o mundo. In: _____. *Pensamento pedagógico brasileiro*. São Paulo: Atlas, 1987. p. 58-73. cap. 5.

Capítulo de livro - autor diferente

RONCA, Antônio Carlos C. O modelo de ensino de David Ausubel. In: PENTEADO, Wilma Millan Alves. *Psicologia e ensino*. São Paulo: Papalivros, 1980.

Artigo de periódico

CALLADO, Antônio. Da grande mentira às primeiras sílabas da verdade. *Em Aberto*, Brasília, v. 26, n. 90, p. 169-174, jul./dez. 2013.

Artigo de periódico com mais de três autores

DEL BARRIO, Cristina et al. Representaciones acerca del maltrato entre iguales, atribuciones emocionales y percepción de estrategias de cambio a partir de un instrumento narrativo: SCAN-Bullying. *Infancia y Aprendizaje*, v. 26, n. 1, p. 63-78, 2003.

Textos on-line

LAEVERS, Ferre (Ed.). *Well-being and involvement in care settings*. A Process-oriented Self-evaluation Instrument (SiCs). Bruxelas: Kind & Gezin, 2008. Disponível em: <<http://www.kindengezin.be/img/sics-ziko-manual.pdf>>. Acesso em: 13 abr. 2015.

Artigo de jornal

NAVES, Paulo. Lagos andino dá banho de beleza. *Folha de São Paulo*, São Paulo, 28 jun. 1999. Folha Turismo, Caderno 8, p. 13.

Trabalho apresentado em evento

MOREIRA, A. F. B. Multiculturalismo, Currículo e Formação de Professores. In: SEMINÁRIO ESTADUAL DE EDUCAÇÃO BÁSICA, 2., 1998, Santa Cruz do Sul. *Anais...* Santa Cruz do Sul: EDUNISC, 1998. p. 15-30.

Autor institucional, legislação e jurisprudência

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*. Brasília, DF: Senado, 1988.

BRASIL. *Lei n. 9.394, de 20 de dezembro de 1996*. Estabelece as Diretrizes e Bases da Educação Nacional. Brasília, DF: MEC, 1996. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/19394.htm>. Acesso em: 15 mar. 2013.

BRASIL. *Estatuto da criança e do adolescente*. Organização dos textos, notas remissivas e índices por Juarez de Oliveira. 5. ed. São Paulo: Saraiva, 1995. 206 p.

BRASIL. Ministério da Educação. Conselho Nacional de Educação. *Resolução n.: 2, de 30 de janeiro de 2012*. Define Diretrizes Curriculares Nacionais para o Ensino Médio. Brasília, DF: MEC/CNE/CEB, 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Estudo exploratório sobre o professor brasileiro com base nos resultados do Censo Escolar da Educação Básica 2007*. Brasília, DF: Inep, 2009.

SÃO PAULO (Estado). Projeto de lei n. 15, de 2003. Dá denominação à Rodovia SP-046-125, que interliga São Luiz de Paraitinga-Catuçaba-Alto da Serra. *Diário Oficial [do] Estado de São Paulo*, Poder Legislativo, São Paulo, v. 113, n. 31, p. 2, fev. 2003.

SÃO PAULO (Estado). Decreto n. 42.822, de 20 de janeiro de 1998. *Lex*: coletânea de legislação e jurisprudência, São Paulo, v.62, n. 3, p. 217-220, 1998.

Verbetes de enciclopédia eletrônica e dicionário

FOULKES, H.; CARTWRIGHT, R. Sleep. In: _____. *Encyclopedia Britânica On-line*. Disponível em: <<http://www.britanica.com/bcom/eb/article>>. Acesso em: 5 fev. 2000.

Íncscio. In: FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário Aurélio da Língua Portuguesa*. 2.ed. Rio de Janeiro: Nova Fronteira, 1986. p. 950.

Entrevista/Depoimento

SILVA, A. A. Mulheres no ataque: depoimento. [9 de junho, 1996]. São Paulo: *Revista da Folha de São Paulo*. Entrevista concedida a Cristiana Couto.

Teses

FERREIRA, J. L. *Homens ensinando crianças: continuidade-descontinuidade das relações de gênero na escola rural*. 2008. 171 f. Tese (Doutorado em Educação) - Programa de Pós-Graduação em Educação, Universidade Federal da Paraíba, João Pessoa, 2008.

Resenha

CARONE, I. *Psicanálise fim de século. Ensaio crítico*. São Paulo: Hacker, 1998. Resenha de: FRAYZE-PEREIRA, J. A. Da possibilidade da crítica à cultura: psicanálise e filosofia. *Revista Brasileira de Psicanálise*, v. 35, n. 2, p. 403-405, 2001.

Trabalho publicado em CD/DVD

LIMA NETO, Newton. A Universidade e os trabalhadores. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA, 49., 1997, Belo Horizonte. *Anais...* Belo Horizonte: Videolar, 1991. CD-ROM.

DIREITO DE RESPOSTA

Estudos em Avaliação Educacional acolhe comentário(s) a artigo publicado na revista. Se o comentário for aceito para publicação, a revista oferecerá ao autor igual espaço para réplica, que poderá ser publicada no mesmo número do comentário ou no número subsequente. Ambos estão sujeitos ao mesmo processo de avaliação dos demais textos. Não são aceitos comentários ou réplicas a resenhas.

.....

ASSINE A REVISTA

ESTUDOS EM AVALIAÇÃO EDUCACIONAL

Tel. (11) 3723-3084

www.fcc.org.br

