

Temas em Avaliação
Educativa



PESQUISA
E
DEBATE
EM

REVISTA DO PROGRAMA DE
PÓS-GRADUAÇÃO PROFISSIONAL
EM GESTÃO E AVALIAÇÃO DA
EDUCAÇÃO PÚBLICA

VOL.10
N.1
2020

■ PESQUISA APLICADA

AVALIAÇÃO DA QUALIDADE NA EDUCAÇÃO INFANTIL: IMPASSES E PERSPECTIVAS NO BRASIL

Maria Malta Campos

AVALIAÇÃO DA APRENDIZAGEM NA EDUCAÇÃO INFANTIL: ESTUDO DE CASO EM UM MUNICÍPIO BRASILEIRO

Bruna Ribeiro

EDUCAÇÃO INFANTIL E AVALIAÇÃO: PROCESSOS E PERSPECTIVAS

Rita de Cássia de Freitas Coelho
Alice de Paiva Macário

A RESTITUIÇÃO COMO PROCESSO CENTRAL NA AVALIAÇÃO DE CONTEXTO

Catarina Moro

AVALIAÇÕES DA EDUCAÇÃO INFANTIL EM MUNICÍPIOS PAULISTAS: LIMITES E POTENCIALIDADES PARA CONTRIBUIR COM A GARANTIA DO DIREITO À EDUCAÇÃO DAS CRIANÇAS PEQUENAS

Cláudia Oliveira Pimenta

AVALIAÇÃO DE CONTEXTO NA EDUCAÇÃO INFANTIL: UMA ESCOLHA PARA PENSAR A MELHORIA DAS PRÁTICAS EDUCATIVAS

Patrícia Corsino
Jordanna Castelo Branco

INFRAESTRUTURA ESCOLAR E SATISFAÇÃO PROFISSIONAL: PERCEPÇÃO DE PROFESSORES DA EDUCAÇÃO BÁSICA BRASILEIRA

Lívia Maria Fraga Vieira
Edmilson Antônio Pereira Junior

OBRIGATORIEDADE DA PRÉ-ESCOLA EM PERSPECTIVA: AVALIAÇÃO DO ATENDIMENTO NO ESTADO DO RIO DE JANEIRO

Bruno Tovar Falciano
Maria Fernanda Rezende Nunes

AVALIAÇÕES EM LARGA ESCALA DE LÍNGUA PORTUGUESA: UMA PESQUISA SOBRE A COMPLEXIDADE DOS TEXTOS QUE DÃO SUPORTE A ITENS QUE AVALIAM A LEITURA

Begma Tavares Barbosa
Hilda Linhares Micarello
Rosângela Veiga Ferreira

A AVALIAÇÃO EXTERNA E OS NOVOS SUJEITOS DA EDUCAÇÃO

Marcelo Baumann Burgos

AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E ACCOUNTABILITY: UMA BREVE ANÁLISE DA EXPERIÊNCIA BRASILEIRA

Reynaldo Fernandes
Amaury Patrick Gremaud

AVALIAÇÃO EDUCACIONAL: FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR EM CENÁRIOS DE ALTA VULNERABILIDADE SOCIAL

Edivaldo Cesar Camarotti Martins
Adolfo-Ignacio Calderón

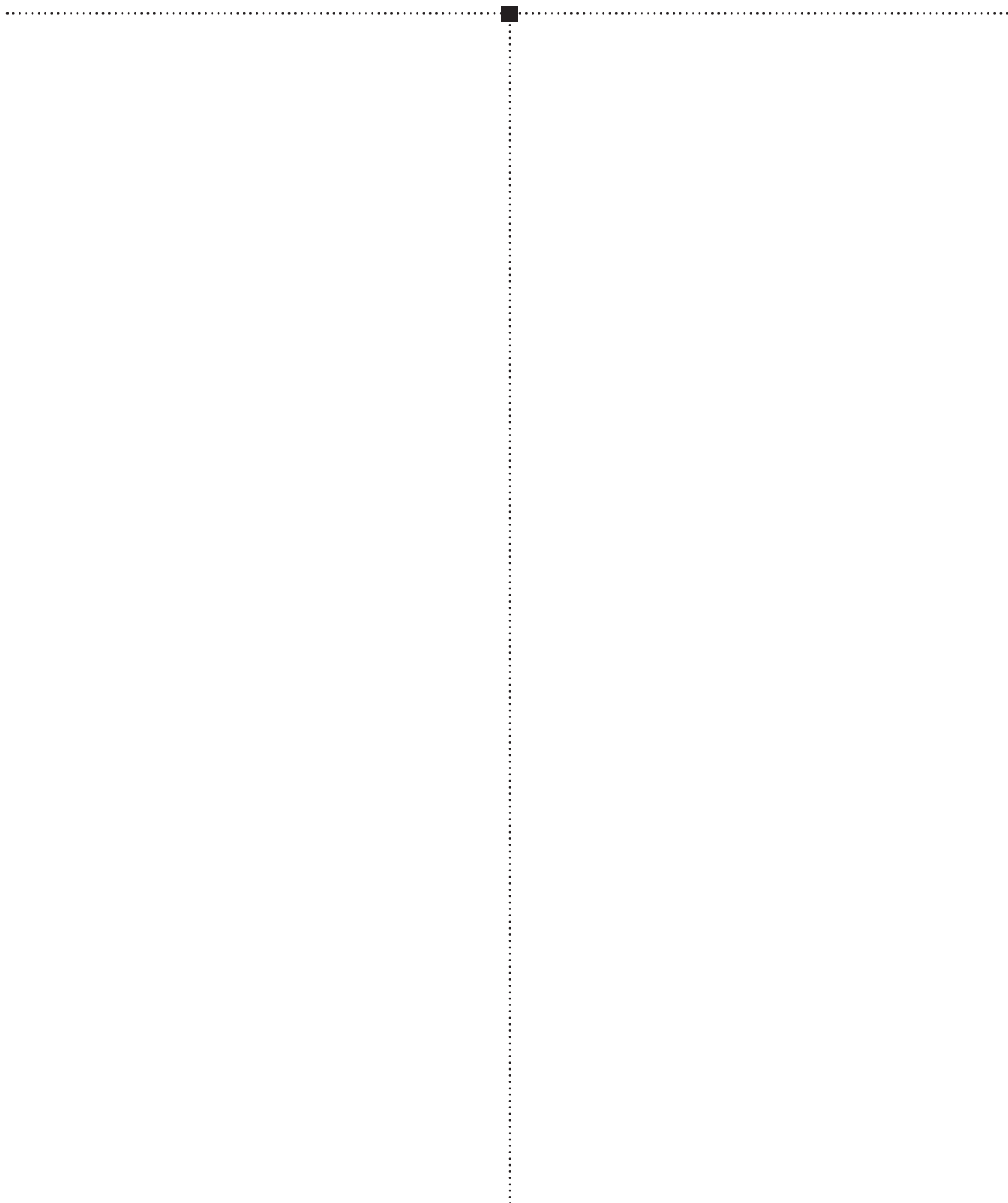
ESCALAS ESPECÍFICAS PARA MEDIR A INFRAESTRUTURA ESCOLAR DA EDUCAÇÃO BÁSICA

Joaquim José Soares Neto
Elianice Silva Castro

MENSURAÇÃO EM SISTEMAS ADMISIONAIS NO ENSINO SUPERIOR E AVALIAÇÕES DE IMPACTO

Tufi Machado Soares
Mariana Calife Nóbrega Soares

PESQUISA E DEBATE EM EDUCAÇÃO



Faculdade de Educação

Centro de Políticas Públicas e Avaliação da Educação
Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública
Avenida Eugênio do Nascimento, 620 – Dom Orione CEP: 36038-330 – Juiz de Fora – MG
Telefone: (32) 4009-9326
Site: <https://periodicos.ufjf.br/index.php/RPDE/>
E-mail: revista@caed.ufjf.br

Editor-Chefe

Frederico Braidá, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Editores Adjuntos

Liamara Scortegagna, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Marcelo Ferreira, Universidade de Brasília, Brasília, Distrito Federal, Brasil

Editor Convidado

Wagner Silveira Rezende, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Conselho Editorial

Alexandre Chibebe Nicoletta, Universidade de São Paulo, São Paulo, São Paulo, Brasil
Beatriz de Basto Teixeira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Bernadete A Gatti, Fundação Carlos Chagas, São Paulo, São Paulo, Brasil
Clarilza Prado de Souza, Universidade de São Paulo, São Paulo, São Paulo, Brasil
Fernando M Reimers, *Harvard Graduate School of Education, Cambridge, Massachusetts*, Estados Unidos da América
Lina Kátia Mesquita de Oliveira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Marcelo Tadeu Baumann Burgos, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil
Nigel Brooke, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil
Robert Evan Verhine, Universidade Federal da Bahia, Salvador, Bahia, Brasil
Tufi Machado Soares, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Colaboradores

Amanda Sangy Quiossa, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Amélia Gabriela Thamer Miranda Ramos, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Juliana Alves Magaldi, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Helena Rivelli, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Leonardo Ostwald Vilardi, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Luísa Gomes de Almeida Vilardi, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Mayanna Auxiliadora Martins Santos, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Priscila Campos Cunha, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Daniel Eveling da Silva, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Diovana de Paula de Jesus Bertolotti, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Vitor Fonseca Figueiredo, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Marcilene Barros Lima, Universidade de Brasília, Brasília, Distrito Federal, Brasil
Marina Furtado Terra, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil
Mônica da Motta Salles Barreto, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Revisão técnica e coeditoração

Mayanna Auxiliadora Martins Santos, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Revisão

Luís Carlos de Oliveira

Projeto Gráfico

Edna Rezende S. de Alcântara, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Diagramação/Capa

Coordenação de Design da Comunicação, CAEd/Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Ficha catalográfica:

Revista Pesquisa e Debate em Educação – Programa de Pós-Graduação Profissional / Gestão e Avaliação da Educação Pública / UFJF. v.10 n.1 (jan./jun. 2020) Juiz de Fora: Universidade Federal de Juiz de Fora, 2020.

Semestral

ISSN 2237-9444

1.Educação. 2.Gestão Educacional. 3.Educação – Avaliação.

CDU 371.26

SUMÁRIO

1

AVALIAÇÃO DA QUALIDADE NA EDUCAÇÃO INFANTIL: IMPASSES E PERSPECTIVAS NO BRASIL	891
Maria Malta Campos	
AVALIAÇÃO DA APRENDIZAGEM NA EDUCAÇÃO INFANTIL: ESTUDO DE CASO EM UM MUNICÍPIO BRASILEIRO	917
Bruna Ribeiro	
EDUCAÇÃO INFANTIL E AVALIAÇÃO: PROCESSOS E PERSPECTIVAS	940
Rita de Cássia de Freitas Coelho Alice de Paiva Macário	
A RESTITUIÇÃO COMO PROCESSO CENTRAL NA AVALIAÇÃO DE CONTEXTO	961
Catarina Moro	
AVALIAÇÕES DA EDUCAÇÃO INFANTIL EM MUNICÍPIOS PAULISTAS: LIMITES E POTENCIALIDADES PARA CONTRIBUIR COM A GARANTIA DO DIREITO À EDUCAÇÃO DAS CRIANÇAS PEQUENAS	978
Cláudia Oliveira Pimenta	
AVALIAÇÃO DE CONTEXTO NA EDUCAÇÃO INFANTIL: UMA ESCOLHA PARA PENSAR A MELHORIA DAS PRÁTICAS EDUCATIVAS.....	1012
Patrícia Corsino Jordanna Castelo Branco	
INFRAESTRUTURA ESCOLAR E SATISFAÇÃO PROFISSIONAL: PERCEÇÃO DE PROFESSORES DA EDUCAÇÃO BÁSICA BRASILEIRA.....	1027
Lívia Maria Fraga Vieira Edmilson Antônio Pereira Junior	
OBRIGATORIEDADE DA PRÉ-ESCOLA EM PERSPECTIVA: AVALIAÇÃO DO ATENDIMENTO NO ESTADO DO RIO DE JANEIRO	1047
Bruno Tovar Falciano Maria Fernanda Rezende Nunes	

AVALIAÇÕES EM LARGA ESCALA DE LÍNGUA PORTUGUESA: UMA PESQUISA SOBRE A COMPLEXIDADE DOS TEXTOS QUE DÃO SUPORTE A ITENS QUE AVALIAM A LEITURA	1064
Begma Tavares Barbosa Hilda Linhares Micarello Rosângela Veiga Ferreira	
A AVALIAÇÃO EXTERNA E OS NOVOS SUJEITOS DA EDUCAÇÃO.....	1082
Marcelo Baumann Burgos	
AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E ACCOUNTABILITY: UMA BREVE ANÁLISE DA EXPERIÊNCIA BRASILEIRA... 	1103
Reynaldo Fernandes Amaury Patrick Gremaud	
AVALIAÇÃO EDUCACIONAL: FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR EM CENÁRIOS DE ALTA VULNERABILIDADE SOCIAL.....	1138
Edivaldo Cesar Camarotti Martins Adolfo-Ignacio Calderón	
ESCALAS ESPECÍFICAS PARA MEDIR A INFRAESTRUTURA ESCOLAR DA EDUCAÇÃO BÁSICA.....	1160
Joaquim José Soares Neto Elianice Silva Castro	
MENSURAÇÃO EM SISTEMAS ADMISIONAIS NO ENSINO SUPERIOR E AVALIAÇÕES DE IMPACTO	1190
Tufi Machado Soares Mariana Calife Nóbrega Soares	

CARTA AOS LEITORES

Prezadas Leitoras e Prezados Leitores,

Com este número, comemoramos o décimo ano de existência da Revista Pesquisa e Debate em Educação, periódico científico vinculado ao Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública (PPGP), do Centro de Políticas Públicas e Avaliação da Educação (CAEd), da Faculdade de Educação (FACED), da Universidade Federal de Juiz de Fora (UFJF).

O primeiro número da Revista Pesquisa e Debate em Educação foi lançado em 2011. Desde 2012, foram publicados, ininterruptamente, dois números por ano, totalizando, com esta edição, 18 números. A Revista Pesquisa e Debate em Educação segue o seu propósito de publicar artigos que versem sobre temáticas educacionais contemporâneas, frutos de reflexões acadêmicas e debates entre profissionais atuantes nos espaços escolares.

Neste momento, com um olhar retrospectivo, a partir do resgate dos títulos dos números já publicados, listados a seguir, percebe-se a diversidade temática abordada, a qual abrange um amplo espectro das discussões sobre a gestão e a avaliação da educação pública, sobretudo no Brasil.

- v. 1 n. 1 (2011) Desafios de gestão: expectativas de (in) formação
- v. 2 n. 1 (2012) Avaliação dos sistemas educacionais e autonomia dos entes federados
- v. 2 n. 2 (2012) Responsabilização na educação
- v. 3 n. 1 (2013) Gestão universitária
- v. 3 n. 2 (2013) Gestão educacional
- v. 4 n. 1 (2014) Competências do século 21
- v. 4 n. 2 (2014) Gestão e avaliação
- v. 5 n. 1 (2015) Rendimento, desempenho e desigualdades educacionais
- v. 5 n. 2 (2015) Mestrados profissionais em educação

- v. 6 n. 1 (2016) Equidade, políticas e financiamento da escola pública
- v. 6 n. 2 (2016) Formação docente e ensino superior
- v. 7 n. 1 (2017) Desafios da educação no século XXI
- v. 7 n. 2 (2017) Docência, políticas públicas e legislação educacional no Brasil
- v. 8 n. 1 (2018) PIBID: a formação de docentes e o cotidiano da educação brasileira contemporânea
- v. 8 n. 2 (2018) Dossiê ensino de ciências da natureza e matemática
- v. 9 n. 1 (2019) Dossiê ensino de ciências da natureza e matemática (continuação)
- v. 9 n. 2 (2019) Financiamento da educação pública

Devemos salientar que, ao longo de todos esses anos, a Revista Pesquisa e Debate em Educação tem buscado se ajustar às políticas editoriais científicas de vanguarda, mantendo seu acesso livre e gratuito, com conteúdo distribuído via internet. Atualmente, a revista encontra-se disponibilizada no Portal de Periódicos da UFJF (<https://periodicos.ufjf.br/index.php/RPDE>) e tem ampliado a sua indexação nas bases de dados, diretórios, catálogos e portais científicos, bem como procurado se manter presente nas redes sociais. Além disso, temos trabalhado incansavelmente para ajustar a política editorial da revista aos preceitos e pressupostos da ciência aberta.

Cumpre, ainda, deixar registrado nossos sinceros agradecimentos a todas e a todos que têm contribuído, direta ou indiretamente, para a manutenção deste periódico, em especial ao público leitor, autores, avaliadores, editores, membros do conselho editorial, colaboradores, revisores, designers e diagramadores.

Esperamos que a Revista Pesquisa e Debate em Educação tenha vida longa e continue cumprindo a sua missão.

Boa leitura!

Frederico Braidá (editor-chefe)

Liamara Scortegagna (editora adjunta)

Marcello Ferreira (editor adjunto)

Mayanna Martins (coeditora e revisora técnica)

braços da política de expansão do capitalismo contemporâneo, encontrando na escola um *locus* estratégico para a formação da mentalidade ajustada a um mundo que opera e se constitui em torno do mercado. No mais das vezes, foi vinculada umbilicalmente ao gerencialismo, cujas críticas que recebeu da educação são conhecidas e extrapolam o objeto dessa apresentação. Somente a referência a esse vínculo é suficiente para entender a natureza das críticas.

A partir desse pano de fundo, as críticas à avaliação destinaram-se a diversos pontos: os altos custos da política, a imposição por parte dos governos, o incentivo desmedido a uma competição na escola e entre as escolas, o falseamento dos resultados para fins de propaganda política, a insuficiência do que é avaliado pelos instrumentos, a supremacia do argumento estatístico sobre o pedagógico, a linguagem autorreferenciada e incompreensível, o aprofundamento de desigualdades, a modelagem para o mercado, a preocupação com os resultados e não com o processo de aprendizagem, o uso punitivo em relação aos professores, a constituição de uma hierarquia entre as disciplinas escolares, o ensinar para o teste e muitos outros. É possível enfrentar essas críticas uma a uma, reconhecendo preocupações válidas e, também, exageros. Uma análise mais profunda, no entanto, mostraria que, na maior parte das vezes, tais críticas são preestabelecidas sem lastro empírico e ancoragem em evidências (curiosamente, também uma crítica dirigida à avaliação). Elas se direcionam mais ao contexto macropolítico no qual vivemos do que propriamente ao instrumento e à forma como a política de avaliação desenvolveu-se no país.

Em rápidas palavras, as avaliações têm mostrado resultados preocupantes em todas as etapas de escolaridade, evidenciando problemas profundos de aprendizagem (o que descarta, de saída, seu uso como propaganda governamental); o ensinar para o teste, quando existe, não tem surtido efeitos; os custos da política são ínfimos em relação aos gastos gerais destinados à educação; se houve uma implementação *top/down* da política de avaliação, hoje ela é apoiada, em grande parte, pelos atores das redes de ensino, e seus resultados têm sido utilizados para o desenho de políticas públicas destinadas a enfrentar os problemas identificados; não há evidências do aumento de desigualdades educacionais provocado pela avaliação, assim como não há evidências acerca do aumento da competição entre os alunos no interior das escolas; a limitação dos instrumentos é um ponto de partida da avaliação, não um problema incontornável – todo instrumento é limitado, assim como qualquer medida produzida com base nele. Sem propor uma defesa cega da avaliação, que, evidentemente, apresenta limitações, cumpre notar que a política de avaliação assumiu um papel de protagonismo nas últimas décadas no Brasil, que não pode ser atribuído somente a imposições governamentais e à alienação provocada pela ideologia capitalista.

Um caminho promissor para a compreensão da avaliação em larga escala é analisar em detalhe suas capacidades e vicissitudes, produzindo evidências sobre seus usos e efeitos. A maior parte dos problemas apontados pela crítica está vinculada mais diretamente ao uso que se faz do instrumento do que a ele próprio, ou à ideia de avaliação em si, embora o ato de avaliar ainda não tenha encontrado um tratamento adequado pela teoria educacional, levantando divergências mesmo quando se trata da avaliação interna à sala de aula, conduzida pelo professor. Uma incursão antropológica pela avaliação na escola poderia mesmo dizer que ela é uma espécie de *persona non grata* presente, considerada por vezes necessária, mas definitivamente incômoda.

Todos esses fatores levaram a avaliação educacional a uma posição secundária nos cursos de Pedagogia e licenciaturas, de modo que o tema nunca assumiu protagonismo na formação inicial dos profissionais que integram as fileiras da estrutura educacional, da equipe das secretarias às salas de aula. Entretanto, um dos primeiros desafios apresentado a esses já profissionais é lidar com os resultados da avaliação e as pressões (muitas vezes, desmedidas) que deles decorrem. A lacuna em torno da avaliação é, então, sentida especialmente pelo diretor de escola, em regra, a via de entrada da avaliação nas escolas e de quem se espera que organize o trabalho com base nessas informações, e pelos professores.

A despeito das críticas contínuas que recebeu ao longo dos anos, desde a implementação do SAEB, no início dos anos 1990, a avaliação em larga escala consolidou-se com uma das principais políticas educacionais no país, como pode ser observado por sua longevidade e continuidade. Atravessando governos de diferentes orientações ideológicas, a avaliação não apenas sobreviveu às três últimas décadas, mas desenvolveu-se e ampliou-se. Esse movimento é notório pelo surgimento de sistemas estaduais e municipais de avaliação, paralelamente à expansão do próprio sistema nacional. Os resultados da avaliação passaram a ser utilizados como suporte para a tomada de decisões nos mais diferentes níveis educacionais. As secretarias estaduais de educação tomaram a frente desse processo, assumindo protagonismo ao estabelecer parcerias com os municípios no que diz respeito à avaliação e ao compromisso com a melhoria da qualidade da educação, reforçando, assim, o pacto federativo. Esse processo de expansão é tributário muito mais da percepção da utilidade real da avaliação como base para orientar a tomada de decisões, concedendo a elas maior legitimidade, do que de uma dominação ideológica de organismos internacionais.

Ao tempo em que se tornou política pública consolidada na agenda educacional brasileira, a avaliação experimentou seu desenvolvimento como fonte e objeto de pesquisa. Impulsionada, mas não somente, pela pesquisa em eficácia escolar, que tem na avaliação uma fonte privilegiada de dados, a avaliação, como política e instrumento, passou a angariar espaço como tema de pesquisa. O Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública – CAEd/UFJF, ao qual a Revista Pesquisa e Debate em Educação está vinculada, é um exemplo disso. A utilização dos resultados e sua incorporação, no cotidiano das redes e escolas, como informação de suporte para o planejamento educacional tornaram-se temas centrais de pesquisa no âmbito da avaliação, o que não pode esconder a variedade de outros objetos de pesquisa que a área enseja: a elaboração de instrumentos, a produção de medidas, a aplicação dos testes, as características dos testes, a produção de indicadores, os questionários contextuais, a forma de divulgação dos resultados, a maneira como os resultados são lidos e interpretados, as ações de estímulos para a apropriação dos dados na escola, as políticas associadas aos resultados da avaliação, a escolha das habilidades a serem avaliadas, as parecerias estabelecidas entre os entes federados, as ações desenvolvidas a partir dos resultados para o enfrentamento de problemas por eles identificados, ações de capacitação dos atores escolares para o trabalho com a avaliação, a avaliação (política) como instrumento de avaliação de outras políticas, a avaliação como instrumento para o diagnóstico de desigualdades, o estudo do efeito-escola, as relações estabelecidas com o currículo e com as ações pedagógicas dos professores, entre muitos outros. Tudo isso é multiplicado pelas etapas e modalidades de ensino (avaliação na educação infantil, no ciclo de alfabetização, nos anos finais do ensino fundamental, no ensino médio, na educação profissional, na educação superior, na educação de jovens e adultos, nas escolas de tempo integral) e pelos tipos de avaliação (formativa, somativa, diagnóstica, digital, impressa etc.). A avaliação, como objeto de estudo, é um fenômeno multifacetado, não podendo ser reduzida ao mau uso de seus resultados, como faz parecer a crítica.

É em reconhecimento a essa pluralidade de caminhos, entradas e interesses que o presente dossiê foi organizado. Nele, uma parte dessa diversidade de pesquisas em torno da avaliação é retratada. Os autores aqui presentes são pesquisadores, professores, acadêmicos e intelectuais que tomam a avaliação como tema de pesquisa e como fonte de dados, não para o estabelecimento de uma posição que a defende a qualquer custo, mas compreendendo-a como um instrumento e uma política que, ao reconhecer seus próprios limites, coloca-se a serviço da melhoria da qualidade da educação. Seus artigos representam uma contribuição inestimável para a pesquisa em avaliação educacional no país.

O dossiê conta com uma contribuição robusta de autores destinados a analisar a avaliação na educação infantil. O artigo de Maria Malta Campos discute o tema da avaliação das políticas de educação infantil no Brasil, descortinando o contexto histórico de formação dos consensos que estão sendo questionados atualmente e analisando as propostas de avaliação da qualidade do atendimento a essa etapa de qualidade. Bruna Ribeiro apresenta os resultados de um estudo de caso realizado em um município da região metropolitana de Salvador/BA, por meio do qual é investigado um conjunto de variáveis que afetam a avaliação da aprendizagem das crianças na educação infantil. Rita de Cássia de Freitas Coelho e Alice de Paiva Macário apresentam as propostas de avaliação da educação infantil que estão sendo construídas pelo CAEd/UFJF, que podem ser aplicadas complementar ou independentemente: pesquisa de campo com o uso de questionários *online*, avaliação por meio de observação e avaliação das condições de oferta. O artigo de Catarina Moro discute, a partir da abordagem avaliativa do contexto, as vicissitudes e virtudes da restituição (entendida numa lógica negocial, participativa e formativa), contrapondo-a à lógica classificatória e excludente associada à escola e, particularmente, à avaliação. Cláudia Oliveira Pimenta analisa iniciativas de avaliação da educação infantil em dois municípios de São Paulo, apontando seus limites e suas potências, mostrando que, em parte, essas ações reproduzem elementos de avaliações destinadas a outras etapas de escolaridade. O artigo de Patrícia Corsino e Jordanna Castelo Branco analisa os resultados de uma avaliação de contexto em uma escola de educação infantil da rede pública municipal do Rio de Janeiro. As autoras dão ênfase às práticas educativas de oralidade, leitura e escrita, chamando atenção para a potência formativa da proposta de avaliação analisada. Lívia Fraga e Edmilson Pereira apresentam uma análise sobre a percepção de professores da educação básica acerca das condições de infraestrutura escolar e da satisfação profissional, como foco na educação infantil. O artigo de Bruno Tovar Falciano e Maria Fernanda Rezende Nunes debruça-se sobre os efeitos da obrigatoriedade escolar na educação infantil em uma amostra das redes municipais do estado do Rio de Janeiro, avaliando o atendimento à pré-escola, suas formas de expansão, a infraestrutura e os impactos no atendimento à creche.

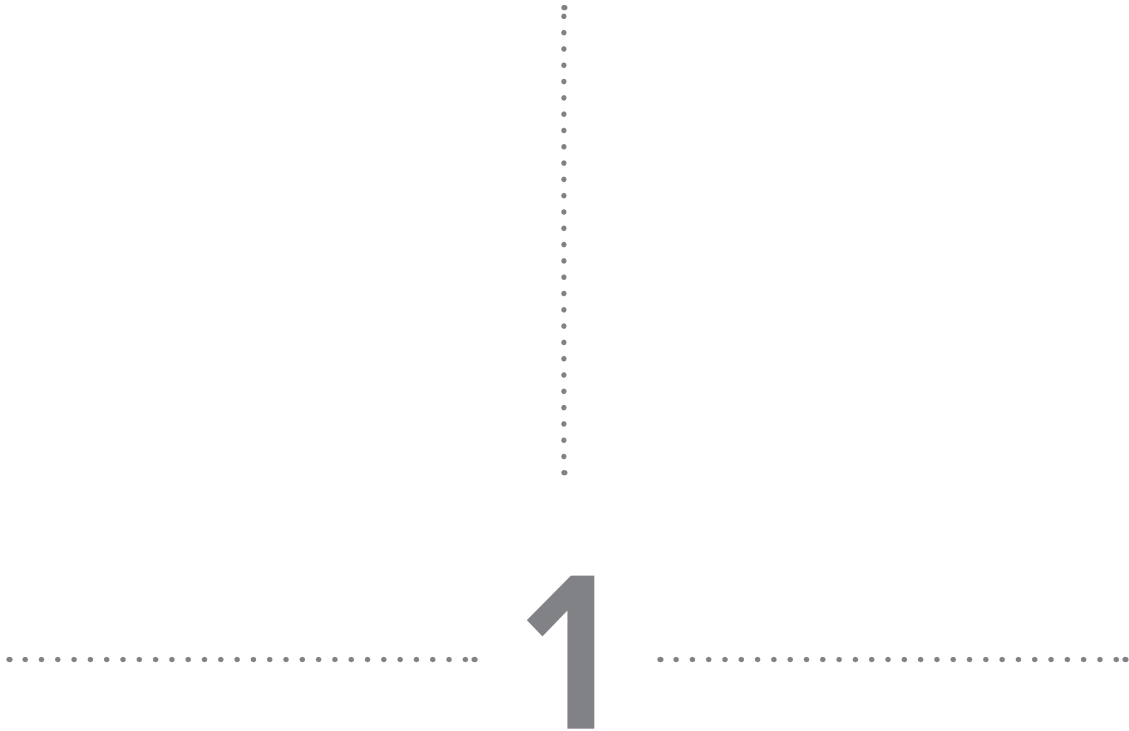
Somam-se a eles, artigos de diferentes abordagens, da análise de itens a políticas de *accountability*. Begma Tavares, Hilda Micarello e Rosângela Ferreira discutem o tema da complexidade dos textos utilizados como suporte para os itens de avaliações em larga escala em Língua Portuguesa, definindo quatro níveis de complexidade. O artigo de Marcelo Burgos analisa a mudança da avaliação externa no sistema educacional brasileiro, passando de instrumento

de controle externo a ferramenta fundamental no cotidiano da escola, fazendo surgir, nesse movimento, novos sujeitos no mundo da educação, dotados de uma nova linguagem. Reynaldo Fernandes e Amaury Patrick Gremaud analisam o movimento das políticas de *accountability* e da avaliação nas duas últimas décadas no Brasil, apresentando uma revisão da literatura internacional sobre o tema e avaliando os impactos desse movimento sobre a melhoria da aprendizagem. Com base em pesquisa qualitativa, Edivaldo Cesar Camarotti Martins e Adolfo-Ignácio Calderón analisam os fatores contextuais de eficácia escolar que contribuíram para que uma escola pública do estado de São Paulo, localizada em uma região de alta vulnerabilidade, fosse capaz de superar as metas do IDEB projetadas para os anos de 2007, 2009 e 2011. Com o objetivo de orientar políticas públicas para a educação básica, Joaquim José Soares Neto e Elianice Silva Castro apresentam uma proposta de escalas de infraestrutura específicas para todos os segmentos de ensino, definindo uma infraestrutura mínima para cada um deles. Tufi Machado e Mariana Calife Nóbrega Soares dedicam-se a analisar os sistemas de admissão (exames, provas e métodos de mensuração empregados) ao ensino superior em diversos países, incluindo o Enem.

Em todos os casos, o que está em jogo não é uma defesa irrefletida da avaliação ou a crença ingênua de que não possuindo limites, ela assume um caráter soteriológico em relação à escola e à educação. O dossiê é um convite ao leitor que deseja conhecer mais sobre a avaliação educacional, fornecendo um mapeamento para um primeiro contato, e, ao mesmo tempo, oferece ao leitor iniciado um leque de pesquisas conduzidas por importantes autores na área, retratando o quão multifacetado é seu objeto.

Wagner Silveira Rezende

Juiz de Fora, junho de 2020



PESQUISA APLICADA

AVALIAÇÃO DA QUALIDADE NA EDUCAÇÃO INFANTIL: IMPASSES E PERSPECTIVAS NO BRASIL

Maria Malta Campos¹

¹ Pesquisadora Senior – Consultora do Departamento de Pesquisas
Educaionais da Fundação Carlos Chagas, São Paulo, São Paulo, Brasil.
E-mail: c_mcampos@fcc.org.br.

Resumo

Este artigo busca levantar algumas questões sobre o tema da avaliação das políticas de educação infantil no Brasil. Parte do contexto atual sobre esse debate, no país e no exterior, mostrando como princípios e consensos estabelecidos a partir da década de 90 estão hoje sendo questionados por diversos motivos e por diferentes atores sociais. A partir dessa conjuntura, este texto aponta alguns fatores que tencionam as propostas de avaliação da qualidade do atendimento nessa etapa da educação básica, examinando os diversos argumentos que são utilizados para discutir as interfaces entre as avaliações das condições de oferta e as avaliações do desenvolvimento infantil.

Palavras-chave: Educação Infantil. Avaliação da qualidade de programas. Avaliação do Desenvolvimento Infantil.

Abstract

The article examines some questions about the evaluation of early childhood policies and programs in Brazil. The focus is on the existing context of this discussion, in this country and internationally, that shows changes in some principles and orientations established since the 1990's, which are now questioned by different social actors, based in various reasons. In this context, the article points out some influences that question the adopted evaluation policies orientations in Brazil, trying to identify important aspects to be considered in early childhood programs context evaluations and in children development assessments.

Keywords: Early Childhood Education. Program Quality Evaluation. Children Development Assessment.

INTRODUÇÃO

As políticas de educação infantil no Brasil, nos últimos anos, têm sofrido o efeito de diversos fatores importantes que se originam tanto dos contornos que o atendimento existente assumiu desde a Constituição de 1988 (BRASIL, 1988) e da Lei de Diretrizes e Bases da Educação Nacional – LDB de 1996 (BRASIL, 1996), quanto de algumas mudanças mais recentes na legislação e em programas governamentais federais, que incidem nesse campo.

A atual agenda de questões sobre o tema da avaliação das políticas de educação infantil no Brasil tornou-se mais complexa e contraditória, na medida em que princípios e consensos estabelecidos a partir da década de 90 do século passado estão hoje sendo questionados por diversos motivos e por diferentes atores sociais. Para discutir o contexto atual do debate sobre a avaliação da qualidade de creches e pré-escolas no país, é importante considerar alguns traços mais salientes das mudanças que se ensaiam sobre essa realidade, ao lado de impasses e dificuldades que remontam há muitas décadas e que nunca foram de fato superados.

Este artigo busca levantar, tentativamente, algumas questões sobre essa problemática, apontando alguns caminhos possíveis nesse contexto carregado de riscos e interrogações. Parte de uma breve caracterização da evolução recente da educação infantil no Brasil, procurando apontar alguns fatores que tencionam essa etapa da educação básica nos anos 2000, para depois focalizar algumas questões mais específicas sobre a avaliação da qualidade do atendimento.

A HERANÇA DO PROJETO DE EDUCAÇÃO INFANTIL INSCRITO NA CONSTITUIÇÃO E NA LDB

É possível identificar hoje, em muitas das disputas e mal entendidos que emergem continuamente a respeito da educação infantil no Brasil, algumas das posições que se contrapõem nos debates sobre o novo estatuto legal adotado para a educação da criança pequena no período da redemocratização. Esses dissensos focalizavam principalmente a creche, uma instituição que historicamente estava fora do âmbito de ação do campo educacional.

A inclusão da creche na primeira etapa da educação básica, estendendo o direito à educação às crianças a partir de seu nascimento, era rejeitada por muitos especialistas e militantes da área. Foi a pressão de outros grupos sociais, aliados a setores minoritários do campo educacional, que logrou inscre-

ver na legislação uma nova estrutura organizacional, baseada no critério da faixa etária atendida, um critério de caráter universal, que não mais utilizava a definição do tipo de atendimento por características da criança nem de sua família, assim como não distinguia essas subetapas da educação infantil pelo horário de atendimento integral ou parcial.

Se a inclusão da creche na educação foi consagrada na legislação e nas políticas educacionais que foram desenvolvidas desde a redemocratização de maneira consistente, ainda hoje muitas manifestações de autoridades governamentais, tanto federais como estaduais e municipais, revelam um certo desconforto e mesmo desconhecimento sobre essas orientações inscritas na Constituição e na LDB.

Um exemplo são programas que surgem, em vários níveis de governo, a partir do Marco Legal da Primeira Infância (Lei 13.257, de 2016), muitos dos quais parecem estar gerando movimentos um tanto contraditórios nessa trajetória. Por um lado, essa lei reforça uma orientação importante para as políticas sociais que atingem a criança pequena, no sentido de integrar programas de áreas como saúde, assistência social e educação dirigidos à “primeira infância”, nomenclatura que vai se tornando mais conhecida e utilizada no âmbito nacional e internacional. Contudo, em uma conjuntura de corte de recursos públicos para programas sociais, de fortalecimento de ideologias conservadoras a respeito do papel da mulher na sociedade e de empobrecimento da população, a nova lei parece estar estimulando propostas que podem enfraquecer e desviar recursos de muitas políticas de caráter universal implantadas de forma consistente desde a década de 1990, em especial na educação, na saúde e na assistência social.

Essas tendências mais recentes, paradoxalmente, se afirmam em uma fase em que a demanda social pela educação infantil adquire uma crescente importância nas dinâmicas políticas locais, especialmente nos grandes centros. Até mesmo por conta dessa pressão da sociedade, a cobertura da creche e da pré-escola vem crescendo de forma significativa nos últimos vinte anos, indicando também a emergência de um novo campo de trabalho para professores e especialistas em educação em todo o país.

O CRESCIMENTO DA COBERTURA DA EDUCAÇÃO INFANTIL NO PAÍS

Em que pese a persistência de desigualdades de acesso por renda familiar, raça, zonas rural e urbana e região geográfica, o aumento do número de matrículas em escola para a população infantil de 0 a 5 anos foi bastante significativo a partir de 2000, tanto na faixa etária da pré-escola (4 e 5 anos), quanto na faixa da creche (0 a 3 anos). Para as crianças de 4 anos e mais, o atendimento aproxima-se da universalização, como prevê a legislação que ampliou a obrigatoriedade escolar para a faixa etária de 4 a 17 anos.

Em quinze anos, entre 2002 e 2017, a porcentagem de atendimento passou de 15% para 34% na faixa etária de 0 a 3 anos, que corresponde à creche, e de 68% para 93% na faixa correspondente à pré-escola. Hoje 8,7 milhões de crianças de 0 a 5 anos estão matriculadas em escolas no Brasil. Para trabalhar com essas crianças na educação infantil, em dez anos, de 2008 a 2018, o número de professores cresceu de 363 mil para 590 mil, um aumento de 62%; comparativamente. O número de professores nos anos iniciais do ensino fundamental teve um aumento de 3,5% no mesmo período, chegando a 763 mil (TODOS PELA EDUCAÇÃO, 2019).

Ainda que abaixo da meta de 50% definida no Plano Nacional de Educação, o aumento verificado na cobertura da creche é bastante significativo. Isso revela, por um lado, a persistência de uma demanda crescente por educação infantil para crianças cada vez mais jovens e, por outro lado, representa uma nova demanda por pessoal habilitado, por novas construções, por parcerias do setor público com entidades sociais e filantrópicas, por assessorias técnicas, por novos materiais pedagógicos, inclusive brinquedos e livros apropriados para crianças bem pequenas, ou seja, toda uma gama de serviços para atender tanto às instituições privadas como públicas e conveniadas. Paralelamente, aumentou também o número de grupos de pesquisa nas universidades, a quantidade de publicações especializadas, a frequência das matérias jornalísticas e o interesse de organizações da sociedade civil. Dessa forma, a creche e a educação infantil em geral ganharam um espaço importante na agenda pública, na medida em que esse atendimento se expandiu no país.

Ao longo dos anos 2000, o governo federal atuou por meio de apoio a diversas mudanças importantes: em primeiro lugar, ampliando o escopo do financiamento público, com a aprovação do Fundeb, que incluiu a educação infantil, abrangendo também as creches, tanto públicas como conveniadas com o setor público; depois, gradativamente incluindo as creches e pré-escolas nos

diversos programas gerenciados pelo FNDE, que financiam bibliotecas escolares, transporte e alimentação, o Programa Dinheiro Direto na Escola, criando um programa de construção de novas unidades, o Proinfância, e de formação de educadores, o Proinfantil (MORO, 2017). Além disso, foram tomadas iniciativas no sentido de inserir contrapartidas em outro importante programa federal, o Bolsa Família, para incentivar a matrícula de crianças dessas famílias na educação infantil².

Vem contribuindo para essa nova posição que ocupa a educação infantil nas questões em pauta na sociedade, o discurso difundido internacionalmente sobre a importância da primeira infância para o desenvolvimento humano, especialmente nos aspectos cognitivos. A difusão das ideias do economista James Heckmann, ganhador do prêmio Nobel, sobre as vantagens econômicas de se investir nas crianças pequenas, o que representaria, segundo seus estudos, uma economia futura significativa em gastos sociais, tem sido um poderoso argumento a reforçar esse foco nos primeiros anos de vida (NCPI, 2014, p. 7–8).

O DEBATE SOBRE A QUALIDADE DO ATENDIMENTO NA EDUCAÇÃO INFANTIL

A preocupação com a qualidade dos serviços de educação e cuidado voltados para a criança de 0 a 6 anos tem acompanhado a evolução desse atendimento desde anos anteriores à Constituição de 1988. Mesmo sem o amparo legal que conquistaria com a nova legislação adotada de 1988 em diante, a precariedade verificada em muitos programas implantados em décadas anteriores, especialmente nas creches, que naqueles anos atendiam à faixa etária até os 6 anos de idade, motivou grupos de profissionais e movimentos sociais a denunciar e lutar por superar as condições que caracterizavam o atendimento em muitos dos serviços voltados para a população mais pobre.

Diversos levantamentos, diagnósticos e pesquisas constatavam o emprego de pessoal não qualificado, a excessiva quantidade de bebês e crianças pequenas por adulto, os locais insalubres, a falta de materiais, a ausência de preocupação com a programação adotada com as crianças, entre outros problemas. Muitas vezes, as próprias militantes dos movimentos de luta por creche que atuavam nas periferias das grandes cidades, ao se responsabilizarem

² Segundo Rita Coelho, O programa Brasil Carinhoso foi "um investimento pactuado entre três ministérios (Saúde, Educação e Desenvolvimento Social)" que financiou novas matrículas na educação infantil para crianças do Programa Bolsa Família (entrevista em CENPEC, 2018).

pelo funcionamento das creches comunitárias que surgiam nos bairros, tomavam consciência dessa precariedade e dos riscos que representavam para as crianças, e buscavam apoio especializado, inclusive nos órgãos municipais, para organizar melhor esse atendimento (CAMPOS, 1990).

No caso da pré-escola, que sempre esteve vinculada ao setor da educação, essa preocupação tocava mais de perto em polêmicas sobre as orientações do currículo, com o debate sobre a concepção exclusivamente preparatória para o ensino fundamental, e a emergência de novas abordagens sobre o processo de alfabetização. Correspondia a uma maior consciência sobre os graves problemas de atraso e repetência registrados nos primeiros anos da escolaridade obrigatória e a urgência de encontrar caminhos para enfrentá-los.

Mesmo com essas diferenças, tanto na creche quanto na pré-escola, a preocupação com a qualidade do atendimento não levou, como ocorreu no ensino fundamental e médio, à busca de processos de avaliação que monitorassem a evolução dos sistemas escolares a partir de medidas externas sobre o nível de aprendizagem dos alunos, as quais passaram a ser adotadas principalmente a partir da década de noventa.

Eram outras as precariedades denunciadas nas creches, fossem elas coletivas ou familiares, comunitárias ou municipais: questões sobre os cuidados de saúde e alimentação, sobre o acolhimento e a atenção às necessidades afetivas das crianças, conflitos entre educadoras e famílias, prédios inadequados, falta de material pedagógico, entre muitos outros aspectos. Ou seja, o foco sempre esteve mais nas variáveis relativas ao contexto dos serviços existentes do que no nível de aproveitamento e desenvolvimento das crianças.

Rosemberg (2013) refletiu sobre essa especificidade do debate sobre “qualidade”, no campo da educação da criança pequena, em comparação com o debate sobre “avaliação” nas etapas seguintes da educação básica. Segundo ela, “desde as primeiras manifestações públicas contemporâneas em prol das creches – militantes e acadêmicos a partir dos anos 1970, governamentais alguns anos depois –, vêm ocorrendo, em determinados setores do país, uma intensa mobilização pela expansão da oferta e melhoria de sua qualidade com base em avaliações que nem sempre receberam essa denominação” (ROSEMBERG, 2013, p.48). Como bem observou em seu artigo, “o termo diagnóstico foi usado mais amplamente como avaliação de uma determinada situação de oferta” (ROSEMBERG, 2013, p.49) na literatura que tratava das condições de funcionamento de creches e pré-escolas, ao passo que para as etapas seguintes da educação, a discussão sobre “avaliação” inclui tanto o foco nos ganhos de aprendizagem dos alunos, quanto a avaliação das condições de oferta do ensino.

É interessante observar que um documento sobre qualidade na creche foi editado pelo MEC antes mesmo que surgisse a primeira proposta curricular oficial para a educação infantil pós Constituição de 1988: “Critérios para um atendimento em creches que respeite os direitos fundamentais das crianças” foi publicado em 1995 (BRASIL, 1995) e “Referencial curricular para a educação infantil” foi publicado em 1997 (BRASIL, 1998).

Na sequência, observa-se uma reafirmação, em diversos documentos orientadores, da concepção de avaliação na educação infantil restrita ao acompanhamento da criança pela professora, no âmbito da creche ou pré-escola, sem a finalidade de retenção ou promoção, inclusive na passagem para o ensino fundamental, como expressamente disposto no texto da LDB de 1996³. Nessa década, os dados recolhidos pelo INEP e pelo IBGE revelavam um grande contingente de crianças acima de 7 anos retidas nas chamadas “classes de alfabetização”, especialmente nas regiões menos desenvolvidas, agravando as altas taxas de atraso escolar no país. Foi a mudança no sistema de financiamento da educação, com a adoção do Fundef em 1996, que induziu as redes locais a matricularem as crianças de 7 anos no ensino fundamental e desestimulou essas práticas de retenção, as quais seriam também combatidas por meio da introdução dos ciclos escolares no ensino fundamental (CAMPOS, 2012).

Enquanto isso, a consolidação dos sistemas de avaliação de resultados de aprendizagem por meio da aplicação externa de testes padronizados aos alunos dos demais níveis de ensino pelo INEP, assim como por diversos sistemas estaduais e até municipais, fez com que o tema da avaliação educacional ganhasse grande legitimidade pública no país.

No entanto, até bem recentemente, a educação infantil tinha permanecido fora do alcance dessas medidas de “prestação de contas” (*accountability*, em inglês), disseminadas internacionalmente, que justificam esses sistemas de avaliação externa de políticas públicas. É possível que isso tenha ocorrido tanto por uma maior importância socialmente atribuída à aprendizagem dos conteúdos tradicionalmente abordados na escola básica a partir dos 6 e 7 anos, quanto por estratégias utilizadas por profissionais e militantes da educação infantil para manter essa etapa “protegida” dos grandes sistemas de avaliação educacional existentes.

³ A Lei de Diretrizes e Bases da Educação Nacional, em seu art. 31, Seção II, Da Educação Infantil, define que nessa etapa “a avaliação far-se-á mediante acompanhamento e registro de seu desenvolvimento, sem o objetivo de promoção, mesmo para o acesso ao ensino fundamental”.

Essa situação vem se modificando nos últimos anos. Vários fatores podem ser considerados na busca de um melhor entendimento sobre esse processo. De forma tentativa, neste texto serão sugeridas algumas hipóteses sobre as causas dessas mudanças.

Em primeiro lugar, a disponibilização pelo IBGE e pelo INEP de dados mais precisos sobre o atendimento na educação infantil, especialmente na faixa da creche, que até o início dos anos 2000 tinha ficado praticamente excluída das estatísticas oficiais, permitiu que diversas pesquisas revelassem aspectos da cobertura e das condições de oferta desses serviços que trouxeram novos elementos para a avaliação de sua qualidade. Assim, tornou-se evidente as inequidades na oferta da educação infantil, segundo a faixa de renda das famílias, sua pertinência racial e a localização de seu domicílio; as deficiências de pré-escolas e creches puderam ser analisadas com base em dados recolhidos pelo Censo Escolar, com informações sobre pessoal, instalações físicas e organização dos estabelecimentos, permitindo comparações entre regiões e estados, entre redes públicas e privadas, entre zonas rurais e urbanas. Como uma grande lente que se tornava gradativamente mais potente, inclusive por pressão de muitos pesquisadores e estudiosos do tema, o exame dessas informações já apontava para questões que motivavam novas pesquisas e a busca de instrumentos para avaliar esses serviços.

Segundo, o crescimento das redes de educação infantil, tanto públicas como privadas, estimulou a atuação de especialistas e organizações da sociedade civil em projetos de assessoria e formação em serviço do contingente cada mais numeroso de profissionais empregados em estabelecimentos de educação infantil. Muitas vezes ligadas a universidades e instituições de pesquisa, essas intervenções permitiram um maior conhecimento sobre o cotidiano das creches e pré-escolas, evidenciando, de um lado, seus problemas e precariedades, mas de outro permitindo vislumbrar suas potencialidades no cuidado e educação das crianças pequenas em ambientes coletivos. A aproximação com a vivência de bebês e crianças pequenas naquelas unidades com problemas mais sérios de funcionamento e com educadores sem preparo adequado fortaleceu também a preocupação com os riscos ao desenvolvimento infantil, especialmente naquelas instituições localizadas em bairros e locais onde vivem as famílias com piores condições de vida. Na percepção de muitos especialistas, seria necessário contar com outras abordagens e novos instrumentos para avaliar de forma mais objetiva a qualidade de creches e pré-escolas, permitindo identificar e corrigir aquelas situações que apresentam riscos ao desenvolvimento infantil.

Outro fator importante, a divulgação internacional de estudos sobre os efeitos da frequência à educação infantil sobre os resultados de aprendizagem na continuidade da escolaridade das crianças, especialmente pesquisas de larga escala que acompanharam coortes de alunos por longos anos, trouxeram novas evidências sobre quais variáveis – o tipo de frequência, a modalidade de atendimento, a orientação curricular, o grau de formação do pessoal, a relação entre escola e família – estão associadas a resultados mais positivos e, ainda, qual o tipo de ambiente familiar das crianças que contribui para esses resultados, controlando-se o nível social e outras variáveis relativas à população atendida. Evidências de programas de pesquisas como o projeto EPPE (*Effective Pre-school and Primary Education*) no Reino Unido ou as realizadas pela rede de pesquisas apoiada pelo NICHD (*National Institute of Child Health and Human Development*) nos Estados Unidos permitem gerar recomendações relevantes para as políticas que incidem sobre as crianças pequenas, especialmente as educacionais, que demandam investimentos públicos importantes a médio e longo prazo (SYLVA et al., 2010; MELHUIH, 2013).

Melhuish (2013), um dos pesquisadores responsáveis pelo projeto EPPE, resume as “cinco áreas que se mostraram particularmente importantes” em relação a resultados mais significativos para as crianças em fases posteriores à sua frequência na pré-escola, constatados por estudos de caso que fizeram parte desse projeto:

1. qualidade da interação verbal adulto – criança;
2. conhecimento e entendimento do currículo, pela equipe;
3. conhecimento da equipe sobre como as crianças aprendem;
4. habilidades dos adultos para ajudar as crianças a solucionarem conflitos;
5. apoio aos pais para favorecer a aprendizagem em casa (MELHUIH, 2013, p. 137).

Esses são fatores cruciais que deveriam ser priorizados na gestão das redes de educação infantil, nos currículos de formação inicial e continuada de professores, nas práticas de supervisão pedagógica nas unidades, nas relações entre escolas e famílias. Note-se que, no Reino Unido, as crianças são matriculadas na pré-escola a partir dos 3 anos, idade que no Brasil ainda compõe a faixa etária atendida por creches.

Contribuem também, para esse conhecimento, as revisões que examinam conjuntos grandes de investigações em diversos países, procurando encontrar tendências gerais que se confirmam em diferentes contextos, por exemplo, os levantamentos publicados em documentos de agências europeias, como aqueles que constam na série *Starting Strong*, especialmente os volumes II, III e IV (OECD, 2006, 2011 e 2015). De acordo com o comentado em artigo anterior (CAMPOS, 2013), uma revisão extensiva de pesquisas realizada por Paul Leseman (2009, apud CAMPOS, 2013, p.38) traz evidências “que detalham quais características dos programas explicam os diferentes resultados observados no acompanhamento longitudinal dos alunos, em sua vida escolar e fora dela”. Características relativas ao formato dos programas, aos modelos de currículos mais adequados a cada faixa de idade, ao tipo de envolvimento das famílias, assim como dados sobre as diferenças de impacto em grupos de nível de renda distintos contribuem para subsidiar decisões de políticas educacionais com maiores potenciais de impactos positivos no desenvolvimento infantil. Além disso, Leseman (2009), assim como Melhuish (2013), não deixam de mencionar aqueles aspectos sobre os quais as pesquisas oferecem resultados contraditórios e/ou não conclusivos, como no caso dos efeitos da frequência a creches domiciliares e coletivas de crianças com menos de 2 anos de idade, fase na qual a qualidade do atendimento parece ser crucial para determinar o tipo de impacto sobre o desenvolvimento infantil (CAMPOS, 2013).

Em quarto lugar e somando-se ao anterior, o foco interdisciplinar na abordagem da situação da primeira infância, que começa a ganhar força entre nós, atraiu para a questão da educação infantil um interesse crescente de pesquisadores de outras áreas científicas, como a psicologia, a medicina e a economia, as quais não enfrentam as mesmas resistências que o campo da educação apresenta em relação ao emprego de instrumentos padronizados de avaliação externa de creches e pré-escolas ou mesmo de avaliações do desenvolvimento infantil, com testes aplicados às crianças, muitas vezes utilizando abordagens metodológicas experimentais ou quasi-experimentais.

Nesse novo contexto, a área da pesquisa sobre a educação infantil no Brasil tem reagido de diversas formas. Alguns grupos, retomando tendências já adotadas em outras décadas, têm procurado desenvolver novas parcerias com países que contam com maior experiência em processos de intervenção e avaliação em redes de educação infantil, como a Itália, os países nórdicos, o Reino Unido, Portugal, Espanha, a Austrália e os Estados Unidos, entre outros. Muitas experiências que associam processos internos de melhoria da qualidade com adoção de instrumentos padronizados ou adaptados para cada caso, com apoio em equipes externas de pesquisa e assessoria, têm sido

divulgadas em encontros e congressos no Brasil, motivando projetos de intercâmbio e colaboração, assim como novas publicações que ajudam a difundir tais iniciativas.

Por sua vez, no período que transcorreu até 2016 e nos dois anos seguintes, o MEC adotou diversas orientações na área da avaliação da educação infantil. De um lado, continuou a reafirmar a posição sobre a avaliação da criança como sendo de responsabilidade dos educadores no contexto das instituições, com o caráter exclusivo de registro e acompanhamento, sem consequências para promoção ou retenção dos alunos. Por outro lado, foram promovidas iniciativas sobre avaliação de políticas e programas de educação infantil no Brasil, entre as quais destacam-se três linhas de ação: uma primeira que, em colaboração com o Banco Interamericano de Desenvolvimento (BID), financiou pesquisa que, pela primeira vez, realizou uma avaliação externa das condições de funcionamento de amostras representativas de creches e pré-escolas públicas, conveniadas e privadas, em seis capitais brasileiras (CAMPOS et al, 2010); a segunda, que apoiou a criação de um instrumento de avaliação institucional participativa para uso de equipes de creches e pré-escolas, os Indicadores de Qualidade da Educação Infantil, cujo uso nas redes municipais foi posteriormente objeto de ampla pesquisa realizada em colaboração com a UNDIME; e a terceira iniciativa, que criou um grupo de estudos encarregado de formular uma proposta de política de avaliação da educação infantil no país, a ser adotada pelo INEP, que culminou na Avaliação Nacional da Educação Infantil (ANEI) (BRASIL, 2016b), a qual deveria ter sido realizada pela primeira vez em 2017, depois em 2019, mas acabou sendo adiada pela administração que assumiu o governo federal nesse ano. O projeto da ANEI prevê a avaliação das condições de oferta da educação infantil, aferindo “a infraestrutura física, o quadro de pessoal, as condições de gestão, os recursos pedagógicos, a situação de acessibilidade, entre outros indicadores contextuais relevantes” (BRASIL, 2016b, p. 6).

Em anos mais recentes, o crescente protagonismo de organizações da sociedade civil, que atuam na área educacional com apoio em fundos privados e promovendo projetos que contam com financiamentos de agências públicas nacionais e com o suporte de universidades e centros de pesquisa internacionais, vem enfatizando a importância da avaliação na agenda de questões sobre políticas de educação infantil e de atenção à primeira infância, incluindo a avaliação do desenvolvimento infantil, no contexto da família, e/ou em creches e pré-escolas. Seminários, cursos, publicações, notícias na imprensa e na internet têm divulgado essas abordagens. Ao mesmo tempo, no contexto de muitas investigações e projetos de assessoria, diversos instrumentos inter-

nacionais de avaliação de serviços de educação infantil e de crianças na faixa etária de 0 a 5 anos têm sido adaptados para uso no Brasil por várias equipes de pesquisa.⁴

AS DIVERSAS ABORDAGENS DE AVALIAÇÃO NO CAMPO DA EDUCAÇÃO INFANTIL: DISPUTAS E CONVERGÊNCIAS

Em trabalho anterior, Campos (2016) havia tentado elencar algumas questões ligadas aos diversos âmbitos em que se desenvolvem as avaliações na educação infantil, sugerindo caminhos para mudanças necessárias e identificando divergências entre posições distintas em alguns casos. A seguir, são retomadas algumas questões abordadas naquele texto.

Avaliação da criança no contexto da instituição de educação infantil

A diretriz que indica que essa avaliação deve se constituir no acompanhamento e nos registros que a professora faz sobre o desenvolvimento das crianças ao longo do ano está expressa em uma página das Diretrizes Curriculares Nacionais para a Educação Infantil de 2009 (BRASIL, 2009a) e em poucas linhas da Base Nacional Comum Curricular (BRASIL, 2016a). Essas orientações contam com um relativo consenso entre especialistas, mas sendo bastante sintéticas, não parecem ter sido suficientes para subsidiar os sistemas e as instituições nas práticas constatadas por algumas pesquisas.

Levantamento realizado por Ribeiro (2016), com dados sobre 125 municípios de 22 estados, analisou 23 instrumentos de avaliação de crianças em uso nas regiões Sudeste e Nordeste. Uma das informações trazidas pela pesquisa é reveladora: “grande parte dos mais de 200 gestores públicos participantes do estudo reconhece que possui pouco ou nenhum repertório e referência sobre avaliação em educação infantil e pede auxílio com maiores informações advindas do governo federal” (RIBEIRO, 2016, p. 232). O exame dos 23 instrumentos localizados pela pesquisa, a maioria deles utilizados pelas professoras, sugere tendências interessantes: os tópicos abordados parecem estar apoiados em conteúdos do Referencial Curricular Nacional para a Educação Infantil (BRASIL, 1998), e não nas Diretrizes Curriculares Nacionais para a Educação Infantil (BRASIL, 2009a) elaboradas pelo Conselho Nacional de Educação e de

⁴ Uma entidade que vem assumindo crescente protagonismo nesse campo é a Fundação Maria Cecília Souto Vidigal – FMCSV, de São Paulo. A autora deste texto faz parte do Comitê Científico do Núcleo Ciência pela Infância, um grupo colaborativo multidisciplinar de pesquisadores, apoiado pela FMCSV e outras organizações.

caráter mandatório, publicadas anos depois. Além disso, a pesquisa localizou alguns documentos que mencionavam a promoção das crianças para o nível seguinte da educação infantil, o que é vedado pela LDB e por pareceres do Conselho Nacional de Educação. Outro levantamento, realizado por Pimenta (2017) sobre 42 municípios paulistas, encontrou uma preponderância de desenhos avaliativos com foco na aprendizagem das crianças, com ênfase nas habilidades de leitura, muitas vezes com uso de fichas elaboradas pelo órgão central, confirmando muitas tendências apontadas por Ribeiro (2016).

Souza e Pimenta (2018, p.1295), em artigo que apresenta uma revisão de estudos sobre o tema, concluem “serem ainda incipientes as iniciativas implementadas por municípios, na medida em que estas tendem a se restringir à aprendizagem das crianças, elegendo este indicador como expressão da qualidade da educação infantil”. Elas argumentam que outros aspectos importantes, relativos à avaliação do contexto da oferta do atendimento, que incluía aspectos relativos à gestão municipal da educação infantil, teriam de ser contemplados.

Como já indicado em texto anterior (CAMPOS, 2016), esses resultados sugerem que a maioria das creches e pré-escolas não sabe traduzir as orientações mais gerais de forma a: avaliar as crianças individualmente ao longo do ano; utilizar os registros para subsidiar o trabalho pedagógico; comunicar adequadamente e mais frequentemente aos pais os resultados de seus filhos; permitir que as crianças participem desse processo.

Um levantamento realizado pela OCDE mostra que, na maioria dos países estudados, as práticas de acompanhamento do desenvolvimento infantil são definidas localmente e não reguladas nacionalmente. Elas diferem muito entre si, sendo que diversos instrumentos são utilizados, desde aqueles elaborados no local até os adaptados de instrumentos internacionais padronizados. Avaliações das crianças baseadas em observações e registros descritivos são as mais comuns. Elas são realizadas por profissionais da unidade, algumas vezes complementadas por gestores e pessoas externas e, em alguns casos, prevendo consulta aos pais (OECD, 2015).

Em Portugal, foi adaptado um sistema de avaliação desenvolvido no Reino Unido por Bertram e Pascal (2004; 2006), para crianças de 3 a 5 anos e para a faixa etária de até 3 anos, respectivamente o EEL (*Effective Early Learning*) e o BEEL (*Baby Effective Early Learning*). O sistema inclui diversos instrumentos, permitindo uma utilização flexível, adaptada à realidade de cada rede ou serviço. O sistema Desenvolvendo a Qualidade em Parcerias, apresentado por Oliveira–Formosinho et al. (2016, p.14) também permite uma diversidade de usos, integral e parcelados, e foi planejado “para ser contextualizado em diferentes

realidades e situações”. Para testar o sistema, foram realizados quinze estudos de caso naquele país. Folque (2012) descreve as diversas abordagens da avaliação de qualidade que esse sistema abrange: fornece roteiros de entrevistas para profissionais das unidades, para as crianças e para seus familiares, sugerindo um modelo de investigação/ação, com participação da equipe e de um observador externo; inclui duas escalas de observação, desenvolvidas originalmente por Laevers, a Escala de Envolvimento da Criança e a Escala de Empenhamento do Adulto (PORTUGAL; LAEVERS, 2010).

Trata-se de um sistema, portanto, que associa a avaliação das condições de oferta e a observação de situações do dia a dia que focalizam interações entre o adulto e as crianças, que são importantes indicadores da qualidade do trabalho pedagógico ali desenvolvido. São três as categorias do comportamento do professor cobertas pela segunda escala: sensibilidade, estimulação e autonomia⁵. O sistema inclui ainda uma ficha de observação, a Ficha de Oportunidades Educativas, que orienta a observação das crianças, permitindo uma avaliação de suas experiências no ambiente da instituição. Os estudos de caso demonstram que o sistema pode ser utilizado em diferentes contextos e contribuir para um processo de melhoria da qualidade adaptado às necessidades e possibilidades de cada realidade (FOLQUE, 2012).

Como mostram Pascal e Bertram (2016, p. 82),

existe uma gama ampla de métodos de avaliação de unidades e de crianças que são apropriados em abordagens pedagógicas participativas. Nenhum método isoladamente pode oferecer toda a evidência, nem acomodar a participação de todos os grupos interessados, e por isso um repertório de métodos é necessário para permitir interrogações e julgamentos coletivos de contextos pedagógicos, processos e resultados.

Em seguida, os autores relacionam diversos métodos, instrumentos e recursos que podem ser utilizados em processos de avaliação participativos de avaliação de contextos pedagógicos, processos pedagógicos, resultados pedagógicos e avaliação de unidades.

Esses exemplos demonstram que é possível combinar a avaliação da unidade com a avaliação da qualidade das interações entre educadores e crianças, que a maioria das pesquisas aponta como sendo cruciais para favorecer o desenvolvimento infantil nessa etapa. Como mostra Yoshikawa et. al. (2013), cada vez mais acumulam-se evidências de que interações adulto-criança que estimulam pensamentos mais complexos que levam a novas aprendizagens e

⁵ Para a fundamentação teórica em que se apoiam esses instrumentos, ver Laevers (2014).

ao mesmo tempo são calorosas e encorajam respostas das crianças são os principais fatores que permitem obter efeitos positivos no desenvolvimento infantil. Segundo esses autores, aspectos do atendimento que dizem respeito às condições de oferta, como infraestrutura, tamanho das turmas e qualificação do professor ajudam a tornar essas interações possíveis, mas não garantem por si só que elas de fato ocorram.

Experiências como as citadas, entre muitas outras disponíveis na literatura internacional, sugerem como as equipes das secretarias municipais e das unidades de educação infantil podem ser subsidiadas com orientações e materiais, para que possam traduzir as diretrizes mais gerais adotadas nos documentos oficiais, em práticas já testadas em diversos contextos e que podem ser adaptadas a cada situação.

Auto-avaliação institucional participativa na educação infantil

Modalidades de avaliação de condições de oferta que valorizam práticas participativas que envolvem não somente as equipes das unidades, mas também os pais das crianças e pessoas da comunidade, foram divulgadas no Brasil, principalmente com a publicação da série de documentos conhecidos como Indique's: inicialmente destinada ao ensino fundamental e publicada pela organização da sociedade civil Ação Educativa. Na continuidade, uma parceria com o MEC levou à publicação dos Indicadores da Qualidade na Educação Infantil em 2009 (BRASIL, 2009b). O documento propõe um trabalho coletivo de auto-avaliação, cobrindo sete dimensões de qualidade, descritas na forma de indicadores, cada um contendo perguntas que devem ser respondidas durante uma dinâmica de grupo prevista para durar um dia, e que deve resultar na elaboração de um plano de ação para cada unidade.

Essa publicação foi amplamente divulgada no país, como mostra uma pesquisa realizada com a colaboração da União Nacional dos Dirigentes Municipais de Educação (UNDIME) que enviou questionários a todos os municípios brasileiros, obtendo um índice de respostas de 44%. Os resultados mostraram que o documento estava sendo utilizado em metade desses municípios, em auto-avaliações, mas também como material de apoio para supervisão das unidades, formação de educadores e avaliação externa das instituições (BRASIL, s./d.).

Diversas experiências têm sido desenvolvidas no país a partir dessa proposta, seja com uso do documento do MEC, seja com versões adaptadas a partir dele. Uma em especial é importante para esta discussão, pois traz subsídios para estimar os potenciais, mas também as limitações dessa abordagem. Foi

desenvolvida na rede municipal de São Paulo, no período de 2013 a 2016, incluindo a elaboração, testagem e publicação de um documento próprio, e está descrita e comentada em dois relatórios (CAMPOS; RIBEIRO, 2016 e 2017). Este processo teve continuidade em São Paulo até a data de preparação deste artigo.

As conclusões apontam progressos verificados na formação continuada dos profissionais, no diálogo entre educadores e familiares das crianças, no fortalecimento de práticas de gestão democrática, no aperfeiçoamento do projeto pedagógico das unidades e na colaboração entre as equipes regionais e das creches e pré-escolas sob sua supervisão. No entanto, reconhecem que esses progressos não ocorreram da mesma forma em todas as unidades, tendo sido identificada uma tendência já registrada em outras experiências: as unidades com equipes mais maduras e bem preparadas tendem a ser mais rigorosas em suas avaliações, ao contrário de outras, com maiores dificuldades de desenvolver o processo de auto-avaliação, que paradoxalmente tendem a não reconhecer que existem problemas de qualidade em sua unidade. Dada a dimensão da rede paulistana de educação infantil – mais de 2.500 unidades na época da pesquisa, entre pré-escolas, creches diretas e conveniadas com entidades – não foi possível, como em outras experiências, contar com uma pessoa externa à unidade que atuasse como “mediador” ou “facilitador”, trazendo um olhar mais crítico ao processo de autoavaliação (CAMPOS; RIBEIRO, 2016).

A natureza dessa abordagem não permite que seus resultados sejam utilizados para comparações entre unidades ou para constatar melhorias de qualidade ao longo do tempo, ao contrário de metodologias que se apoiam em avaliações externas, com uso de instrumentos padronizados e aplicados por pessoas treinadas. Por outro lado, a metodologia proposta abre um leque grande de possibilidades a serem exploradas na gestão pedagógica das unidades, na formação em serviço do pessoal, na supervisão realizada pelas equipes da secretaria, no planejamento da expansão da rede, e pode oferecer importantes oportunidades para a pesquisa.

Na Itália, uma experiência relatada por Antonio Gariboldi e Marina Maselli, desenvolvida na região da Emilia Romagna, entre os anos de 2012 e 2015, mostra como nas condições mais favoráveis de redes de menor escala e no contexto de uma região famosa internacionalmente por suas experiências na educação infantil, é possível associar em um sistema de avaliação construído de forma negociada, duas funções: “uma de caráter certificativo da qualidade estrutural de elementos de serviço, isto é, aqueles elementos que representam condições e garantias para a qualidade dos processos educacionais; e

outra, formativa dos processos ativados em contextos educacionais para a primeira infância” (GARIBOLDI; MASELLI, 2018). No caso da certificação, os autores explicam que esta exigência tornou-se necessária em função das políticas de terceirização de serviços públicos, que compõem um “sistema plural com recursos e abordagens diversificados” (GARIBOLDI; MASELLI, 2018, p. 2-3).

Outras iniciativas em regiões próximas do norte da Itália, cujas equipes trabalham em parceria com colegas de universidades brasileiras, lideradas por pesquisadoras da Universidade Federal do Paraná, adotam abordagens semelhantes, associando o suporte de pesquisadores de universidades ao trabalho das redes de atendimento à primeira infância, promovendo processos participativos de melhoria de qualidade que incluem o emprego de instrumentos de avaliação de contexto construídos e/ou adaptados a partir de escalas já existentes (BONDIOLI, 2018). Entre outras características, as experiências desenvolvidas por pesquisadoras da Universidade de Pavia utilizam também apoio externo ao trabalho avaliativo realizado nas unidades, “com funções de mediação” (MORO; SOUZA, 2014; SAVIO, 2018). Um projeto de colaboração com grupos de pesquisa no Brasil procurou testar abordagens semelhantes em uma amostra de instituições, identificando os principais potenciais e dificuldades que esse tipo de proposta encontra em nossa realidade (SOUZA; MORO; COUTINHO, 2015).

Avaliação externa de instituições de educação infantil

A pesquisa Educação infantil no Brasil: avaliação qualitativa e quantitativa, promovida pelo MEC com financiamento e apoio do BID, foi realizada por equipe sediada na Fundação Carlos Chagas, com participação de seis equipes de pesquisadores de universidades federais situadas nas seis capitais de cinco regiões do país que foram cobertas pela amostra (CAMPOS et al., 2011). Pela primeira vez, de forma mais abrangente, foi realizada uma avaliação externa de amostras representativas de estabelecimentos públicos, conveniados e privados que atendiam crianças de 0 a 6 anos no Brasil, com o uso de dois instrumentos padronizados de avaliação que são amplamente utilizados em todo o mundo há muitos anos: as escalas ECERS-R (HARMS; CRYER; CLIFFORD, 2003) e ITERS-R (HARMS; CLIFFORD; CRYER, 1998). Na média, as instituições avaliadas apresentaram pontuações insuficientes, sendo que em alguns aspectos elas foram muito baixas. As entrevistas realizadas com professores, coordenadores pedagógicos e diretores de unidades, mostraram que a maioria deles não seria capaz de atuar sobre os pontos críticos revelados pela pesquisa sem apoio externo (CAMPOS, 2013).

Em muitos países, tem sido desenvolvidos sistemas de credenciamento e monitoramento da qualidade de serviços de educação e cuidado para crianças pequenas. A Austrália é um caso que vem despertando muito interesse internacionalmente, por ter adotado um sistema abrangente, proposto centralmente, mas executado pelos governos locais, que inclui autorização de funcionamento e acompanhamento de processos de melhoria de qualidade junto aos diversos tipos de atendimento que ali coexistem, desde aqueles sediados nos domicílios, até creches e pré-escolas públicas e privadas com e sem fins lucrativos. Como explica Tayler (2014), que participou da elaboração do sistema, a educação infantil em seu país comporta um leque amplo de diversos tipos de serviços. Mesmo aquelas instituições privadas com fins lucrativos podem receber indiretamente benefícios governamentais, via subsídios às famílias. Procurou-se, assim, implantar um sistema em que todos os serviços e programas tem sua qualidade avaliada e monitorada, estando previstas obrigações legais e consequências nos casos de não haver cooperação das unidades. O processo desenvolvido, segundo Tayler (2014), inclui um trabalho colaborativo com as equipes, mas o resultado da avaliação é de responsabilidade de avaliadores externos treinados. As unidades podem apelar e solicitar revisão de suas pontuações que, por lei, podem ser divulgadas. A autora informa que, em 2014, 35% dos serviços atendiam aos requisitos nacionais, e 26% os superavam. A área identificada como aquela que apresenta a maior dificuldade para os serviços é a que se refere ao programa e à prática educacional.

Nos Estados Unidos, a organização não governamental *National Association for the Education of Young Children (NAEYC)* oferece um sistema de credenciamento que pode ser contratado por unidades ou sistemas locais. É um sistema que inclui o emprego de vários instrumentos e permite que os serviços adotem programas internos de melhorias dos aspectos que apresentam problemas por ocasião das primeiras avaliações e requeiram revisões desses resultados após mudanças introduzidas em suas condições de funcionamento e em suas práticas.

Esses exemplos são importantes para sugerir formas de avaliação que podem ser adotadas pelas administrações municipais no Brasil para monitorar a qualidade das creches geridas por entidades privadas sem fins lucrativos que estabelecem convênios com o poder público. De fato, esta é uma modalidade de oferta de serviços de creche muito disseminada no país e que nem sempre conta com condições básicas de funcionamento, situação que se agrava quando cresce a pressão da demanda e as prefeituras passam a improvisar formas de expandir rapidamente as matrículas.

Avaliação individual de crianças em contextos de pesquisa

A divulgação de algumas pesquisas realizadas no Brasil que utilizaram instrumentos de medida do desenvolvimento infantil como indicadores da qualidade de instituições que atendem à pequena infância provocaram grande polêmica nos meios educacionais. Parte das críticas focalizava o instrumento utilizado por um desses projetos, julgado inadequado para esse uso (o ASQ-3 – *Agas and Stages Questionnaire*⁶). *Mas grande parte da rejeição se dirige à avaliação em si, como se houvesse um impedimento ético, além de uma inadequação metodológica geral no uso de instrumentos de avaliação do desenvolvimento de crianças pequenas em pesquisas.*

Questões éticas relativas às pesquisas com crianças pequenas são relevantes e requerem especial atenção de todo pesquisador. Elas se encontram especificadas em diversos protocolos, nacionais e internacionais. Esses cuidados sempre podem ser melhor explicados e divulgados com o objetivo de proteger as crianças e suas famílias de abordagens que as prejudiquem. Exemplos de protocolos de ética que devem ser observados em pesquisas com crianças em contextos de creche e pré-escola podem ser encontrados em documentos da *National Association for the Education of Young Children* e em proposta elaborada por pesquisadores filiados à *European Early Childhood Education Research Association* (EECERA) (FORMOSINHO et al., 2016).

Porém, hoje é importante argumentar em favor da necessidade de se contar com dados mais precisos sobre os efeitos de diversos fatores que afetam o desenvolvimento infantil nas condições específicas do contexto brasileiro, tanto no ambiente da família, como no ambiente das creches e escolas, e em outros espaços, como em abrigos e hospitais. Assim como no Reino Unido foram identificados, entre os modelos de serviços que atendem às crianças pequenas, aqueles que produziram os melhores resultados na continuidade de sua escolaridade, seria muito importante identificar, dentre as diversas formas de atendimento existentes no Brasil, quais apresentam as melhores possibilidades de serem aperfeiçoadas e expandidas com o objetivo de trazer os maiores benefícios para a população infantil no presente e no futuro. Mais ainda, neste momento em que se procura rever os currículos dos cursos de formação de professores, seria muito oportuno contar com resultados de pesquisa que subsidiassem os futuros educadores quanto àquelas práticas

⁶ O ASQ-3 é um instrumento de triagem de distúrbios de desenvolvimento para crianças de 0 a 5 anos de idade (MORO; SOUZA, 2014.). Esse instrumento foi utilizado em uma pesquisa realizada na rede de creches do Rio de Janeiro (BARROS et al., 2011).

pedagógicas que mais favorecem o desenvolvimento infantil nas condições que caracterizam a realidade das famílias e das creches e pré-escolas no país.

Não necessariamente a realização de pesquisas que se utilizam de medidas do desenvolvimento infantil levam à adoção de sistemas abrangentes de avaliação externa nos moldes daqueles adotados nas etapas seguintes da educação. Uma coletânea que reúne estudos com dados atualizados sobre políticas de atendimento para crianças de 0 a 3 anos, em doze países de diversos continentes, não encontrou nenhuma modalidade de avaliação de alunos generalizada, com uso de grandes amostras ou de tipo censitário. O que foi constatado é que, em muitos países, existe um uso mais disseminado de instrumentos padronizados para uso dos educadores nas unidades (BHERING; ABUCHAIM, no prelo).

À GUIA DE CONCLUSÃO

O panorama descrito sobre as diversas abordagens existentes no campo da avaliação de programas e de avaliação de crianças em ambientes de creches e pré-escolas mostra diversas perspectivas que se cruzam e se contradizem, mas muitas vezes revelam aspectos complementares e abrem possibilidades interessantes para promover melhorias de qualidade na educação infantil no Brasil.

Muitas iniciativas poderiam ser adotadas de forma a adotar alguns sistemas e instrumentos já testados em outras realidades e que apresentam condições para serem adaptados ou para sugerir caminhos próprios que levem a melhorias da qualidade das instituições de educação infantil, onde mais de 8 milhões de crianças pequenas brasileiras passam parte de sua infância.

Desenhos de pesquisa que combinam a avaliação externa da qualidade de instituições de educação infantil, utilizando instrumentos padronizados, com observação e avaliação do desenvolvimento das crianças ali atendidas, preferentemente com o acompanhamento de sua evolução na continuidade de sua escolaridade, podem gerar avanços no conhecimento desses processos no contexto brasileiro. Isso posto, apenas a evidência de que a frequência à educação infantil está associada a melhores resultados no futuro não é suficiente para orientar aspectos importantes das políticas educacionais que demandam mudanças no país.

REFERÊNCIAS

BARROS, Ricardo P. et al. Uma avaliação do impacto da qualidade da creche no desenvolvimento infantil. **Pesquisa e Planejamento Econômico**. Brasília, v.41, n. 2, ago. 2011, p. 2013 – 232.

BERTRAM, A. D.; PASCAL, C. **Effective Early Learning (EEL)**: A handbook for evaluating, assuring and improving quality in settings for Three to Five Year Olds. Birmingham: Amber Publishing, 2004.

BERTRAM, A. D.; PASCAL, C. **Baby Effective Early Learning (BEEL)**: A handbook for evaluating, assuring and improving quality in settings for Birth to Three Year Olds. Birmingham: Amber Publishing, 2006

BHERING, Eliana et al. **Educação infantil**: políticas internacionais para crianças de 0 a 3 anos. (no prelo).

BONDIOLI, Anna M. Avaliação e cultura de qualidade nos serviços italianos para a infância: orientações e experiências dos anos 1990 aos dias de hoje. **Pro-Posições**. Campinas, v. 29, n. 2, mai./ago. 2018.

BRASIL. **Constituição da República Federativa do Brasil**. Brasília: Senado Federal, 1988.

BRASIL. **Lei de Diretrizes e Bases da Educação**. Brasília: Senado Federal, 1996.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, DF: MEC, 2016a. Disponível em: <http://portal.mec.gov.br>. Acesso em: 20 nov. 2019.

_____. Ministério da Educação. INEP. **Avaliação Nacional da Educação Infantil (ANEI)**. Brasília: MEC/INEP, 2016b. Disponível em: anei@inep.gov.br. Acesso em: março 2019.

_____. Ministério da Educação. Conselho Nacional de Educação. **Diretrizes Curriculares Nacionais para a Educação Infantil**. Brasília: MEC/CNE, 2009a.

_____. Ministério da Educação. **Indicadores da qualidade na educação infantil**. Brasília: MEC/SEB, 2009b.

_____. Ministério da Educação. **Referencial Curricular Nacional para a Educação Infantil**. Brasília: MEC/SEF, 1998.

_____. Ministério da Educação. **Critérios para um atendimento em creches que respeite os direitos fundamentais das crianças**. Brasília: MEC, 1995.

BRASIL. Ministério da Educação. **Monitoramento do uso dos Indicadores da qualidade na educação infantil**. Brasília: MEC/SEB, s./d. Disponível em: <<http://portal.mec.gov.br>> Acesso em: 08 mai. 2016.

BRASIL. **Marco Legal da Primeira Infância**. Lei no. 13.257 de 8 de março de 2016. Disponível em: <http://www.planalto.gov.br>. Acesso em 17 out. 2018.

CAMPOS, Maria M. Avaliação na primeira infância: contribuições para o debate. In: **Utilização de métodos e instrumentos padronizados de avaliação na primeira infância: convergências e divergências**. Fortaleza: SESC, 2016, p. 33–45.

_____. Entre as políticas de qualidade e a qualidade das práticas. **Cadernos de Pesquisa**. São Paulo, v. 43, n. 148, jan./abr. 2013, p. 22 – 43.

_____. Políticas educativas e avaliação para as primeiras idades no Brasil. In: Cardona, M. J.; Guimarães, Célia M. (coord.). **Avaliação na educação de infância**. Viseu, Portugal: Psicosoma, 2012, p. 61–76.

_____. A questão da creche: história de sua construção na cidade de São Paulo. **Revista Brasileira de Estudos Pedagógicos**. Brasília, v. 71, n. 169, set./dez. 1990, p. 212 – 231.

CAMPOS, Maria M. et al. A qualidade da educação infantil: um estudo em seis capitais brasileiras. **Cadernos de Pesquisa**. São Paulo, v. 41, n. 142, jan./abr. 2011, p. 20 – 54.

CAMPOS, Maria M.; RIBEIRO, Bruna. **Autoavaliação institucional participativa em unidades de educação infantil da rede municipal de São Paulo**. São Paulo: Fundação Carlos Chagas, 2016. (Textos FCC: relatórios técnicos, v. 48).

_____. **Autoavaliação institucional participativa em unidades de educação infantil da rede municipal de São Paulo – II**. São Paulo: Fundação Carlos Chagas, 2017. (Textos FCC: relatórios técnicos, v. 51).

CENPEC. (org.) **Razões e desafios do gestor da Educação**. Dez entrevistas sobre a prática da gestão educacional. São Paulo: Fundação Santillana/Moderna, 2018.

FOLQUE, Maria A. Desenvolver a qualidade em parcerias (DQP) – Um referencial com potencialidades múltiplas. **Cadernos de Educação de Infância**, n 95, 2012, p. 14–19. Disponível em: <space.uevora.pt>. Acesso em: dia. nov. 2019.

FORMOSINHO, Júlia et al. Ethical principles for holistic pedagogic evaluation. In: Formosinho, Júlia; Pascal, Christine. **Assessment and evaluation for transformation in Early Childhood**. Londres: EECERA/Routledge, 2016, p. 131-141.

GARIBOLDI, Antonio; MASELLI, Marina. Construindo a avaliação em conjunto: a experimentação de uma abordagem participativa para avaliação na região italiana da Emília Romanha. **Pro-Posições**. Campinas, v. 29, n 2, mai./ago. 2018. <http://scielo.br> Acesso em 24/09/2019.

HARMS, Thelma; CLIFFORD, Richard M; CRYER, Debby. **The early childhood environment rating scale: revised edition (ECERS – R)**. Nova York: Teachers College Press, 2005.

HARMS, Thelma; CRYER, Debby; CLIFFORD, Richard M. **Infant-Toddler environment rating scale: revised edition (ITERS - R)**. Nova York: Teachers College Press, 2006.

LAEVERS, Ferre. Fundamentos da educação experiencial: bem-estar e envolvimento na educação infantil. **Estudos em Avaliação Educacional**. São Paulo, v.25, n. 58, mai./ago. 2014, p. 152 – 185.

LESEMAN, Paul P. M. L'impact d'une offre d'éducation et d'accueil de qualité sur le développement des jeunes enfants. Synthèse de la recherche. In: EACEA. **L'éducation et l'accueil des jeunes enfants en Europe: réduire les inégalités sociales et culturelles**. Bruxelas: EACEA – Eurydice, Commission Européenne, 2009, p. 17–49.

MELHUIH, Edward. Efeitos de longo prazo da educação infantil: evidências e política. **Cadernos de Pesquisa**. São Paulo, v. 43, n. 148, jan./abr. 2013, p. 124 – 149.

MORO, Catarina. Avaliação de contexto e políticas públicas para a educação infantil. **Laplage em Revista**. Sorocaba, vol. 3, n. 1, jan./abr. 2017, p. 44 – 56.

MORO, Catarina; SOUZA, Sandra Z. Produção acadêmica brasileira sobre avaliação em educação infantil: primeiras aproximações. **Estudos em Avaliação Educacional**. São Paulo, v.25, n. 58, mai./ago. 2014, p. 100 – 125.

NCPI. Núcleo Ciência pela Infância. Comitê Científico. **O impacto do desenvolvimento na primeira infância sobre a aprendizagem**. São Paulo: NCPI, 2014.

OLIVEIRA-FORMOSINHO, Júlia. A avaliação da qualidade como garantia do impacto da provisão na educação de infância. In: Ministério da Educação. **Manual DQP - Desenvolvendo a Qualidade em Parcerias**. Portugal: Ministério

da Educação, s./d., p. 9–14. Disponível em: dge.mec.pt. Acesso em: dia? Nov. 2019.

OECD. Organisation for Economic Co-operation and Development. **Starting Strong IV: Monitoring quality in early childhood education and care**. Paris: OECD Publishing, 2015.

_____. Organisation for Economic Co-operation and Development. **Starting Strong III: A quality toolbox for early childhood education and care**. Paris: OECD Publishing, 2011.

_____. Organisation for Economic Co-operation and Development. **Starting Strong II: Early childhood education and care**. Paris: OECD Publishing, 2006.

PASCAL, Christine; BERTRAM, Tony. Participatory methods for assessment and evaluation. In: Formosinho, Júlia; Pascal, Christine. (ed.) **Assessment and evaluation for transformation in Early Childhood**. Londres: EECERA/Routledge, 2016, p. 74 – 92.

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

PORTUGAL, Gabriela; LAEVERS, Ferre. **Avaliação em Educação Pré-escolar**. Sistema de acompanhamento de crianças. Porto: Porto Editora, 2010.

RIBEIRO, Bruna. Avaliação da aprendizagem na educação infantil: um estudo exploratório em 125 municípios brasileiros. **Revista Linhas**. Florianópolis, v. 19, n. 40, , mai./ago. 2018, p. 218 – 245.

ROSEMBERG, Fúlvia. Políticas de educação infantil e avaliação. **Cadernos de Pesquisa**. São Paulo, v.43, n. 148, jan./abr. 2013, p. 44 – 75.

SAVIO, Donatella. “Promover a partir de dentro”: uma abordagem reflexiva e participativa da avaliação de contextos educativos. **Pro-Posições**. Campinas, v. 29, n. 2, mai./ago. 2018. (Dossiê: Desafios para a avaliação de contexto na educação infantil: Itália e Brasil)

SOUZA, Gizele de; MORO, Catarina; COUTINHO, Angela S. (orgs.) **Formação da rede em educação infantil: avaliação de contexto**. Curitiba: Appris, 2015.

SOUZA, Sandra Z.; PIMENTA, Claudia O. Avaliação e gestão da educação infantil em municípios brasileiros. **Educação & Realidade**. Porto Alegre, v. 43, n. 4, out./dez. 2018, p. 1277–1300.

SYLVA, Kathy et al. **Early Childhood Matters**. Evidence from the Effective Pre-school and Primary Education Project. Londres: Routledge, 2010.

TAYLER, Collette. Avaliação da qualidade da educação infantil na Austrália. **Estudos em Avaliação Educacional**. São Paulo, v.25, n. 58, mai./ago. 2014, p. 126 – 150.

TODOS pela Educação. **Anuário Brasileiro de Educação Básica 2019**. São Paulo: Moderna, 2019.

YOSHIKAWA, Hirokazu et al. **Investing in our future**: the evidence base on preschool education. Society for Research in Child Development/Foundation for children development, Out. 2013.

Recebido em: 10 de dezembro de 2019

Aceito em: 19 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÃO DA APRENDIZAGEM NA EDUCAÇÃO INFANTIL: ESTUDO DE CASO EM UM MUNICÍPIO BRASILEIRO

Bruna Ribeiro¹

¹ Doutoranda na Faculdade de Educação (FEUSP) da Universidade de São Paulo, São Paulo, Brasil. Assessora da frente de Educação Infantil da Secretaria Municipal de Educação de São Paulo e coordenadora do Grupo de Estudo e Trabalho sobre escuta de crianças (SME-SP). E-mail: brunaribeiro@usp.br

RESUMO

O presente artigo visa a socializar parte dos resultados advindos de um estudo de caso de caráter exploratório que teve por objetivo investigar as variáveis intervenientes no processo de avaliação da aprendizagem das crianças em um município brasileiro, com vistas a produzir insumos que contribuam para fomentar o debate na área da educação infantil. Os dados provenientes do estudo evidenciam nuances e apontam interfaces importantes a serem levadas em consideração, como currículo-avaliação e avaliação educação infantil-ensino fundamental, bem como problematiza políticas de responsabilização dos profissionais, apontando como perspectivas as políticas de participação. O estudo conclui evidenciando alguns desafios a serem superados e propõe caminhos a serem explorados com vistas à construção e consolidação de uma avaliação da aprendizagem na educação infantil alinhada às finalidades dessa etapa.

Palavras-chaves: Avaliação da aprendizagem. Educação infantil. Avaliação de crianças.

Abstract

This article aims to socialize part of the results from an exploratory case study that aimed to investigate the intervening variables in the process of assessing children's learning in a Brazilian municipality, with a view to producing inputs that contribute to foster debate in the field area of early childhood education. The data from the study show nuances and point to important interfaces to be taken into consideration, such as curriculum-evaluation and early childhood-elementary school evaluation, as well as problematizing policies of professional accountability, pointing out the perspectives of participation policies. The study concludes by highlighting some challenges to be overcome and proposes ways to be explored with a view to building and consolidating a learning assessment in early childhood education aligned with the purposes of this stage.

Keywords: Evaluation of learning. Early childhood education. Evaluation of children

Introdução

A educação infantil no Brasil, reconhecida como primeira etapa da Educação Básica pela Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), no que tange à temática da avaliação, vem buscando definir mais claramente seus contornos através da delimitação do campo avaliativo desta etapa.

Nesse contexto, a avaliação da aprendizagem das crianças é um tema que merece maior atenção e destaque, tanto de pesquisadores e estudiosos da temática da avaliação, como de instituições de educação infantil, Universidades Públicas, Secretarias e Conselhos Municipais de Educação que, embora com diferentes papéis e responsabilidades, podem (e devem) colaborar para uma avaliação que a reafirme a identidade da educação infantil e atue na perspectiva da garantia dos direitos das crianças.

Sousa e Pimenta (2018), ao realizarem importante mapeamento das produções acadêmicas brasileiras que tomaram a avaliação da educação infantil como objeto de estudo, evidenciaram ser ainda incipientes as iniciativas implementadas por municípios, pois, embora haja registros de proposições que abarcam aspectos como a infraestrutura física, recursos humanos e materiais, entre outros, essas avaliações ainda tendem a se restringir à aprendizagem das crianças, elegendo este indicador como expressão da qualidade da educação infantil.

Nesse artigo, nos interessa discorrer especificamente sobre a questão das iniciativas municipais de avaliação da criança, através do cruzamento de dados provenientes de um estudo de caso (YIN, 2001; AMADO, 2013), realizado por ocasião do Projeto UNESCO BRZ/1041 (Organização das Nações Unidas para a Educação, a Ciência e a Cultura) de apoio ao desenvolvimento de estratégias de implementação do Plano Nacional de Educação no tocante às políticas públicas de Educação Básica, com vistas a produzir insumos sobre avaliação da aprendizagem das crianças que fomentem o debate da área.

O estudo realizado por solicitação da Coordenação Geral da Educação Infantil (COEDI), pertencente à Secretaria de Educação Básica do Ministério da Educação (SEB/MEC), envolveu a análise da legislação e orientações legais vigentes a respeito da temática da avaliação na educação infantil, levantamento de estudos e pesquisas nacionais e internacionais sobre avaliação da aprendizagem, envio de *Survey* para 300 Conselheiros e Dirigentes Municipais de Educação, análise documental de instrumentos utilizados para avaliação das crianças na Educação Infantil e realização de um estudo de caso sobre as avaliações das crianças em curso em um município brasileiro.

Parte dos resultados dessa pesquisa foram socializados em trabalho publicado em 2018 (RIBEIRO, 2018). Entretanto, devido aos limites da publicação, não foi possível discorrer sobre o estudo de caso.

Nesse sentido, nos interessa neste artigo socializar e refletir sobre os dados advindos do estudo de caso realizado, mas para esse intuito é importante retomarmos, ainda que de forma breve, os principais achados do estudo anterior (RIBEIRO, 2016; 2018), a fim de que possamos situar e ler os dados do estudo de caso à luz do contexto macro da pesquisa.

Buscando verificar as tendências e repercussões da temática da avaliação na primeira etapa da Educação Básica na produção científica nacional, sobretudo em relação às iniciativas municipais de avaliação da aprendizagem das crianças, foi realizado um levantamento de estudos e pesquisas nacionais sobre avaliação da aprendizagem, onde se constatou escassez de materiais acadêmicos e de literatura geral nacional que abordem a avaliação sob a perspectiva específica da educação infantil e sua quase invisibilidade quando se trata de estudos mais amplos que realizam balanço sobre avaliação na Educação Básica (RIBEIRO, 2018).

O estudo buscou, ainda, levantar dados que auxiliassem na compreensão das ações em curso relacionadas à avaliação das crianças nos diferentes municípios brasileiros, através da distribuição¹ de um *Survey* para 300 conselheiros e dirigentes municipais de educação. Os 206 questionários recolhidos eram representativos de 125 municípios, distribuídos em 22 estados situados nas cinco regiões brasileiras.

A análise² dos dados dos questionários evidenciou, entre outras coisas, que as crianças egressas da educação infantil estavam sendo avaliadas através de uma concepção estreita de currículo e aprendizagem, restringindo-os à mera aquisição de competências e habilidades de português e matemática, de forma fragmentada e descontextualizada.

Com base nos questionários, foi possível recolher 23 instrumentos de avaliação da aprendizagem utilizados na educação infantil que foram analisados à luz da legislação e orientações legais vigentes, com base em análise metodológica já descrita em estudo³ anterior, sendo possível, então, verificar que

¹ Os questionários foram distribuídos durante a realização de um Encontro Nacional dos Conselheiros Municipais de Educação (UNCME) realizado em Pernambuco em 2015.

² Ver: RIBEIRO, B. Avaliação da aprendizagem na educação infantil: um estudo exploratório em 125 municípios brasileiros. **Revista Linhas**. Florianópolis, v.19, n.40, p.218-245, maio/ago, 2018.

³ *Ibidem*.

as crianças dos municípios analisados estavam sendo avaliadas na educação infantil tomando-se por referência uma noção de qualidade que não condiz com todo o arcabouço legal da área.

Visando melhor investigar os fenômenos descritos anteriormente, buscou-se realizar um estudo de caso de caráter exploratório objetivando compreender as interfaces entre as avaliações realizadas na educação infantil e ensino fundamental através da análise documental dos instrumentos avaliativos utilizados e da escuta dos sujeitos envolvidos no processo avaliativo.

Esse estudo se justifica, uma vez que, conforme já afirmado por Sousa e Pimenta (2018), a avaliação cumpre papel indutor de políticas e práticas, mas isso não se realiza sem consequências. Quais são as consequências para as crianças e os profissionais da educação infantil das escolhas das avaliações municipais? Longe de buscar a resposta final para questão tão ampla e complexa, este estudo almeja apenas lançar luz sobre a temática das avaliações municipais que estão focando na aprendizagem da criança através de um exemplo que elucidar os desafios a serem enfrentados e superados na constituição de uma avaliação condizente e coerente com os fins da educação infantil, conforme preconizado “pelo arcabouço legal vigente, as quais se ancoram em subsídios da literatura nacional e estrangeira relativa a esse campo de estudo” (SOUSA; PIMENTA, 2018, p.1296).

METODOLOGIA

Dentre os 125 municípios investigados através do *Survey*, um deles foi escolhido para realização de um estudo de caso (YIN, 2001) de caráter exploratório (AMADO, 2013), visando, assim, obter mais informações sobre a avaliação no interior de um município, através da compreensão sobre como as diferentes profissionais⁴ envolvidas no processo (professoras de todas as faixas etárias atendidas na educação infantil e coordenadora geral da educação infantil do município) percebem e se relacionam com o fenômeno da avaliação.

Os critérios utilizados para definição do município obedeceram às seguintes condições: ter participado da etapa anterior do estudo (*Survey*), possuir sistema municipal próprio, realizar avaliações na educação infantil e ensino fundamental há mais de 2 anos e aderir livremente à pesquisa, colaborando, assim, com a entrada no campo e a coleta de informações no município.

⁴Optamos pelo uso do feminino ao se referir as profissionais participantes do estudo, uma vez que todos os sujeitos escutados eram mulheres.

Não foi pretensão deste estudo obter dados visando a generalizações e sim circunscrever o objeto investigado, buscando elementos que auxiliassem na reflexão mais ampla sobre as possíveis variáveis intervenientes no processo de avaliação municipal da aprendizagem das crianças, contribuindo assim para o objetivo⁵ principal da pesquisa que era produzir insumos que fomentem o debate na área.

Com esse intuito, foram delineados os seguintes procedimentos metodológicos:

- Caracterização do município e da política pública de educação infantil, com base em fontes secundárias.
- Análise documental dos instrumentos destinado à avaliação da aprendizagem das crianças matriculadas na educação infantil do município e instrumentos destinados à avaliação da aprendizagem das crianças matriculadas no primeiro ano do ensino fundamental.
- Entrevista em profundidade (AMADO, 2013) com a coordenadora geral da equipe técnica de educação infantil do município.
- Grupo focal/*Focus Group Studies* (MORGAN, 1998; 2004) com professoras da rede municipal de educação infantil.

Os principais resultados provenientes da metodologia adotada serão discutidos na seção a seguir.

RESULTADOS: O MUNICÍPIO, A REDE E A AVALIAÇÃO

O município selecionado para o estudo de caso localiza-se na Região Metropolitana de Salvador e possui uma área territorial de 759,8 km², sendo 42 km de faixa costeira. De acordo com dados do IBGE (2010), a população urbana é de 231.973 habitantes e a rural e costeira somam 10.997, o que o coloca como o quarto município mais populoso da Bahia.

O município é sede do maior complexo industrial integrado do hemisfério sul e é considerado o município mais industrializado do Estado, respondendo sozinho por 35% das exportações e concentrando o maior PIB industrial do Nordeste.

Apesar do PIB elevado, o município tem 67% de sua população em situação de pobreza e indigência. O número de pessoas em situação de pobreza (46,7%) é muito superior, se comparado à média no Nordeste brasileiro, que

⁵ De acordo com edital 01/2015, Projeto 914 UNESCO/1009.4, consultor 13.

é de 24,1%, e mais alta ainda, se comparada à média da população nacional, que é de 12,9%.

A Secretaria Municipal de Educação (SME) era centralizada, e a equipe de educação infantil era composta, na época da coleta de dados, por sete profissionais cujo foco principal de atuação era, segundo informado pela própria equipe, a formação continuada das profissionais e a realização de visitas às unidades educacionais.

A rede de educação infantil do município era composta por 156 unidades educacionais, sendo 110 diretas e 46 denominadas comunitárias (conveniadas com o poder público municipal), perfazendo um total de 4.559 crianças matriculadas nessa etapa.

Em relação à avaliação da aprendizagem das crianças, foram elaboradas orientações pela equipe técnica de educação infantil, no início de 2013, e realizada revisão dessas orientações no decorrer de 2015, visando a sua atualização.

As orientações se distinguem de acordo com a faixa etária da criança, sendo organizadas em quatro tipos de orientações destinadas a subsidiar a observação, o registro e avaliação das professoras em relação aos seguintes agrupamentos:

- Orientação de avaliação para grupo inicial e grupo 1 (crianças até 2 anos).
- Orientação de avaliação para grupos II e III (crianças de 2 e 3 anos).
- Orientação de avaliação para grupos IV e V (crianças de 4 e 5 anos).
- Orientação de avaliação grupo V (registro final das crianças de 5 anos)⁶.

Esses registros devem ser realizados na Caderneta Pedagógica, também denominado Diário Pedagógico, que possui um campo específico para inserção dos dados dessa avaliação. A caderneta é “um documento Oficial que deve ser mantido em local apropriado, de forma que assegure a sua inviolabilidade, não podendo ser retirado da Instituição de Ensino sem a devida autorização da Equipe Gestora” (CEAP, 2015, p.5).

O registro é realizado trimestralmente pelo professor, nos meses de março, julho e dezembro. Além do campo para preenchimento da avaliação, o documento serve também para controle da frequência da criança e do educador, que deve indicar os dias nos quais faltou e registrar os “temas das aulas”.

⁶ O grupo V era mencionado em duas orientações, sendo que uma delas destinava-se a orientar exclusivamente sobre a avaliação final.

Vale ressaltar, ainda, que o texto introdutório da Caderneta de todas as faixas etárias esclarece aos profissionais que a proposta de avaliação das crianças matriculadas na educação infantil no município deve se dar através do “registro do desempenho do educando através de relatórios descritivos que deverão estar pautados nas habilidades constantes nas orientações curriculares municipais para educação infantil” (CEAP, 2015, p.5).

As habilidades mencionadas no documento referem-se ao que eles denominam “indicadores de aprendizagem na educação infantil” que apresentam características que, segundo o documento (CEAP, 2015), as crianças, de modo geral, apresentam em cada faixa etária. Com base nas características de cada faixa etária explicitadas nas cadernetas, são apresentados itens para observação e registro do professor, conforme foram transcritos e sistematizados no quadro 1, a seguir:

Quadro 1: Itens para observação e registro das professoras das crianças de 0 a 5 anos.

GRUPO INICIAL (crianças até 2 anos)	GRUPO II E III (crianças de 2 e 3 anos)	GRUPO IV E V (crianças de 4 e 5 anos)	GRUPO V FINAL (crianças de 5 anos)
1. Adaptação/ desfralde/ brincadeira interação/ comunicação	1. Adaptação/ desfralde/ brincadeira/interação/ participação	1. Relação consigo e com os outros: brincadeira/ interação/ participação/ autonomia	Síntese de uma página com o registro final do acompanhamento do desenvolvimento/ aprendizagem da criança, contendo ainda espaço para registro da criança e de sua família.
2. Linguagem corporal: como se movimenta	2. Linguagem corporal: movimento, desenvolvimento físico	2. Linguagem corporal: expressão corporal, movimento, desenvolvimento físico	
3. Linguagem oral: como expressa o que quer e sente	3. Linguagem oral: fala espontânea, expressão inicial do que precisa e sente. Reconto inicial de histórias, repetição de textos que sabe de cor.	3. Linguagem oral: expressão das necessidades, sentimentos, ideias e preferências. Narração, reconto de histórias, fatos, repetição de textos que sabe de cor.	
4. Outras observações	4. Interesse e manuseio de materiais gráficos.	4. Leitura: interesse e manuseio de materiais gráficos, estratégias de leitura.	
	5. Linguagem plástica- iniciação no manuseio de lápis de cor, tinta, modelagem.	5. Linguagem escrita	
	6. Participação em brincadeiras que envolvam elementos da linguagem matemática.	6. Linguagem plástica: expressão pelo desenho, pintura, escultura, colagem, produção de objetos.	

Fonte: CEAP, 2015. Elaborado pela autora, 2019.

Conforme podemos constatar no quadro 1, é possível identificar que os grandes temas para observação, registro e avaliação das crianças dialogam, em certa medida, com o proposto nas Diretrizes Curriculares Nacionais para Educação Infantil/DCNEI (BRASIL, 2010) e com o currículo municipal, ao trazer os dois grandes eixos da educação infantil, entendidos como as brincadeiras e as interações, e ainda propor o acompanhamento e o registro do progresso das crianças em relação a importantes temas, como a adaptação e desfralde dos bebês e o trabalho com as múltiplas linguagens (corporal, oral, plástica, leitura, escrita.). No entanto, é possível observar também a fragmentação dessas linguagens que, ao serem tratadas de modo desarticulado e descontextualizado, em conteúdos estanques e, muitas vezes desprovidos de sentido, fragilizam não só a avaliação, mas principalmente as vivências cotidianas das crianças, que passam a ser pautadas por um rol prescritivo de ações e atividades que não dialogam com suas vidas, histórias, singularidades. Enfim, não dialogam com as crianças reais e concretas.

É importante ainda refletir e problematizar sobre a opção do município em dividir as competências esperadas das crianças por faixa etária, partindo da premissa de uma construção linear do conhecimento, em que as crianças somente podem ter acesso a determinado tipo de conhecimento quando completam determinada idade (como no exemplo da iniciação ao lápis de cor que, segundo o documento, deve ocorrer somente a partir dos 2 anos).

Sobre essa concepção de criança e conhecimento, é sempre bom lembrar os inúmeros estudos (VYGOTSKY, 1995; MELLO, 2015) que já apontaram para a necessidade de superação das formas de se tratar as crianças através das etapas de desenvolvimento, enfatizando que “as potencialidades dos bebês e das crianças se ampliam à medida que vivem experiências que sejam organizadas e apoiadas por educadoras e educadores e ao mesmo tempo sejam desafiadoras das ações das possibilidades das crianças” (SÃO PAULO, 2015, p.16). Ou seja, a criança avança “à medida que se criam percursos de aprendizagem significativos e socialmente relevantes que se somam desde a educação Infantil até o ensino fundamental” (SÃO PAULO, 2015, p.16).

E as professoras, o que pensam e como fazem uso desses instrumentos? Para investigar essa questão, foram realizados grupos focais que escutaram ao todo 24 professoras (que atuavam nas distintas faixas etárias atendidas na educação infantil), divididas em 3 grupos com 8 profissionais cada. Os debates nos grupos tiveram como objetivo compreender o que pensavam e sentiam as profissionais da educação infantil em relação à temática da avaliação da aprendizagem das crianças e ainda como percebiam e utilizavam os ins-

trumentos avaliativos propostos pela SME. As conversas transitaram em torno de questões relativas às funções, aos usos, à frequência e ao conteúdo dos instrumentos.

A escolha dessa estratégia se mostrou relevante, uma vez que, conforme já apontado por Amado (2013), o corpo de dados proveniente dos grupos focais possibilitou uma validade ecológica que não se encontra nos questionários, com informações forjadas na vida cotidiana dos participantes e expressas na própria linguagem das respondentes e do seu contexto.

O quadro 2, a seguir, sintetiza as principais contribuições advindas das discussões nos grupos focais, onde as falas foram agrupadas em categorias analíticas que permitem uma melhor visualização do conteúdo debatido pelas participantes. A análise do discurso (BARDIN, 2011) permitiu que as falas fossem agrupadas por temáticas que deram origem a sete categorias de análise, agrupadas em quatro dimensões.

Quadro 2: Quadro-síntese das categorias temáticas elaboradas a partir dos Grupos Focais de professoras

DIMENSÃO INVESTIGADA NO GF	CATEGORIAS ANALÍTICAS
Função da avaliação	Avaliação como sinônimo de ação burocrática
Frequência da avaliação	Frequência excessiva
Conteúdo da avaliação	Conteúdo padrão
	Competências não desenvolvidas
Interfaces avaliação da EI e EF	Distinções entre avaliações de diferentes etapas
	Influência no cotidiano, no currículo e na avaliação
	Acordos informais entre profissionais

Fonte: A autora, 2019.

Em relação à **função** da avaliação das crianças através do instrumento proposto pela secretaria, as profissionais escutadas foram unânimes em situar a “avaliação como sinônimo de ação burocrática”, afirmando que a ação de registrar e avaliar as aprendizagens das crianças realizada por elas estava vinculada à necessidade de terceiros (seja coordenadora pedagógica, equipe SME ou famílias). Em nenhum momento, foi relatado o uso das avaliações/registros para fomentar o planejamento, visitar práticas, conhecer e auxiliar as crianças. As falas revelaram apenas a percepção de que a avaliação é um dispositivo meramente burocrático para dar mais trabalho para o professor. Apareceram, ainda, analogias se referindo ao ato de avaliar como ‘tortura’ imposta a elas. As transcrições a seguir ilustram as falas que foram agrupadas nessa categoria.

Não tem função, é uma chatice, não serve para nada! É para pegar o professor em falha! (professora grupo focal A).

Tomam grande tempo, mais não ajuda é mais uma função burocrática para dar mais trabalho para o professor! (professora grupo focal B).

Acho que é tortura mesmo! Em 2014 teve uma “princesa Isabel” que nos absolveu de fazer isso por um tempo! (professora grupo focal C).

No que diz respeito à **frequência** da atividade de registro e avaliação das crianças, cuja proposta da SME era trimestral (março, julho e dezembro), as profissionais revelaram descontentamento, afirmando ser esta uma frequência excessiva e desnecessária que as colocavam em uma situação em “que nem tinham o que dizer”, uma vez que pouco se tinha para acrescentar de um relatório para outro.

Algumas participantes sugeriram, ainda, que a frequência fosse alterada para apenas dois registros anuais, sendo um no início e o outro no fim do ano.

Oito itens para preenchimento trimestral, dá quase 300 escritas, é absurdo! Faz as mesmas perguntas do mesmo tópico, fica difícil desenvolver, nem tem o que falar de diferente! (professora grupo focal B).

A discussão sobre o **conteúdo** do instrumento gerou a criação de duas categorias: “conteúdo padrão” e “competências não desenvolvidas”. Ambas as categorias guardam relação com a percepção das participantes em relação à dimensão frequência, pois as estratégias por elas narradas para definição do conteúdo estão relacionadas, em parte, com sua percepção de que a frequência da avaliação é excessiva.

Nesse sentido, as falas agrupadas na categoria “conteúdo padrão” revelam que, diante do fato de que têm pouco o que acrescentar de um relatório para outro, as professoras utilizam como estratégia a troca de nomes das crianças, repetindo o mesmo conteúdo para crianças diferentes, ou então não mudam os nomes, e sim as frases, fazendo “montagens” em que misturam aleatoriamente frases de diversos relatórios.

A gente só trocava o nome [das crianças] nos relatórios! (professora grupo focal A).

Eu pegava um pouquinho daqui...um pouquinho da li... [se referindo as frases de diferentes relatórios] (professora grupo focal C).

Já a categoria “Competências não desenvolvidas” revela uma percepção das participantes em relação à avaliação como um instrumento que deve apenas

focar no que a criança ainda não sabe ou não consegue fazer. Justificam, assim, a necessidade de diminuição da frequência dos registros sob o argumento de que não é preciso olhar para uma “competência que a criança já tem”.

Se ela [criança] já tem habilidade em março não preciso mais tocar no assunto (professora grupo focal B).

Por fim, foi necessário criar a dimensão **Interfaces avaliação da educação infantil e ensino fundamental**, contendo três categorias, para dar conta da temática da avaliação realizada no primeiro trimestre do ano com as crianças egressas da educação infantil, que foi um tema recorrente e reincidente nos 3 grupos focais.

Na categoria denominada “distinções entre avaliações das diferentes etapas”, foram agrupadas as falas que expressam a percepção de que a avaliação realizada na educação infantil tem um “peso” menor do que a realizada no ensino fundamental, sendo a primeira mais uma ação “pró-forma”, sem grandes consequências, enquanto a segunda é a que realmente gera impactos, ações e visibilidade pública e política.

Na educação infantil a avaliação é mais burocrática, fazemos por fazer, porque tem que cumprir, já no fundamental tem maior visibilidade pública e política (professora grupo focal A).

A categoria “influência no cotidiano, no currículo e na avaliação” expressa a percepção das professoras de que a avaliação realizada no 1º ano do ensino fundamental tem gerado impactos negativos na educação infantil, reorientando o currículo, as práticas e vivências cotidianas e a própria avaliação das profissionais que passam a pautar sua observação nos itens das provas do 1º ano.

O pessoal da SME [equipe EI] é comprometido, mas não tem jeito...a gente olha para a avaliação do Fundamental para saber o que dar na educação infantil porque depois caem matando na gente, dizendo que educação infantil não ensina nada, só brinca. As crianças vão mal na prova [do 1º ano] e a educação infantil é quem paga o pato! (professora grupo focal C).

A política de bonificação das professoras do 1º ano, cujos alunos se saíssem bem na prova, foi implementada no município como uma estratégia “piloto”, buscando ver se a medida surtia efeito positivo nas avaliações das crianças. A lógica da gestão meritocrática de responsabilizar quase que exclusivamente os professores pelos resultados produz efeitos na contramão da lógica pedagógica, como já discorrido por Machado e Alavarse (2015, p.74):

Em tal perspectiva, as avaliações externas, além da associação mecânica entre desempenho em provas e trabalho docente, desprezando, frequentemente, as condições das quais emergem esses resultados, confundem-se com um modelo de gerenciamento de recursos humanos, retirando-lhes todo o potencial pedagógico.

No município investigado não foi diferente. Segundo relatos das participantes dos grupos focais, a prova do 1º ano gerou uma série de “acordos informais entre profissionais” de ambas as etapas que, visando corresponder à expectativa do resultado da avaliação e assim receber o bônus, realizavam “acordos informais”, em que o dinheiro era dividido entre professoras da educação infantil e do ensino fundamental, caso as primeiras “auxiliassem” trabalhando com as crianças na educação infantil o conteúdo da prova do primeiro ano.

Uma professora do 1º ano ofereceu ‘rachar’ o bônus comigo se eu mandasse as crianças da minha turma [5 anos] já pré-silábicas (professora grupo focal B).

Mas do que se tratava a avaliação do ensino fundamental?

Parte da metodologia proposta para o estudo de caso contemplava a análise do instrumento de avaliação das crianças matriculadas no 1º ano do ensino fundamental, com vistas a melhor compreender se a política de avaliação do município atuava na perspectiva de um *continuum* (SILVA; FLORES, 2015), como o proposto pelo Parecer CNE/CEB nº 4/2008 que, em seu item 9.1, afirma que a avaliação nos três primeiros anos do ensino fundamental deverá ser processual, participativa, formativa, cumulativa e diagnóstica (BRASIL, 2008), bem como o disposto nas Diretrizes Curriculares Nacionais da Educação Básica (BRASIL, 2013) que prevê que a Educação Básica se organize, levando-se em consideração as dimensões da organicidade, sequencialidade e articulação, sendo necessário, para isso, que as diferentes etapas compartilhem das mesmas concepções e princípios.

A equipe técnica da secretaria municipal de educação do município em questão passou a realizar a avaliação da aprendizagem das crianças matriculadas no 1º ano da rede municipal através da aplicação de uma prova, a partir do ano de 2014.

A prova era aplicada no primeiro trimestre do ano e tinha por objetivo aferir os conhecimentos de português e matemática das crianças egressas da educação infantil. Para isso, era organizada em oito páginas que continham 10 questões de português e 10 de matemática, respectivamente.

No início da prova, eram apresentadas as duas matrizes de referência que indicavam os eixos trabalhados, o tema a que se referia cada questão e o descritor, explicitando, assim, o que era esperado das crianças.

As informações dos quadros 3 e 4, a seguir, foram transcritas na íntegra das provas analisadas.

Quadro 3: Matriz de referência de português

eixo	Tema	questão	Descritor
Apropriação do sistema de escrita	Reconhecimento de letra	1	D1-Identificar as letras do alfabeto
		2	D2-Diferenciar letras de números e de outros símbolos
		3	D3-Identificar letra inicial
		4	D4- Conhecer a sequência alfabética
	Desenvolvimento da consciência fonológica	5	D5- identificar o número de sílabas que formam uma frase
Leitura	Leitura de palavras	6	D6-Ler palavras em padrão canônico e não canônico
Escrita	Escrita do nome	7	D7-Escrever o próprio nome
	Escrita de palavras, frases e textos	8	D8-Escrever palavras
		9	D9-Escrever frases
		10	D10-Escrita de texto de memória

Fonte: Matriz de Referência Avaliação Municipal, 2017. Elaborado pela autora, 2019.

Quadro 4: Matriz de referência de matemática

Eixo	Tema	Questão	Descritor
1.Espaço e forma	Localização espacial	1	D1- Identificar e localizar objetos no espaço
	Representações de figuras geométricas	2	D2- Reconhecer as representações de figuras geométricas espaciais
Grandezas e medidas	Identificação, comparação, relação e ordenação de medidas e grandezas	3	D3- Identificar e comparar objetos, tomando como atributo unidades de medidas (comprimento, massa, capacidade)
		4	D4- Reconhecer cédulas e moedas do Sistema Monetário Brasileiro e suas possíveis trocas entre cédulas e moedas em função dos seus valores
		5	D5- Identificar e relacionar medidas de tempo
Números e operações	Construção do significado do número e suas representações	6	D6- Associar a denominação do número a sua respectiva representação simbólica
		7	D7- Relacionar objetos e coleções e representar quantidades
	Resolução de problemas por meio da adição e subtração	8	D8- Resolver problemas com ações de tirar ou diminuir quantidades
		9	D9- Resolver problemas com ações de juntar ou acrescentar unidades
Tratamento da informação	Leitura e interpretação de dados em gráficos e tabelas	10	D10- Identificar informações apresentadas em gráficos e tabelas

Fonte: Matriz de Referência Avaliação Municipal, 2017. Elaborado pela autora, 2019.

Como foi possível constatar pelas informações dos quadros anteriores, espera-se que as crianças, em relação à língua portuguesa, saibam identificar as letras e a sequência do alfabeto, identificar o número de sílabas em uma frase, escrever seus nomes, palavras, frases e até um texto de memória.

Já em relação à matemática, é esperado que as crianças saibam identificar cédulas do sistema monetário, analisar gráficos, resolver problemas e representar quantidades numericamente, entre outras habilidades.

Esse modelo de avaliação vem causando grande desconforto nas profissionais que atuam na educação infantil, conforme evidenciado nos grupos focais e, inclusive, na coordenadora geral da educação infantil municipal que, em entrevista, informou que essa prova gerou muita polêmica, não só entre as profissionais da rede, de ambas as etapas, mas também entre as técnicas da secretaria responsáveis pela educação infantil e ensino fundamental, que divergem em relação às expectativas sobre os conhecimentos e experiências esperados que as crianças tenham acesso na educação infantil.

Porque em última instância quem está sendo avaliada é a etapa da educação infantil, uma vez que a prova ocorre até março... São as profissionais da educação é que estão sendo implicitamente avaliadas, suas práticas, nosso currículo. (...) Não fomos consultadas, não fomos chamadas para discutir e pensar junto, não fomos sequer informadas com antecedência (Coordenadora Geral da Educação Infantil Municipal).

Essa correlação de forças e de disputa explicitada na fala da coordenadora geral evidencia um movimento em andamento no município em que as avaliações do ensino fundamental muito rapidamente passam a capturar e pautar o currículo e as práticas da educação infantil.

Essa situação já era evidenciada na análise dos questionários respondidos pelos conselhos e dirigentes municipais e, no município em questão, foi confirmada pela coordenadora geral e pelas participantes dos três grupos focais.

A coordenadora afirmou, ainda, que tem observado uma crescente preocupação das profissionais da educação infantil em alfabetizar, focando suas práticas quase que exclusivamente nas atividades da linguagem escrita, em detrimento das demais linguagens, uma vez que, como elucida:

só consegue realizar a avaliação [do 1º ano] quem está alfabetizado, o que gera muita ansiedade para as professoras do grupo 5, pois mesmo recebendo orientação da equipe de Educação Infantil para não ter a preocupação de alfabetizar, é contraditório esperar que a criança realize uma prova dessas. (Coordenadora Geral da Educação Infantil Municipal).

Avaliações como as propostas na educação infantil e, principalmente, no ensino fundamental do município evidenciam que as especificidades de cada etapa estão sendo levadas em consideração, nem os aspectos compartilhados por ambas, na formulação e proposição de políticas avaliativas. Situação semelhante já havia sido identificada por Silva e Flores (2015, p.194), ao realizarem uma análise das políticas públicas educacionais brasileiras.

Ademais, é importante frisar que, além de a avaliação do ensino fundamental partir de uma concepção redutora do processo de alfabetização, esta ainda corrobora com “a equivocada visão, ainda presente em redes de ensino e escolas, de que a Educação Infantil se presta à preparação das crianças para o ingresso no Ensino Fundamental (...)” (SOUSA, 2018, p.76).

Essa questão é abordada por diversos autores (SOUSA, 2018; RIBEIRO, 2016; 2018) que evidenciam que “iniciativas de avaliação nos anos iniciais do ensino fundamental, direcionadas a julgar o nível de alfabetização das crianças, podem influenciar os planos de ensino e, em especial, as práticas de avaliação da aprendizagem na Educação Infantil” (SOUSA, 2018, p.76).

Nesse sentido, concordamos com Machado e Alavarse (2015) que, ao se apoiarem em Freitas (2011), defendem que somente quando os resultados das avaliações puderem “contribuir com a construção de alternativas pedagógicas para as políticas e as escolas cumprirem suas funções junto à sociedade democrática, no sentido de oferecer educação pública de qualidade para todos seus alunos e alunas” e não com a responsabilização que busca punir ou premiar, é que “que vamos criar condições para ampliar a qualidade do ensino” (MACHADO; ALAVARSE, 2015, p.76).

Iniciativas de responsabilização que não incluem as diferentes instâncias do sistema educacional são, como nos alertam Sousa e Pimenta (2018), um desvirtuamento dos processos de *accountability*.

Portanto, não se trata de não realizar avaliações externas e de larga escala, mas sim de colocá-las a serviço da aprendizagem das crianças, de forma a apoiar o profissional nesse processo, tendo clareza, entretanto, de que

uma avaliação em larga escala, que tenha como foco a avaliação da criança em um contexto precário e desigual de oferta da educação infantil, apoia diagnósticos equivocados e leva à prescrição de “remédios” que, certamente, não terão efeito, ou pior, poderão produzir graves efeitos colaterais (SOUSA, 2018, p.74).

Esses “efeitos colaterais” citados pela autora já podem ser sentidos tanto pelas crianças da educação infantil, ao terem seu cotidiano ditado pelos conteúdos da avaliação do 1º ano, quanto pelas crianças do ensino fundamental, que passam a ser “julgadas” e rotuladas por um instrumento técnico e pedagogicamente frágil, e também pelas profissionais de ambas as etapas, que têm sua autonomia subtraída e passam a pautar/nortear sua práxis em medidas externas e não, contraditoriamente, na busca pelo desenvolvimento da aprendizagem das crianças.

Assim, é urgente a reflexão sobre a natureza das avaliações às quais as crianças são submetidas, pois como dito, “a forma de fazer a avaliação pode falsear dados e induzir as redes municipais não a acertos, mas, sim, a equívocos (SOUSA; PIMENTA; MACHADO, 2012, p. 34).

É imperativo que os processos avaliativos de aprendizagem das crianças não estejam descolados da análise do complexo contexto que envolve determinações intra e extraescolares, refletindo sobre a ampla gama de fatores que incidem na qualidade da educação infantil e forjam as condições objetivas de aprendizagem.

Em entrevista, a coordenadora de educação infantil do município citou que algumas questões que vem enfrentando, como uma equipe reduzida, excesso de demandas e disputas internas no interior da secretaria interferem e fragilizam o desenho e a implementação de ações voltadas para a educação infantil. A profissional afirmou ainda que tem consciência de que a proposta de avaliação na educação infantil tem muitos desafios e está longe de se alinhar aos anseios de sua equipe e do postulado nas DCNEI (BRASIL, 2010). No entanto, sair desse lugar “à sombra do ensino fundamental”, como ela denomina, exige a construção de uma rede de apoio com sujeitos de diferentes instâncias implicados em “apoiar a construção de avaliações potencialmente capazes de servir a um projeto educacional que dê materialidade ao direito à educação de qualidade às crianças pequenas” (SOUSA; PIMENTA, 2018, p.1296).

A fala da coordenadora, transcrita a seguir, pode ser lida como uma síntese que evidencia que a construção de uma avaliação coerente com as finalidades da educação infantil demanda ações que envolvem diferentes atores e instâncias implicados na consecução de políticas avaliativas respaldadas em um projeto de educação e de sociedade democrática.

O Acompanhamento e avaliação das crianças da educação infantil têm sido um dos maiores desafios da nossa rede municipal. Há alguns anos estamos discutindo insistentemente, defendendo, em meio a um campo

duro de disputas cotidianas, as concepções de criança e de infância que acreditamos. Temos pouquíssimas pessoas perto de nós que podem nos apoiar nestas discussões. Com nossa pequena equipe técnica da educação infantil, já conseguimos conquistas muito valiosas, mas, no cotidiano, este é um processo solitário. Somos nós que temos que propor e defender perante pessoas com muito maior poder de decisão do que nós. Algumas chefias ainda não compreendem que esta é uma discussão nacional. Continuamos acreditando, resistindo e insistindo. Precisamos de ajuda para continuar a construção de formas de acompanhamento e avaliação cada vez mais próximas do que é bom para a criança. (...), portanto, reite-ramos, precisamos de ajuda, a pressão é muito forte. Só sustentamos com ajuda da Teia! (Coordenadora da Educação Infantil da Secretaria Municipal de Educação).

Nesse cenário de pouco diálogo e de pouca construção coletiva no âmbito da Secretaria Municipal de Educação, os discursos, as falas e as práticas das profissionais parecem refletir as contradições, fragilidades e desafios de articulação no âmbito da política municipal.

CONSIDERAÇÕES FINAIS

Com base nos dados provenientes do estudo de caso realizado em um município brasileiro, podemos tecer algumas considerações, relacionando-as com os dados macro da pesquisa, que nos ajudam a iluminar a temática da avaliação da aprendizagem das crianças na etapa da educação infantil.

Os elementos provenientes do estudo de caso evidenciam fragilidades das propostas avaliativas municipais referentes à avaliação da aprendizagem das crianças, tanto da educação infantil como do ensino fundamental, no universo pesquisado.

Nesse sentido, os dados ratificam as evidências obtidas por meio do *Survey* que indicam que as crianças egressas da educação infantil estão sendo avaliadas tomando-se por referência uma concepção estreita de currículo e aprendizagem, focada em apenas duas áreas de conhecimento (português e matemática) em que se toma, equivocadamente, a avaliação como sinônimo de medida.

O estudo de caso traz indícios, ainda, de que essa avaliação tende a impactar negativamente tanto as práticas cotidianas quanto a avaliação na educação infantil, que passa a ser pautada pelo conteúdo das avaliações do primeiro ano do ensino fundamental.

A análise dos instrumentos avaliativos utilizados pelo município para avaliar a aprendizagem das crianças de 0 a 6 anos revelou o prevalectimento de:

avaliações padronizadas que desconsideram a singularidade de cada criança, seu percurso de aprendizagem individual e grupal e os condicionantes desse processo, mostrando-se, assim, insuficientes não só como instrumento de documentação e avaliação do percurso, mas também como procedimento que busca favorecer a aprendizagem. (RIBEIRO, 2018, p.241).

Apesar da avaliação da aprendizagem na educação infantil, no município investigado, ser realizada por um instrumento de observação e registro, através de relatórios descritivos individuais das crianças, prática essa alinhada ao apregoado pelas orientações normativas federais, foram muitas as fragilidades identificadas no instrumento (fragmentação do conteúdo, concepção linear de conhecimento, concepção de criança pautada em fases estanques, ausência de temáticas fundamentais, entre outros).

No entanto, outras questões emergiram para além daquelas já identificadas com o uso de outros instrumentos e metodologias (RIBEIRO, 2018), trazendo à tona importantes nuances e questões que contribuem para dar pistas a respeito de investigações futuras que visam a contribuir com a produção de conhecimento relativo à avaliação da aprendizagem das crianças atendidas na educação infantil.

Questões que alertam sobre a importância do debate da avaliação na educação infantil não ser feito de forma descolada do ensino fundamental, evidenciam a importância e a necessidade da articulação entre a concepção de criança e educação que subjaz nessas etapas, para que estas trabalhem na perspectiva de um *continuum*, onde possam, apesar de guardadas as devidas diferenças de objetivos das etapas, preservar e compartilhar o eixo que as une e justifica essa articulação: os bebês e crianças, estejam na educação infantil ou no ensino fundamental.

Situar os bebês e crianças como o eixo central que une as etapas conduz à busca pela “construção e a consolidação de um currículo integrador que supere modelos curriculares fragmentados e descontínuos e considere bebês e crianças na sua integralidade” (SÃO PAULO, 2015, p.9)

A constituição de um currículo e, conseqüentemente, de uma avaliação integradora requer o compartilhamento de concepções e princípios sobre as diversas infâncias e seus direitos (SÃO PAULO, 2015).

Essa concepção, no entanto, precisa ser respaldada e fomentada por políticas públicas que atuem nessa perspectiva de articulação, a fim de que os profissionais não sejam considerados os únicos responsáveis pela qualidade da educação.

A articulação e a integração precisam ser assumidas no bojo das políticas como um princípio indispensável à noção de qualidade que sustenta essas etapas, pois, como nos lembra Sousa (2018, p.67), “tanto a noção de qualidade adotada, quanto o delineamento assumido para sua avaliação remetem a posicionamentos políticos e ideológicos”.

O desafio que se coloca para a área é de que esse posicionamento assumido na avaliação seja coerente com os pressupostos legais vigentes e com os conhecimentos nacionais e internacionais que respaldam o campo da educação infantil.

O estudo de caso mostrou que são muitas as variáveis intervenientes no processo de avaliação da aprendizagem na educação infantil que devem ser levados em consideração. Por exemplo, no município estudado, havia um contexto onde as profissionais não se apropriaram da importância da avaliação por meio da observação e do registro na educação infantil, situando-a como um dispositivo meramente burocrático e, por isso mesmo, dispensável, somado a um instrumento frágil e incipiente de avaliação utilizado nessa etapa.

A esse cenário, atrela-se uma configuração do atendimento municipal em que, em muitos casos, as unidades de educação infantil se situam como salas anexas dentro de instituições de ensino fundamental. Acrescente-se a isso a grande visibilidade pública e política da avaliação do ensino fundamental (que acarreta pressões de várias ordens) e uma tentativa de instituir um sistema meritocrático de bonificação das professoras do ensino fundamental e assim temos as “condições” propícias para o enfraquecimento da identidade e finalidade da educação infantil.

No entanto, a partir dos desafios identificados e mencionados anteriormente, podemos tecer algumas considerações sobre possíveis caminhos que podem ser melhor explorados e trilhados. Mereceria maiores estudos e investigações a questão da elaboração e disseminação dos instrumentos de avaliação da criança utilizados por professores(as), com foco nas iniciativas municipais que optam pela construção coletiva desses instrumentos, inserindo os profissionais na discussão e na sua elaboração, com vistas a compreender os aspectos que podem colaborar para maior legitimação e apropriação dos instrumentos pelos(as) profissionais.

O estudo de caso deixa clara a necessidade de integração da política municipal relativa à interface currículo-avaliação, educação infantil-ensino fundamen-

tal, e nos provoca a contrapor os modelos focados na responsabilização exclusiva dos profissionais (seja com premiações ou punições) frente às políticas de participação.

Assim, se por um lado reiteramos as conclusões do estudo progresso de que “há ainda um longo caminho a ser percorrido na construção e consolidação de uma avaliação da aprendizagem na educação infantil, na perspectiva da garantia dos direitos fundamentais das crianças” (RIBEIRO, 2018, p.241), por outro, acrescentamos algumas considerações que nos parecem ser imprescindíveis para que essa construção se efetive, como a articulação do debate da aprendizagem com as condições necessárias para a aprendizagem, instituição de processos formativos e apoio aos municípios e unidades educacionais através de orientações do governo federal e disseminação de práticas avaliativas alinhadas aos pressupostos legais e conhecimento da área.

As considerações aqui apontadas evidenciam aspectos que demandam e merecem ser melhor explorados, a fim de que possamos ter mais insumos que fomentem o debate e a proposição de avaliações da aprendizagem na educação infantil que possam realmente cumprir seu mandato de tornar visíveis e, por isso mesmo possíveis, as aprendizagens das crianças.

REFERÊNCIAS

AMADO, J. **Manual de investigação qualitativa em educação**. Coimbra: Universidade de Coimbra, 2013.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2011.

BRASIL. Lei n. 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, 23 dez. 1996.

_____. Conselho Nacional de Educação. **Parecer CNE/CEB nº 4/2008**. Orientação sobre os três anos iniciais do Ensino Fundamental de nove anos. 2008.

_____. Ministério da Educação. **Diretrizes Curriculares Nacionais para educação infantil**. Brasília: MEC, 2010.

_____. Ministério da Educação. Secretaria de Educação Básica. **Diretrizes Curriculares Nacionais para Educação Básica**. Diretoria de Currículo e Educação Integral. Brasília: MEC, SEB, DICEI, 2013.

CEAP (Coordenadoria de Ensino e Aprendizagem). **Caderneta Pedagógica**, 2015. Prefeitura Municipal: Secretaria de Educação/SEDUC, 2015.

FREITAS, L. C. Responsabilização, meritocracia e privatização: conseguiremos escapar ao neotecnicismo? In: **Seminário de Educação Brasileira**, 2011, Campinas. Anais. Campinas: UNICAMP/CEDES, 2011. Disponível em: http://www.cedes.unicamp.br/seminario3/luiz_freitas.pdf Acesso em: 15 jun. 2017.

IBGE/Instituto Brasileiro de Geografia e Estatística. **Censo Demográfico 2010**. Disponível em: <https://censo2010.ibge.gov.br/> Acesso em; 14 de agos. 2019.

MACHADO, C.; ALAVARSE, O. M. Responsabilização ou controle da qualidade do ensino: a que serve a avaliação externa? **Educação: Teoria e Prática**. Rio Claro, vol. 25, n.48, p. 67-79, jan./abr., 2015.

MELLO, S. A. **Uma proposta para pensar um currículo integrador da infância paulistana**. São Paulo: SME/DOT, 2015. (Mimeo).

MORGAN, D. L. **The focus group guidebook**. Thousand Oaks, CA: Sage publications, 1998.

_____. **Focus groups**. Thousand OAKS, CA: Sage publications, 2004.

RIBEIRO, B. Avaliação da aprendizagem das crianças: insumos para o debate. **Relatório Técnico**. São Paulo: MEC, 2016.

_____. Avaliação da Aprendizagem na Educação Infantil: um estudo exploratório em 125 municípios brasileiros. **Revista Linhas**, Florianópolis, v. 19, n. 40, p. 218-245, maio/ago., 2018.

SÃO PAULO (SP). Secretaria Municipal de Educação. Diretoria de Orientação Técnica. **Currículo integrador da infância paulistana**. São Paulo: SME/DOT, 2015.

SILVA, M. B. G.; FLORES, M. L. R. Articulações e tensões entre a educação infantil e o ensino fundamental: análises a partir do contexto recente das políticas educacionais brasileiras. In: FLORES, M. L.; ALBUQUERQUE, S. S. **Implementação do pré-infância no Rio Grande do Sul: perspectivas políticas e pedagógicas**. Porto Alegre: EDIPUCRS, 2015.

SOUSA, S. Z. Avaliação da Educação Infantil: questões controversas e suas implicações educacionais e sociais **Rev. educ. PUC-Camp.**, Campinas, v.23, n.1, p. 65-78, jan./abr., 2018

SOUSA, S. Z.; PIMENTA, C. O. Avaliação e Gestão da Educação Infantil em Municípios Brasileiros. **Educação & Realidade**, Porto Alegre, v. 43, n. 4, p. 1277-1300, out./dez., 2018.

SOUSA, S. Z. L.; PIMENTA, C.; MACHADO, C. Avaliação e Gestão Municipal da Educação. **Estudos em Avaliação Educacional**, São Paulo, v. 23, n. 53, p. 14-36, set./dez., 2012.

VYGOTSKY, L. S. A pré-história do desenvolvimento da linguagem escrita. In: **Obras escogidas**. Madrid: Visor, 1995.

YIN, R. K. **Estudo de caso**: planejamento e métodos. Trad. Daniel Grassi. 2.ed. Porto Alegre: Bookman, 2001.

EDUCAÇÃO INFANTIL E AVALIAÇÃO: PROCESSOS E PERSPECTIVAS

Rita de Cássia de Freitas Coelho¹

Alice de Paiva Macário²

¹ Pesquisadora no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF)- rita.coelho@caed.ufjf.br

² Pesquisadora no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF)- alice.macario@caed.ufjf.br

Resumo

Este artigo tem por objetivo fomentar a discussão sobre avaliação da educação infantil e apresentar as propostas de avaliação da educação infantil em construção no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF), como contribuição para uma metodologia brasileira de avaliação da educação infantil. As propostas estabelecem três estratégias metodológicas de avaliação da educação infantil que podem ser implementadas de forma independente ou complementar: a) pesquisa de campo com uso de questionários aplicados online a diretores(as), coordenadores(as), professores(as) e auxiliares que atuam na educação infantil; b) avaliação por meio de observação em turmas de educação infantil; e c) avaliação das condições da oferta a partir da análise dos dados do Censo Escolar da Educação Básica do INEP, do Censo Demográfico e da PNAD do IBGE. As propostas se constituem com base nas Diretrizes Curriculares Nacionais para a Educação Infantil (BRASIL, 2009), no Plano Nacional de Educação (BRASIL, 2014), na Base Nacional Comum Curricular (BRASIL, 2017), em diferentes documentos do Ministério da Educação e em experiências nacionais e internacionais.

Palavras-chave: Educação Infantil. Avaliação. Qualidade da Educação. Política Educacional.

Abstract:

This article aims to foster the discussion on the evaluation of early childhood education and to present the proposals for the evaluation of early childhood education in development at the Center for Public Policies and Education Evaluation of the Federal University of Juiz de Fora (Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora - CAEd / UFJF), as a contribution to a Brazilian methodology for assessing early childhood education. The proposals establish three methodological strategies for evaluating early childhood education that can be implemented independently or complementarily: a) field research sending online questionnaires to principals, coordinators, teachers and assistants who work in early childhood education; b) evaluation through observation in early childhood education classes; and c) context evaluation based on the analysis of data from the School Census of Basic Education from INEP, the Demographic Census and PNAD from IBGE. The proposals are based on the National Curriculum Guidelines for Early Childhood Education (BRASIL, 2009), on the National Education Plan (BRASIL, 2014), on the Common National Curricular Base (BRASIL, 2017), different documents from the Ministry of Education, and national and international experiences.

Keywords: Early Childhood Education. Evaluation. Quality of Education. Educational policy.

APRESENTAÇÃO

O Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF) é uma instituição de referência nacional, no âmbito da avaliação educacional, e que vem implementando projetos, pesquisas e programas de avaliação, por meio de convênios e contratos com as redes públicas de ensino federal, estaduais e municipais e outras instituições.

A partir do desenvolvimento de metodologias, instrumentos e processos de avaliação, as informações produzidas são discutidas com as equipes interessadas e subsidiam um processo de pesquisa formativo e de apropriação dos resultados, que busca contribuir com a qualidade da educação.

Considerando a significativa expansão e importância da educação infantil (EI), o CAEd/UFJF organizou, a partir de 2017, um trabalho com foco na pesquisa em avaliação da educação infantil, visando a desenvolver metodologias de avaliação que considerem as especificidades dessa etapa da educação básica, os princípios estabelecidos pelas Diretrizes Curriculares Nacionais para a Educação Infantil – DCNEI (BRASIL, 2009), as determinações do Plano Nacional de Educação – PNE (BRASIL, 2014-2024), a BNCC (BRASIL, 2017) e as experiências internacionais e nacionais.

Este texto objetiva analisar as questões em debate sobre a avaliação na/da educação infantil no contexto brasileiro e apresentar as propostas que o CAEd/UFJF vem desenvolvendo.

CONTEXTOS E AÇÕES

Com a promulgação da Constituição Federal de 1988 (BRASIL, 1988), da LDB (BRASIL, 1996), do PNE (BRASIL, 2014) e do Marco Legal da Primeira Infância (BRASIL, 2016), a educação infantil vive um processo crescente e coerente de efetiva integração ao Sistema Educacional Brasileiro. Essa integração é multidimensional e abrange processos de natureza jurídica, administrativa, financeira, cultural e pedagógica.

Iniciada com a gestão das redes de atendimento, sendo assumida pelo setor da educação, financiada como primeira etapa da educação básica no âmbito do FUNDEB³, orientada nacionalmente pelas DCNEI (BRASIL, 2009), pela BNCC (BRASIL, 2017) e regulamentada nos respectivos sistemas de ensino, a integração da educação infantil ao Sistema Educacional vem, gradativamente, consolidando uma identidade pedagógica própria dessa etapa, distinta da do ensino fundamental.

A aprovação da Base Nacional Comum Curricular da Educação Infantil (BRASIL, 2017), que afirma a centralidade da criança e estabelece uma organização curricular específica por campos de experiências, deve ser entendida como uma das fases da integração em curso no país, assim como um marco importante na política nacional da educação infantil.

Nessa trajetória, evidenciam-se inúmeros desafios e disputas, dentre os quais merecem destaque: a demanda por expansão do atendimento e melhoria da qualidade, o debate sobre o papel da educação infantil no processo de alfabetização e a exigência de uma avaliação que abarque as especificidades da educação infantil.

Existe um consenso sobre a necessidade da garantia de um “padrão” de qualidade, de modo a assegurar o **direito à educação**, que não se reduz a vagas, mas está diretamente vinculado às práticas pedagógicas cotidianas ofertadas aos bebês e crianças matriculadas nas instituições de educação infantil, principalmente quando consideradas a diversidade de contextos e a desigualdade social presentes no país. No enfrentamento dessas questões, a avaliação pode trazer significativas contribuições para a política e as instituições. Mas de qual avaliação estamos falando?

É importante que, na construção de uma cultura da avaliação na educação infantil se distingam, conceitualmente, os diferentes níveis possíveis de avaliação: avaliação de sistema, avaliação de políticas, avaliação de programas, avaliação institucional e avaliação da criança, assim como os diferentes tipos de avaliação diagnóstica, formativa e somativa, interna e externa, que geram, conseqüentemente, diferentes ações, estratégias e resultados.

De modo resumido, a área tem demarcado duas perspectivas no debate da avaliação no campo da educação infantil: a avaliação **na** educação infantil [interna, realizada pelos profissionais e professores(as)], podendo, ou não, envolver a avaliação da criança; e a Avaliação **da** educação infantil [externa, realiza-

³ Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação.

da por profissionais de fora do contexto da instituição], de caráter institucional, não envolvendo a avaliação da criança. Entretanto, constata-se que a conceitualização ou compreensão destas perspectivas vem sobrepondo-se uma tensão central sobre a possibilidade de a avaliação da criança ser inserida no âmbito das avaliações externas. Entre essas grandes tendências, inúmeras iniciativas de municípios, universidades, fundações, organizações não governamentais e organismos internacionais têm sido implementadas. Segundo Rosemberg (2013), observa-se um duplo movimento: a incorporação da educação infantil na política de avaliação (movimento externo à EI) e a incorporação da avaliação como parte da política de educação infantil (movimento interno da área da EI).

A utilização de avaliação em larga escala na educação infantil é um tema complexo e controverso pelo desafio existente de se produzir medidas de proficiência, a partir da aplicação de instrumentos avaliativos nas crianças, construídos fora do contexto em que se insere a oferta. Os debates atuais sobre avaliação também destacam a necessidade de se levar em consideração o contexto e a importância de que a avaliação possibilite a compreensão pedagógica da prática educativa. Nesse sentido, a avaliação em larga escala será sempre **da** educação infantil.

A legislação educacional estabelece que a avaliação da criança na educação infantil é uma competência das instituições, visando ao acompanhamento do processo pedagógico e do desenvolvimento dos bebês e das crianças sem o objetivo de seleção, promoção ou classificação, o que nos remete às responsabilidades internas ao processo educativo focalizado nas interações com o bebê e a criança em suas conquistas, possibilidades e desafios. Como demarca o inciso I do artigo 31 da LDB (BRASIL, 1996, p.31) “A educação infantil será organizada de acordo com as seguintes regras comuns: I – avaliação mediante acompanhamento e registro do desenvolvimento das crianças, sem o objetivo de promoção, mesmo para o acesso ao ensino fundamental”. As DC-NEI (BRASIL, 2009, p.22), complementando esse princípio, determinam que:

Art. 10 – As instituições de Educação Infantil devem criar procedimentos para acompanhamento do trabalho pedagógico e para avaliação do desenvolvimento das crianças, sem objetivo de seleção, promoção ou classificação, garantindo: I – a observação crítica e criativa das atividades, das brincadeiras e interações das crianças no cotidiano; II – utilização de múltiplos registros realizados por adultos e crianças (relatórios, fotografias, desenhos, álbuns etc.); III – a continuidade dos processos de aprendizagens por meio da criação de estratégias adequadas aos diferentes momentos de transição vividos pela criança (transição casa/instituição de Educação Infantil, transições no interior da instituição, transição creche/pré-escola e

transição pré-escola/Ensino Fundamental); IV – documentação específica que permita às famílias conhecer o trabalho da instituição junto às crianças e os processos de desenvolvimento e aprendizagem da criança na Educação Infantil; V – a não retenção das crianças na Educação Infantil.

Por outro lado, existe um consenso nacional, expresso no PNE (BRASIL, 2014-2024), sobre a necessidade de uma avaliação nacional da educação infantil, integrando a política nacional de avaliação da educação básica. A avaliação nacional proposta no plano deve ser realizada de dois em dois anos e trata-se de avaliação em larga escala da educação infantil com o uso de instrumentos externos que focalizem a qualidade das instituições “com base em parâmetros nacionais de qualidade, a fim de aferir a infraestrutura física, o quadro de pessoal, as condições de gestão, os recursos pedagógicos, a situação de acessibilidade, entre outros indicadores relevantes” (BRASIL, 2014, p.03). Tais características remetem a um novo desafio em termos do Sistema Nacional de Avaliação da Educação Básica (SAEB).

Simultaneamente, principalmente a partir de propostas de organismos internacionais e demandas de gestores municipais, cresce a pressão por incorporar a avaliação da criança às avaliações externas em larga escala da educação infantil.

Entendida como um dever de Estado, a avaliação da educação infantil é mais do que apenas uma exigência legal, é uma responsabilidade social frente à vulnerabilidade da criança na primeira infância, a importância desse ciclo de idade de formação, a obrigatoriedade de matrícula na pré-escola e o volume de recursos públicos investidos.

É preciso considerar esse debate como mais uma fase do processo de integração da educação infantil ao Sistema Educacional que contribuirá para demarcar a identidade própria como primeira etapa da educação básica com suas especificidades no trabalho pedagógico ou para reforçar o risco de considerá-la, como preparatória para o ensino fundamental.

A opção histórica da integração da educação infantil ao Sistema Educacional foi a resposta política que nossa sociedade construiu para enfrentar a tensão entre o assistencialismo, a filantropia e o direito de todas as crianças à educação emancipatória. Diante disso, a avaliação é uma das características sistêmicas da educação com a qual precisamos dialogar e formular propostas comprometidas com a infância. Negar essa característica que se coloca atualmente como forte pressão, acreditando ser uma estratégia adequada de defesa da identidade da Educação Infantil, é abrir mão de um espaço que será ocupado por outras concepções e interesses.

É nessa perspectiva que o CAEd/UFJF vem atuando, buscando qualificar o debate e construir propostas de avaliação coerentes com a concepção de educação infantil, que garantam o cumprimento do dever do Estado e a igualdade de direito de todos os bebês e crianças à EI de qualidade.

Propostas já implementadas na Austrália, no México, nos Estados Unidos, em Portugal e na Itália estão presentes no debate nacional, contribuindo para compreender os objetivos, os instrumentos, a metodologia e os elementos que podem auxiliar na construção de uma proposta brasileira de avaliação da/na educação infantil.

Dessas experiências, duas foram implementadas em pesquisas no Brasil e se colocam como referências importantes. A primeira, as escalas norte americanas de avaliação de ambiente: *Infant Toddler Environment Rating Scale* (ITERS), com foco nos bebês e crianças de até 2 anos e 6 meses e a *Early Childhood Environment Rating Scale* (ECERS), voltada para as crianças de 2 anos e 6 meses a 5 anos de idade foram objeto de uma pesquisa nacional coordenada pela Fundação Carlos Chagas, em parceria com o Ministério da Educação e com o Banco Interamericano de Desenvolvimento. Esses instrumentos de observação das instituições de educação infantil foram utilizados em cinco capitais brasileiras para avaliar a qualidade da educação infantil, oferecendo parâmetros de avaliação para o observador externo, com possibilidades de graduação da qualidade observada em níveis que vão de 1 a 7, elencados em uma escala: de inadequado, mínimo, bom a excelente. Ambas as escalas estão organizadas em 7 dimensões: **espaço e mobiliário; rotinas de cuidado pessoal, linguagem e raciocínio, atividades, interação, estrutura do programa, pais e equipe.**

A segunda refere-se ao projeto desenvolvido pela Universidade Federal do Paraná, em parceria com a Universidade Federal do Rio de Janeiro, a Universidade do Estado de Santa Catarina, a Universidade Federal de Minas Gerais, a Università degli Studi di Pavia, com coordenação técnica e apoio financeiro do Ministério da Educação. O estudo foi construído a partir da análise dos instrumentos italianos de autoavaliação de contexto: ISQUEN – *Indicatori e Scala della Qualità Educativa del Nido* (Indicadores e Escala da Qualidade Educativa da Creche) e AVSI – *Autovalutazione della Scuola dell'Infanzia* (Autoavaliação da Pré-escola). O ISQUEN é um instrumento formado por 51 descritores, em que cada um apresenta três opções (a, b, c), que podem ser selecionadas. Os descritores são organizados em quatro grupos: **os sujeitos; os contextos e as práticas; os saberes do fazer; as garantias.** O AVSI é um instrumento extenso, constituído por 81 descritores, graduados em níveis de

1 a 9, sendo que cinco deles possuem parâmetros estabelecidos (1-3-5-7-9) e os outros são pontuações intermediárias (2-4-6-8), ordenados em quatro grupos: **a experiência educativa; as atividades profissionais; os adultos e suas relações; as garantias.** Dentre as diferentes ações do projeto, destaca-se o estudo de campo desenvolvido em quatro capitais brasileiras visando à análise metodológica dos instrumentos de avaliação de contexto.

Do ponto de vista nacional, merecem destaque as experiências de uso dos Indicadores da Qualidade na Educação Infantil (BRASIL, 2009). Embora concebido como um instrumento para as escolas, foram utilizados por diferentes segmentos como secretarias municipais, conselhos de educação, fóruns de Educação Infantil e docentes do ensino superior totalizando 10.090 participações. Esse instrumento foi construído para promover a autoavaliação das instituições de educação infantil brasileiras, constituído por 7 dimensões: (1) Planejamento institucional; (2) Multiplicidade de experiências e linguagens; (3) Interações; (4) Promoção da saúde; (5) Espaços, materiais e mobiliários; (6) Formação e condições de trabalho dos(as) professores(as) e demais profissionais; (7) Cooperação e troca com as famílias e participação na rede de proteção social. Cada dimensão ou tema apresenta itens que descrevem diferentes possibilidades com relação aos elementos constituintes de uma instituição de educação infantil de qualidade.

Vale ainda mencionar dois importantes documentos decorrentes de iniciativas do MEC/SEB e do INEP: *Educação Infantil: subsídios para a construção de uma sistemática de avaliação* (2012)⁴ e a *Proposta para Avaliação da Educação Infantil: elementos para sua operacionalização* (2015)⁵.

Embora a Base Nacional Comum Curricular da Educação Infantil (BNCC-EI) esteja em fase inicial de implementação e apropriação, o fato de seu conteúdo estar fortemente embasado nas Diretrizes Curriculares Nacionais da Educação Infantil, na análise de propostas pedagógicas de municípios brasileiros e, recentemente, ter sido homologada pelo Ministério da Educação, após aprovada pelo Conselho Nacional de Educação, a BNCC-EI se coloca como uma referência fundamental para a formulação de uma proposta de avaliação para a Educação Infantil.

⁴ Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=11990-educacao-infantil-sitematica-avaliacao-pdf&category_slug=novembro-2012-pdf&Itemid=30192.

⁵ Relatório com estudos e discussões da Comissão Assessora de Especialistas da Avaliação da Educação Infantil (INEP).

PERSPECTIVAS EM CONSTRUÇÃO NO CAED/UFJF

A partir desses múltiplos referenciais, a construção das propostas de avaliação da educação infantil em curso no CAEd/UFJF reafirma os **pressupostos básicos** apresentados no documento Educação Infantil: subsídios para a construção de uma sistemática de avaliação (2012)⁶. Ou seja, a avaliação da educação infantil:

- » não envolve a avaliação do desempenho da criança;
- » respeita os objetivos propostos pelas Diretrizes Curriculares Nacionais para a Educação Infantil, coerente com as singularidades e características da educação infantil;
- » considera o contexto em que se insere a oferta da educação infantil;
- » está fundamentada em parâmetros/critérios de qualidade negociados;
- » promove um processo participativo e formativo;
- » exige um diálogo multidisciplinar com pesquisadores e profissionais das áreas da educação infantil, estatística, construção de medidas e tecnologia da informação.

A partir de uma perspectiva mais ampla, compreendendo as condições da oferta e as práticas oferecidas às crianças, são objetivos da avaliação da educação infantil:

- produzir diagnóstico acerca da realidade do atendimento ofertado na Educação Infantil;
- buscar informações sobre as práticas pedagógicas implementadas pelos(as) professores(as) de Educação Infantil;
- fornecer aos(as) gestores(as) e professores(as) informações sobre as características da oferta, da gestão e das práticas docentes;
- contribuir com o aprimoramento da competência docente;
- recomendar ações de melhoria da política municipal de Educação Infantil.
- construir uma cultura de avaliação na Educação Infantil como processo participativo e formativo.

⁶ Documento resultado de um processo de discussão do qual participaram: I – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP); II – Secretaria de Educação Básica do Ministério da Educação (SEB); III – Fundo Nacional de Desenvolvimento da Educação (FNDE); IV – União Nacional dos Dirigentes Municipais de Educação (UNDIME); V – Confederação Nacional dos Trabalhadores em Educação (CNTE); VI – Associação Nacional de Pós-Graduação e Pesquisa em Educação (ANPED); VII – Movimento Interfóruns de Educação Infantil do Brasil (MIEIB); e VIII – Rede Nacional Primeira Infância (RNPI).

- A proposta em construção no CAEd estabelece três estratégias metodológicas de avaliação da educação infantil que podem ser implementadas de forma independente ou complementar: a) Pesquisa de campo com uso de questionários aplicados online a diretores(as), coordenadores(as), professores(as) e auxiliares que atuam na educação infantil; b) Avaliação por meio de observação em turmas de educação infantil; c) Avaliação das condições da oferta a partir da análise dos dados do Censo Escolar da Educação Básica do INEP, do Censo Demográfico e da PNAD do IBGE.

Pesquisa de campo com uso de questionários

A primeira proposta tem como base os instrumentos (questionários) já utilizados pelo CAEd/UFJF para a avaliação do ensino fundamental, incluindo questões relevantes relacionadas às práticas pedagógicas e à elaboração de um instrumento específico para os auxiliares/monitores/assistentes da educação infantil.

O primeiro momento da pesquisa avaliativa, por meio de questionários, exige a definição do conteúdo: escolha dos temas, elaboração dos itens, organização dos itens em blocos no questionário, abrangendo o perfil dos profissionais, as condições de infraestrutura das instituições, as práticas pedagógicas, a gestão pedagógica, o ambiente educativo, a relação com a família e comunidade e as expectativas dos profissionais em relação às crianças. A seguir, o instrumento é revisado e editorado.

A partir desses procedimentos, é estruturado o processo de aplicação, cronograma da coleta e os questionários são disponibilizados para os respondentes (diretores(as), coordenadores(as) pedagógicos(as), professores(as) e auxiliares), em uma plataforma online. No período de aplicação, o processo é monitorado e acompanhado por meio da Central de Atendimento do CAEd/UFJF, que atende demandas de esclarecimentos de dúvidas técnicas e de conteúdo pedagógico.

Após o término da aplicação, as respostas são processadas e passam pelo controle de qualidade, verificando se os dados do planejamento (base de dados) conferem com os dados processados (quantidade de itens por questionário, quantidade de opções por item e quantidade de questões por ator). Em seguida, é realizado o tratamento da informação e a validação dos dados.

Depois de validados, os dados são submetidos à análise descritiva, na qual são detalhadas as frequências e as estatísticas (médias, mínimos, totais, porcentagens etc.). Dependendo do tamanho da amostra, as respostas podem passar pela modelagem dos dados (criação de medidas, indicadores, índices etc.).

Posteriormente, as informações geradas e validadas são organizadas em planilhas de Excel, analisadas qualitativamente pela equipe do CAEd/UFJF e sistematizadas em Relatório. Em seguida, é realizado amplo seminário de devolutiva dos principais resultados, iniciando um processo de apropriação dos resultados.

Essa proposta foi implementada em oito municípios da região da Zona da Mata mineira (Belmiro Braga – 3503 habitantes; Coronel Pacheco – 3125 habitantes; Goianá – 3952 habitantes; Guarará – 3938 habitantes; Juiz de Fora – 563769 habitantes; Pequeri – 3347 habitantes; São João Nepomuceno – 26538 habitantes; Simão Pereira – 2651 habitantes), em um município da região metropolitana de Belém, no Pará (Benevides – 60990 habitantes) e em uma capital do Nordeste (São Luís – 1.094.667), totalizando dez municípios com diferentes características, como pode ser observado no quadro abaixo:

Quadro 1 – Indicadores contextuais dos municípios

	Belmiro Braga	Benevides	Coronel Pacheco	Goianá	Guarará	Juiz de Fora	Pequeri	São João Nepomuceno	São Luís	Simão Pereira
População Geral	3.503	60.990	3.125	3.952	3.938	563.769	3.347	26538	1.094.667	2.651
Oferta EImatrículas	83	2.304	66	111	82	19.594	142	1.117	45.296	86
Número de instituições	3	26	2	1	1	262	2	15	440	2
Oferta-0 a 3 anos	Não atende	628	Não atende	Não atende	Não atende	7.813	49	574	17.476	23
Oferta-4 e 5 anos	83	1.676	66	111	82	11.781	93	543	27.820	63
IDHM	0,660	0,665	0,669	0,716	0,652	0,778	0,638	0,708	0,768	0,694
IDEB 2017 Anos Iniciais do EF	5,4	6,2	6,0	6,0	5,2	5,2	5,8	6,4	4,6	3,7
Plano Municipal	Não	Sim	Não	Sim	Sim	Sim	Não	Sim	Sim	Não

Fonte: Dados da Sinopse Estatística do Censo Escolar (2017 e 2018), do IDEB e do Atlas do Desenvolvimento Humano.

Quadro 2 – Participação dos atores escolares

Município	Total de diretores(as)	Total de coordenadores(as)	Total de professores(as)	Total de assistentes técnicos(as) educacionais	Total
Belmiro Braga/MG	0	2	3	-	5
Benevides/PA	4	9	50	40	103
Coronel Pacheco/MG	1	1	3	-	5
Goianá/MG	0	1	9	-	10
Guarará/MG	1	0	8	-	9
Juiz de Fora/MG	26	31	189	-	246
Pequeri/MG	1	1	10	-	12
São João Nepomuceno/MG	7	7	80	-	94
São Luís	90	100	720	30	940
Simão Pereira/MG	1	1	12		14
Total	41	53	364	40	1438

Fonte: Dados das avaliações (CAEd/UFJF).

São quatro tipos de questionários, sendo um para diretores(as), um para coordenadores(as) pedagógicos(as), um para professores(as) e um para auxiliares/monitores/ assistentes que atuam na educação infantil.

Esses instrumentos incluem questões referentes às expectativas dos profissionais em relação às crianças e às opiniões dos respondentes sobre práticas pedagógicas desenvolvidas nas instituições, de modo que os resultados podem oferecer subsídios para a gestão municipal em relação à formação continuada, supervisão, monitoramento e melhoria da qualidade do atendimento, visando a assegurar os direitos e objetivos de aprendizagem.

Apesar dos limites da pesquisa autodeclaratória, esta metodologia permite uma visão geral das condições de oferta, das principais concepções dos profissionais que atuam na educação infantil, explicita diferenças entre a opinião dos(as) gestores(as) e professores(as) e aponta expectativas em relação ao percurso das crianças possibilitando analisar congruências, divergências, contradições e levantar práticas desenvolvidas, ou não, nas instituições.

Com o objetivo de cotejar os resultados a partir das respostas “socialmente desejáveis” e avançando na perspectiva de uma avaliação pedagógica da educação infantil, propomos a utilização de um instrumento de observação das práticas oferecidas pelos(as) professores(as) às crianças.

Observação de turmas de Educação Infantil a partir de instrumento de observação

Construir um instrumento de observação da educação infantil, de forma participativa, exige um significativo grau de escuta e diálogo. Desse modo, buscamos parceria com outras instituições, Universidades e Secretarias Municipais de Educação, na construção compartilhada de um instrumento de observação.

A construção de um instrumento de observação exige a escolha e definição de parâmetros, dimensões, temas e descritores que configuram uma Matriz de Referência. Para tal, optamos por selecionar parâmetros e indicadores de qualidade, preferencialmente regulamentados/normatizados, no âmbito do Sistema Educacional Brasileiro ou relativamente reconhecidos de forma consensual, uma vez que esses parâmetros oferecem um conjunto de referenciais para a oferta da educação de qualidade.

O instrumento de observação organiza-se em três dimensões interdependentes, que são:

1. **Infraestrutura, espaços e materiais**, abordando aspectos da estrutura física dos espaços internos e externos, a diversidade e disponibilidade de materiais e recursos, o mobiliário e as condições de segurança e higiene.
2. **Práticas educativas**, abrangendo: as formas de tratamento e acolhida da criança e da família; a participação e interação das e com as crianças; as oportunidades de deslocamentos e desafios corporais; a exploração de movimentos, gestos e olhares com o corpo; a convivência com diversas manifestações artísticas e culturais; a imersão na cultura escrita; a oralidade; a produção de textos; experiências com a literatura infantil (manipulação de livros); apoio à escrita espontânea; observação e manipulação de elementos da natureza; atividades com conhecimento matemático; identificação de relações espaciais e temporais; experiências de quantificar, classificar e seriar.
3. **Ambiente educativo e trabalho colaborativo**, destacando a relação com as famílias, as relações de cooperação na instituição, os registros individuais de cada criança, a frequência dos professores(as), razão professor(a) – criança, promoção e apoio ao aleitamento materno, ações integradas com serviços da rede de proteção integral e as práticas de saúde e cuidado.

Para cada descritor, é possível identificar quatro níveis (condições) que podem ser observados:

- I. **Abaixo do básico** – padrão de oferta que não atende aos parâmetros de qualidade estabelecidos para a educação infantil no país.
- II. **Básico** – padrão de oferta abaixo do adequado, mas que atende, parcialmente, aos parâmetros e caracteriza as condições mínimas de qualidade da oferta da educação infantil.
- III. **Adequado** – padrão de oferta desejável, segundo os parâmetros de qualidade e as regulamentações educacionais.
- IV. **Avançado** – padrão de oferta acima do adequado, em que são contemplados, plenamente, os parâmetros de qualidade, as regulamentações educacionais e os aspectos apontados por pesquisas.

Antes de ser utilizado em determinada realidade, o instrumento deve ser analisado do ponto de vista conceitual e pedagógico em reunião com gestores(as) e professores(as) da realidade local. Esse momento é importante para o estudo, problematização, revisão, reclassificação dos temas e itens, buscando compreender as discordâncias e estabelecer consensos sobre os critérios de qualidade da educação infantil.

No caso da observação dentro de um processo avaliativo, é preciso ter clareza dos parâmetros estabelecidos, devendo ser relativamente estáveis e consuetudinários. Desse modo, é necessário colocar uma lente que permita compreender o contexto observado e suas diferentes nuances, evidenciando a relevância do roteiro de observação que auxilie o(a) observador(a) nas tomadas de decisão, pressupondo estudo e conhecimento acerca do que será observado.

São objetivos específicos da observação:

- Verificar in loco as condições da infraestrutura das instituições, a organização dos espaços e a disponibilidade de materiais.
- Observar no cotidiano as práticas pedagógicas implementadas na docência.
- Perceber o envolvimento das crianças e as interações entre adultos e crianças.
- Elucidar aspectos analisando as informações recebidas pelos questionários com os aspectos observados pelo(a) pesquisador(a).

A observação foi realizada em uma amostra de instituições sorteadas entre aquelas que responderam aos questionários nos municípios de Juiz de Fora – MG e São Luís – MA. Estabelecidas as instituições, as turmas foram observadas pelo período mínimo de 04 (quatro) horas ou máximo de 08 (oito) horas.

A equipe de Observadores(as) foi composta por técnicos(as) da Secretaria Municipal de Educação de Juiz de Fora e pesquisadores(as) do CAEd/UFJF. O CAEd/UFJF realizou a formação dos(as) técnicos(as) que atuaram como observadores(as) externos(as), com duração de 16 horas, abrangendo as três grandes dimensões: a) Infraestrutura, espaços e materiais; b) Práticas Pedagógicas; c) Ambiente educativo e trabalho colaborativo.

A observação é sempre um recorte da realidade, mas nos permite ter indícios sobre determinadas dimensões pedagógicas, o que exige um compromisso ético e implicado com o processo avaliativo, que não deve ser generalizado ou tomado de forma isolada das variáveis que os constituem.

Nessa perspectiva, o instrumento de observação integra a discussão nacional no que tange à avaliação da educação infantil, principalmente por seu caráter formativo, no qual todos(as) os(as) participantes (observadores(as) externos(as), professores(as), técnicos(as) da secretaria) negociam concepções e se implicam em prol de uma avaliação capaz de evidenciar a qualidade da educação infantil em determinado contexto.

Análise de indicadores sociais nacionais

Indicadores sociais são bens públicos, meritórios. Seu uso por alguns não restringe a utilização pelos demais. Na realidade, quanto mais públicos forem, mais bem-estar podem gerar (JANNUZZI, 2017).

Como terceira proposta de atuação do CAEd/UFJF, consideramos uma avaliação a partir da análise de indicadores sociais relevantes já coletados no âmbito do Censo Escolar da Educação Básica, do INEP⁷, da Pesquisa Nacional por Amostra de Domicílios (PNAD)⁸ e do Censo Demográfico do IBGE⁹.

⁷ Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

⁸ Os indicadores são válidos apenas para os estados, com atualização um ano e meio após a coleta, e forte possibilidade no horizonte de algumas informações não mais existirem se a pesquisa for mesmo substituída pela Pnad Contínua.

⁹ Instituto Brasileiro de Geografia e Estatística.

De acordo com o relatório de 2015 da Comissão Assessora de Especialistas da Avaliação da Educação Infantil, constituída pelo INEP, destacamos os seguintes dados e indicadores que podem compor, de imediato, uma avaliação em larga escala da educação infantil:

Quadro 3 – Eixos da Avaliação da Educação Infantil

Dimensões	Descritores
Dimensão A <i>Acesso / Oferta (por idade)</i>	Demanda potencial e demanda manifesta. Oferta. Localização da oferta (urbano/rural; centro/periferia). Período de atendimento. Jornada. Dependência administrativa. Inclusão de crianças com deficiência. Perfil da criança atendida (nível socioeconômico; idade; cor).
Dimensão B <i>Infraestrutura</i>	Natureza do espaço (prédio próprio; alugado; compartilhado; vinculado). Espaços internos (existência). Banheiro (existência). Espaços externos (existência). Acesso à água filtrada. Abastecimento de água (existência). Condições de saneamento básico (existência). Coleta de lixo (existência). Tratamento do lixo (existência). Abastecimento de energia elétrica. Limpeza dos espaços (adequada). Iluminação (existência). Acústica (adequada).
Dimensão C <i>Recursos Materiais</i>	Brinquedos (existência). Livros de literatura infantil (existência). Mobiliário e equipamentos (existência). Alimentação (tipo; frequência; quantidade).
Dimensão D <i>Profissionais de Educação Infantil</i>	Docente que atua com a turma (formação inicial; formação continuada; regime de contrato; remuneração; plano de carreira; jornada de trabalho). Auxiliar/assistente educacional que atua com a turma (formação inicial; formação continuada; regime de contrato; remuneração; plano de carreira; jornada de trabalho). Gestor da unidade escolar (formação inicial; formação continuada; regime de contrato; remuneração; plano de carreira; jornada de trabalho).

Fonte: Relatório da Comissão Assessora de Especialistas da Avaliação da Educação Infantil (2015).

Quadro 4 – Indicadores definidos a partir das fontes de dados disponíveis

Indicador desenvolvido	Fonte de dados
Dimensão A <i>Acesso / Oferta (por idade)</i>	
A1: Percentual de crianças de zero a três anos de idade que frequentam escola ou creche.	Pnad
A2: Percentual de crianças de quatro e cinco anos de idade que frequentam escola ou creche.	Pnad
A3: Diferença entre as taxas de crianças de zero a três anos de idade que frequentam escola ou creche, residentes em áreas rurais e em áreas urbanas.	Pnad
A4: Diferença entre as taxas de crianças de quatro e cinco anos de idade que frequentam escola ou creche, residentes em áreas rurais e em áreas urbanas.	Pnad
A5: Diferença entre as taxas de crianças negras e brancas de zero a três anos de idade que frequentam escola ou creche.	Pnad
A6: Diferença entre as taxas de crianças negras e brancas de quatro e cinco anos de idade que frequentam escola ou creche.	Pnad
A7: Diferença entre as taxas de crianças de zero a três anos de idade que frequentam escola ou creche do quinto de renda familiar <i>per capita</i> mais elevado e do quinto de renda familiar <i>per capita</i> mais baixo.	Pnad
A8: Diferença entre as taxas de crianças de quatro e cinco anos de idade que frequentam escola ou creche do quinto de renda familiar <i>per capita</i> mais elevado e do quinto de renda familiar <i>per capita</i> mais baixo.	Pnad
A9: Percentual de crianças de áreas rurais que frequentam unidades educacionais rurais.	Censo Escolar – Cadastro de Alunos
A10: Percentual de turmas de educação infantil em tempo integral.	Censo Escolar – Cadastro de Turmas
A11: Percentual de turmas de creches que possuem, no máximo, quinze (15) crianças.	Censo Escolar – Cadastro de Turmas
A12: Percentual de turmas de pré-escolas que possuem, no máximo, vinte e cinco (25) crianças.	Censo Escolar – Cadastro de Turmas
A13: Percentual de estabelecimentos exclusivos de educação infantil.	Censo Escolar – Cadastro de Escolas
Dimensão B <i>Infraestrutura</i>	
B1: Percentual de unidades de educação infantil que funcionam em prédio escolar próprio.	Censo Escolar – Cadastro de Escolas
B2: Percentual de creches que possuem berçário.	Censo Escolar – Cadastro de Escolas
B3: Percentual de unidades educacionais que possuem banheiro(s), dependências e vias adequadas a crianças com deficiência ou mobilidade reduzida.	Censo Escolar – Cadastro de Escolas
B4: Índice de existência de espaços internos nas unidades educacionais.	Censo Escolar – Cadastro de Escolas
B5: Índice de existência de espaços externos nas unidades educacionais.	Censo Escolar – Cadastro de Escolas
B6: Índice de existência de serviços básicos de infraestrutura nas unidades educacionais.	Censo Escolar – Cadastro de Escolas

Indicador desenvolvido	Fonte de dados
Dimensão C <i>Recursos materiais</i>	
C1: Índice de existência de recursos materiais nas unidades educacionais.	Censo Escolar – Cadastro de Escolas
Dimensão D <i>Profissionais de educação infantil</i>	
D1: Percentual de docentes com ensino superior concluído.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D2: Percentual de docentes com ensino médio – Normal.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D3: Percentual de docentes com ensino superior em andamento.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D4: Percentual de docentes com ensino superior concluído em Pedagogia.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D5: Percentual de docentes com ensino superior concluído que possuem formação/complementação pedagógica.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D6: Percentual de docentes de unidades educacionais públicas que são concursados/efetivos/estáveis.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D7: Percentual de auxiliares/assistentes educacionais com ensino superior concluído.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D8: Percentual de auxiliares/assistentes educacionais com ensino médio – Normal.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D9: Percentual de auxiliares/assistentes educacionais com ensino Superior em andamento.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D10: Percentual de auxiliares/assistentes educacionais com ensino superior concluído em Pedagogia.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula
D11: Percentual de auxiliares/assistentes educacionais com curso superior concluído que possuem formação/complementação pedagógica.	Censo Escolar – Cadastro de Profissionais Escolares em Sala de Aula

Fonte: Relatório da Comissão Assessora de Especialistas da Avaliação da Educação Infantil (2015).

Essa perspectiva, além de atender à legislação nacional no que tange à avaliação da educação infantil, contribui para a formulação e implementação de políticas públicas, possibilitando um diagnóstico abrangente por estados e municípios, subsidiando o processo decisório de programas e projetos governamentais.

CONSIDERAÇÕES COMPARTILHADAS

Diante do exposto, é necessário avançar no que tange ao desafio de tentar conhecer/avaliar as práticas implementadas na educação infantil como uma forma de fortalecer a identidade pedagógica desta etapa e seu compromisso em possibilitar experiências significativas aos bebês e crianças, de modo a assegurar os objetivos e direitos de aprendizagem e desenvolvimento integral.

Apesar de a qualidade ser um conceito polissêmico, é possível afirmar que a legislação e os documentos orientadores da educação infantil consolidam uma concepção de educação infantil em seus múltiplos aspectos orientando a construção de metodologias avaliativas.

Nesse movimento, destacamos a importância da mobilização do campo da educação, do posicionamento e participação dos(as) especialistas, dos(as) professores(as) e dos demais envolvidos no processo educativo na construção de políticas de avaliação que contemplem as especificidades da educação infantil.

Assim, as propostas em construção no CAEd/UFJF não pretendem defender uma forma única ou padronizada de avaliação da educação infantil e reconhecem as possibilidades de avaliação do contexto, das condições da oferta, da infraestrutura e das práticas pedagógicas da e na educação infantil, compreendendo a avaliação da criança na educação infantil como uma competência do(a) professor(a) nas interações cotidianas no interior das instituições públicas e privadas.

REFERÊNCIAS

BECCHI, Egle; BONDIOLI, Anna; FERRARI, Mônica. **ISQUEN**: indicadores e escala de avaliação da qualidade educativa da creche. In: CIPOLLONE, Laura (Org.). Instrumentos e indicadores para avaliar a creche: um percurso de análise da qualidade. Curitiba: Editora UFPR, 2014. p. 149-189.

BONDIOLI, Anna; FERRARI, Monica (a cura di). **AVSI – Autovalutazione della Scuola dell’Infanzia: uno strumento di formazione e il suo collaudo**. S. Paolo: Edizioni Junior, 2008.

BRASIL. **Constituição da República Federativa do Brasil**: promulgada e publicada em 5 de outubro de 1988. Brasília, DF: Senado, 1988.

_____. Lei nº. 9.394, de 20 de dezembro de 1996. **Estabelece as diretrizes e bases da educação nacional**. Diário Oficial [da] República Federativa do Brasil, DF, 23 de dez. 1996.

_____. Ministério da Educação. Secretaria de Educação Básica. **Parâmetros Nacionais de Qualidade para a Educação Infantil**. Brasília: MEC/SEB, 2006.

_____. Ministério da Educação. Secretaria de Educação Básica. **Parâmetros básicos de infra-estrutura para instituições de Educação Infantil**. Brasília: MEC/SEB, 2006.

_____. Ministério da Educação. Secretaria de Educação Básica. **Diretrizes Curriculares Nacionais para a Educação Infantil / Secretaria de Educação Básica**. – Brasília: MEC/ SEB, 2009.

_____. Ministério da Educação. Secretaria de Educação Básica. **Indicadores da Qualidade na Educação Infantil**. Brasília: MEC/SEB, 2009.

_____. Ministério da Educação. **Educação infantil: subsídios para construção de uma sistemática de avaliação**. Brasília, DF: MEC/SEB, 2012.

_____. Ministério da Educação, UNICEF, UNDIME, Ação Educativa. **Monitoramento do uso dos Indicadores da Qualidade na Educação Infantil**: Relatório Técnico. Brasília, DF: MEC/SEB, 2013.

_____. Presidência da República. Lei nº 13.005/2014. **Aprova o Plano Nacional de Educação - PNE e dá outras providências**. Brasília, 2014.

_____. Ministério da Educação. **Relatório com estudos e discussões da Comissão Assessora de Especialistas da Avaliação da Educação Infantil**. Brasília, DF: MEC/SEB, 2015.

_____. Lei 13.257, de 08 de março de 2016. **Dispõe sobre as políticas públicas para a primeira infância**. Brasília: MEC, 2016.

_____. Ministério da Educação. Conselho Nacional de Educação. Conselho Pleno. Parecer CNE/CP no 15/17. **Base Nacional Comum Curricular**. Brasília, 2017.

HARMS, Thelma; CLIFFORD, Richard M. **Early childhood environment rating scale (ECERS)**. Chapel Hill: Frank Porter Graham Child Development Center, University of North Carolina, 1980.

HARMS, Thelma; CRYER, Debby; CLIFFORD, Richard M. **Infant/toddler environment rating scale (ITERS)**. Chapel Hill: Frank Porter Graham Child Development Center, University of North Carolina, 1990.

HARMS, Thelma. O uso de escalas de avaliação de ambientes na educação infantil. **Cadernos de Pesquisa**, São Paulo, v. 43, n. 148, p. 76-97, jan./abr. 2013. Tradução de Akemi Kamimura e revisão técnica de Beatriz Abuchaim.

JANNUZZI, Paulo de Martino. **Indicadores Sociais no Brasil**. 6 ed. Campinas, SP: Editora Alínea, 2017.

ROSEMBERG, Fúlvia. Políticas de educação infantil e avaliação. **Cadernos de Pesquisa**, São Paulo, v. 43, n. 148, p. 44-75, jan./abr. 2013.

Recebido em: 06 de dezembro de 2019

Aceito em: 18 de junho de 2020

Publicado em: 30 de junho de 2020

A RESTITUIÇÃO COMO PROCESSO CENTRAL NA AVALIAÇÃO DE CONTEXTO

Catarina Moro¹

¹ Professora da associada ao Departamento de Teoria e Prática de Ensino e ao Programa de Pós-Graduação em Educação da Universidade Federal do Paraná, Curitiba, Paraná, Brasil. E-mail: moro.catarina@gmail.com

Resumo

A intenção do presente artigo é discutir os desafios e potencialidades da “restituição” a partir da abordagem avaliativa de contexto. Nesta abordagem avaliativa, restituir implica uma função específica do formador, que também atua como avaliador externo e constitui-se uma etapa do processo. Entende-se a restituição numa lógica negocial, participativa, formativa; na antinomia à lógica excludente e classificatória, muitas vezes associada à escola e à avaliação, potencializando a participação democrática, a emancipação e o fortalecimento da profissionalidade docente.

Palavras-chave: Avaliação de contexto. Restituição. Avaliação formativa. Participação. Educação Infantil.

Abstract

The purpose of this paper is to discuss the challenges and potentialities of “restitution” from the context evaluative approach. In this evaluative approach, to retribute implies a specific function of the trainer, who also acts as an external evaluator and constitutes a stage in the process. Restitution is understood in a negotiate, participatory, and formative logic; in the antinomy to the exclusionary and classifying logic often associated with school and assessment; enhancing democratic participation, emancipation and the strengthening of teaching professionalism.

Keywords: Context Evaluation. Restitution. Formative evaluation. Participation. Childhood Education.

“Nós não aprendemos com a experiência.
Nós aprendemos refletindo sobre a experiência.”

John Dewey

CONTEXTUALIZANDO A PROPOSIÇÃO DO DEBATE

A avaliação da qualidade da educação infantil vem se tornando tema pujante no Brasil, tanto nas proposições políticas, quanto nos debates e pesquisas acadêmicas nos últimos dez a quinze anos e, recentemente, vem demonstrando ainda maior vigor. Uma indicação dessa maior afluência foi a presença significativa em cursos, mesas-redondas, comunicações orais, painéis e na conferência de encerramento da X Reunião da Associação Brasileira de Avaliação Educacional (ABAVE), realizada em 2019, que teve como tema “Avaliação Educacional no Brasil: O Desafio da Qualidade”². Fúlvia Rosemberg, em 2013, já havia enfatizado a demarcação da avaliação para esta etapa educacional na realidade brasileira como um “novo” problema social, nas suas articulações a diferentes âmbitos – o sujeito, a instituição, o sistema e a política ou as políticas educacionais³. Cabe pontuar a existência de tensionamentos entre as divergentes visões e abordagens avaliativas presentes no cenário nacional, hoje, seja em relação à avaliação da oferta educativa, seja em relação à do desempenho das crianças.

Anteriormente, em 2009, saiu a primeira edição do documento nacional, “Indicadores da Qualidade da Educação Infantil”⁴, que difunde uma proposta de autoavaliação institucional para as unidades educativas, com vistas a efetivação do processo para se refletir sobre a qualidade educativa ali ofertada e propor melhorias. Ainda que, a difusão do documento não tenha alcançado boa parte das instituições públicas de educação infantil no país; a vinda a público dessa proposta fez com que alguns municípios, principalmente as capitais, afrontassem o debate a respeito e do mesmo modo passassem a elaborar e publicizar documentos semelhantes e se colocassem no movimento de efetivação de proposições de avaliação de suas redes. Em meio às ações - de debate, elaboração de indicadores próprios e efetivação de processos

² Ver Programa e Resumos em: https://www.abave.org.br/xabave/programacao_completa.pdf

³ Rosemberg (2013, p.55) discute as questões éticas implicadas e faz, ainda, um alerta e uma denúncia quanto às pesquisas voltadas a avaliar diferentes aspectos, quesitos ou dimensões relacionados às crianças e que tomam as instituições de educação infantil como “ ‘celeiro de sujeitos’ para pesquisas de diversas áreas de conhecimento”.

⁴ O documento é fruto da parceria entre o Ministério da Educação (MEC), pela Coordenação-Geral da Educação Infantil (COEDI)/Secretaria da Educação Básica (SEB), Ação Educativa, Fundação Orsa, União Nacional dos Dirigentes Municipais da Educação (Undime), Fundo das Nações Unidas para a Infância (Unicef).

avaliativos locais - inúmeros têm sido os resultados ou efeitos dessa prática, a depender do modo como o município entende e propõe o uso seja de documento próprio ou do documento nacional.

O artigo em questão discute uma proposição específica de avaliação de contexto, dando ênfase a uma das suas ações-meio e se apresenta organizado em três partes. Inicialmente expõe as concepções de base distintas à abordagem da avaliação de contexto proposta pelo grupo *pavese*⁵. Em seguida, tematiza questões relativas ao valor da restituição feita em mediação pelo formador/avaliador externo. E, à guisa de considerações finais, enfatiza-se a centralidade dos momentos colegiados de restituição como fundante dos percursos de avaliação de contexto pelo caráter autorreflexivo e formativo predisposto aos envolvidos; além de indicar os desafios e dificuldades, quando ainda não se tem uma cultura avaliativa mais disseminada nessa perspectiva.

QUESTÕES DE BASE SOBRE A AVALIAÇÃO DE CONTEXTO

A discussão a ser empreendida aqui se ancora em uma abordagem específica de avaliação e tem como foco serviços de educação e cuidado de crianças até seis anos. Esta abordagem tem sido especificada como avaliação de contexto e ultimamente tem sido cunhada por Donatella Savio e Anna Bondioli por avaliação como “promoção interna”, “promoção a partir do interno” ou ainda “promoção a partir do interior”. Trata-se de uma modalidade autoavaliativa, melhor dizendo, auto e heteroavaliativa, realizada por avaliadores internos (auto) e externos (hetero); sendo, respectivamente, membros da equipe profissional das unidades educativas (podendo incluir a participação de familiares das crianças matriculadas e da comunidade do entorno) e, um ou mais especialistas não vinculados às unidades em avaliação – pesquisadores ligados às universidades, coordenadores pedagógicos do município ou da região, que assumem um tríptico papel, pois, além de avaliador, atuam como mediadores e formadores do grupo implicado no processo (BONDIOLI, 2015; BONDIOLI; SAVIO, 2015; BONDIOLI; SAVIO, 2013).

Desde as primeiras leituras e estudos sobre esta abordagem e da realização do projeto “Formação da Rede em Educação Infantil: Avaliação de Contexto”, com a pesquisa multicampos, realizada entre 2013 e 2015, tem-se buscado discutir sobre sua fundamentação teórico-metodológica, seus pressupostos e

⁵ Grupo de professores e pesquisadores que tiveram anteriormente ou se mantém vinculados a *Università Degli Studi di Pavia*, Itália.

princípios, as questões procedimentais implicadas. A pesquisa referida possibilitou visibilizar e retomar a discussão da abordagem *pavese* no Brasil (BRASIL, 2015; SOUZA; MORO; COUTINHO, 2015)⁶, em termos de avaliação institucional; permitindo abrir uma nova interlocução para pesquisadores e técnicos, em seus âmbitos de intervenção próprios, o da academia e o da proposição e execução de políticas públicas na área. Cabe realçar que essa abordagem oportuniza e potencializa processos de “autorreflexão colegiada” (RUIZ, 1996).

Pode-se encontrar em vários escritos de Savio e Bondioli (dos quais destaco: BONDIOLI; SAVIO, 2015, 2013; SAVIO, 2013) e, também em trabalhos nossos (MORO; COUTINHO, 2018; COUTINHO; MORO; BALDEZ, 2017; SOUZA; MORO; COUTINHO, 2015), com maiores detalhes, o que se reitera a seguir, a partir de perguntas, como elementos de base e distintivos da abordagem de avaliação de contexto, do grupo de pesquisa da Universidade de Pavia.

O que essa abordagem entende por contexto educativo?

É um sistema aberto, interdependente a outros sistemas, que sofre as influências da qualidade das interrelações que estabelece e às quais está sujeito. Para além dos aspectos físicos e materiais (espaço, infra-estrutura, mobiliário, localização etc.), o contexto educativo se faz das dimensões relacionais e simbólicas, definidas e redefinidas dinamicamente, recíproca e continuamente. Fundamentada na perspectiva sistêmica (*bronfenbrianna*), se interessa em reconhecer a instituição de educação infantil (no caso) na relação com os pais e demais familiares, com a gestão da unidade, da rede ou do município e, com os aspectos conjunturais mais amplos – social, político e cultural. “O entrelaçamento desses aspectos no acontecimento dinâmico e ‘vivo’ da situação educacional faz parte do que chamamos de ‘contexto’” (SAVIO, 2011, p. 51, tradução minha). Decorre dessa compreensão, a ideia de “identidade educacional”, a ser reconhecida ou construída pelo coletivo: “Quando os sujeitos (...) se questionam sobre o assunto, tentam traduzi-lo em pensamentos e palavras, o tornam explícito e, portanto, comunicável, pode-se dizer que afrontam a definição da “identidade educacional” desse contexto” (SAVIO, 2011, p. 51, tradução minha).

⁶ O Projeto em questão foi coordenado pelo Núcleo de Estudos e Pesquisas em Infância e Educação Infantil da Universidade Federal do Paraná (NEPIE/UFPR), em parceria com a Coordenação Geral de Educação Infantil, do Ministério da Educação brasileiro (COEDI/MEC), com outras três universidades brasileiras (Universidade Federal de Minas Gerais, Universidade Federal do Rio de Janeiro e Universidade do Estado de Santa Catarina) e com a *Università degli studi di Pavia/Itália*.

⁷ No original: “*l'intrecciarsi di questi aspetti nell'accadere dinamico e 'vivo' della situazione educativa fa parte di ciò che abbiamo chiamato 'contesto'.*”

⁸ No original: “*Quando i soggetti (...), si interrogano su di esso, provano a tradurlo in pensieri e parole, lo rendono esplicito e perciò comunicabile, si può dire che affrontano la definizione dell'“identità educativa”.*”

Colocar-se no processo de avaliação de contexto permite identificar e tornar-se consciente da identidade educacional, singular, da instituição educativa, que mesmo não intencionalmente é constituída e explicitada pela equipe profissional de cada instituição.

Qual concepção de qualidade se articula com a avaliação de contexto?

Pensar sobre uma conceituação de qualidade em relação à abordagem avaliativa aqui em discussão significa despojar-se das ideias de padronização, uniformidade, generalização, objetivação, entre outras, que queiram e possam indicar uma qualidade ideal, reconhecível e universal quanto à oferta educativa para as crianças pequenas.

De acordo com Dahlberg, Moss e Pence (2003), a partir dos anos de 1990, houve um aumento de posicionamentos críticos a respeito e passou-se a problematizar a ideia de uma qualidade única e ideal a ser atingida. Com isso, fortaleceu-se uma visão crítica que compreende a qualidade como um construto social, para o qual é importante sua definição *in loco*, por aqueles que estão envolvidos na oferta educativa de determinada instituição ou rede, ou seja, por aqueles sujeitos que a fazem no dia a dia.

Nessa abordagem, a qualidade como uma entidade estática, definitiva e universal não existe. Concorde-se com a perspectiva de que se trata de um construto subjetivo, relativo e dinâmico, baseado nos valores de uma dada comunidade, sendo, por isso, complexo, plural, valorado e contextualizado espacial e temporalmente.

Em face dessa perspectiva, a qualidade educativa adotada aqui é negocial, intersubjetiva, transativa, reflete valores e crenças, necessidades e escolhas, influências e responsabilidades de um determinado grupo em um momento específico (BONDIOLI; SAVIO, 2013a).

Quais princípios orientam e se articulam à avaliação de contexto?

Tendo em consideração as ideias sobre contexto educativo e qualidade, demarcadas pela interdependência, subjetividade e dinâmica próprias da situação educacional, destacam-se e apresentam-se aqui, como princípios dessa abordagem, a:

- a) **participação democrática e co-responsabilização** – ao longo do trabalho avaliativo, requer-se e ao mesmo tempo se promove o diálogo, a expressão dos pontos de vista de todos aqueles que integram o gru-

di quel contesto”.

po. Importa assim, fomentar o confronto de ideias entre estes participantes, deflagrar um processo amplamente e fortemente participativo, a fim de manifestar-se um compromisso comum e recíproco de todos como co-responsáveis no desafio conjunto de definir e implementar um processo de avaliação da qualidade para fazê-la avançar (GARIBOLDI, 2015). Os princípios de participação e de co-responsabilização estão na base da metodologia de “promoção interna”. A dinâmica do trabalho com o grupo, realizada pela mediação do avaliador externo e formador, ao mesmo tempo permite e requer a construção de um contexto democrático, atento às diferentes vozes dos participantes. Assim, reafirma-se, a partir da co-autoria em trabalho anterior (COUTINHO; MORO; BALDEZ, 2017), a compreensão acerca da participação democrática como um direito que permite a realização de outros direitos e de se “obter justiça, influenciar os resultados e denunciar abusos de poder” (TOMÁS, 2006, p. 207). Daí a necessidade em dar voz aos diferentes participantes, a fim de favorecer que possam tomar parte dentro do grupo para além de fazer parte dele (BORDENAVE, 1992), participando no e do grupo de maneira ativa.

- b) negociação** – quando um grupo de pessoas (representantes de diferentes segmentos da comunidade institucional) se envolve no processo de avaliação de contexto, interessa a explicitação e o compartilhamento dos pontos de vista de cada participante, assim como a discussão acerca destes, no intuito de se definir consensualmente sobre os pontos fortes e críticos da instituição, em seus aspectos estruturais e processuais, que melhor possam revelar uma boa qualidade educativa. A negociação implica a preservação das identidades individuais em meio ao fortalecimento do sentido de pertencimento ao grupo e à instituição, com as escolhas e valores colegiados, conciliados (COUTINHO; MORO; BALDEZ, 2017). A diversidade de ideias e de interesses deve ser entendida como um recurso que fomenta o caráter transacional, em termos deweyanos, dos diferentes pontos de vista em relação à qualidade (BONDIOLI, 2004).
- c) autorreflexividade** – implica uma investigação com procedimentos sistemáticos de análise da realidade educacional compartilhada pelos participantes, no sentido de que aquele que avalia está de fato avaliando a si mesmo. O grupo envolvido no processo avaliativo realiza uma autoanálise, uma autorreflexão com vistas a criar, a trazer inovações. O trabalho realizado pelo grupo implica que este se dê conta, adquira consciência acerca de tudo que surgiu no processo e assu-

ma para si a responsabilidade pela intervenção a partir dos pontos e questões críticas identificadas (GARIBOLDI, 2015); por isso, pode ser entendido como processo de pesquisa-ação. Ruiz (1996), ao se referir à ideia de “autorreflexão colegiada”, fala da importância de criar uma cultura colaborativa no interior da comunidade educativa, o que é sempre desafiador perante situações de individualismo exacerbado (que, contraditoriamente, podem ser muito comuns no campo da educação); assim como, de acolher concepções, pontos de vista, opiniões pessoais e profissionais, distintas que podem tornar-se provocadoras de mudanças.

d) formação, transformação, fortalecimento – a avaliação da qualidade educativa em contexto perspectiva elaborar e aperfeiçoar ideias dentro do grupo de trabalho, criar repertórios compartilhados, propiciar aos participantes o envolvimento em aprendizados transformadores. A avaliação está orientada a mudanças do contexto institucional para melhor, através da elaboração de planos ou projetos inovadores e factíveis para aquela realidade específica. Ao mesmo, permite àqueles que dela participam uma expansão da consciência sobre o fazer profissional. Portanto, pode constituir-se autotransformadora, abrindo possibilidades a cada participante em responsabilizar-se eticamente com seu trabalho. E, conseqüentemente, buscar um agir mais intencional em relação ao fazeres e saberes implicados, refletindo um fortalecimento profissional que não é somente individual, mas, sobretudo, coletivo. (GARIBOLDI, 2015; BONDIOLI; SAVIO, 2013; BONDIOLI, 2004).

Conseqüentemente, importa afirmar que tais princípios, anteriormente destacados, fundamentam e asseguram um percurso comprometido com uma avaliação de contexto participativa, formativa e transformadora. Cabe também asseverar que cada percurso é singular e sua estruturação deve ser planejada a partir das especificidades de cada contexto.

Qual é o lugar do avaliador interno nessa abordagem avaliativa?

Considerando os princípios recém-mencionados, notabiliza-se o avaliador interno como elemento chave do processo, desde o início. Parte-se do pressuposto de que seu envolvimento seja de caráter volitivo, o que exige ter em conta que sua participação no processo não pode ser compulsória, imposta por outrem. Mesmo que a intenção de implementar o processo de avaliação tenha partido da gestão (da instituição ou da rede), ao convocar a participação

dos avaliadores internos, é imprescindível mobilizar, nos futuros integrantes do processo avaliativo, a vontade e o desejo de empenhar-se nessa ação.

Cada participante, no papel de avaliador interno, será demandado a agir de modo “protagônico” (CUSSIÁNOVICH; FIGUEROA, 2005; CUSSIÁNOVICH, 2009): –explicitando seus pontos de vista, ideias, julgamentos, sentimentos, concepções (explícitas ou latentes) e, exemplificando-os com situações concretas e reais da prática educativa cotidiana; – ouvindo os demais participantes de modo atento, respeitoso e curioso, independentemente de concordar ou não com o conteúdo manifesto pelos demais partícipes;– observando e analisando criticamente sua própria realidade com base no instrumento avaliativo escolhido para uso no processo, que foi anteriormente estudado por ele (avaliador interno) e debatido no Grupo.

A restituição, principalmente aquela que acontece nos Grupos de Trabalho (GT)⁹, logo após a observação e a avaliação feitas com base no instrumento, é um momento relevante do percurso e implica a participação ativa e protagônica dos avaliadores internos, como será detalhado mais à frente.

E o que faz o avaliador externo?

O avaliador externo tem, entre outras responsabilidades, a de promover a experiência de avaliação na sua integralidade, começando com a proposição do instrumento (um ou mais) a ser utilizado no processo. Na sequência, cabe a ele propiciar o debate e as reflexões formativas em meio ao processo avaliativo, seja na etapa que precede a observação da realidade ou avaliação propriamente ditas, que implica a leitura, o estudo e a compreensão do instrumento ou instrumentos escolhidos; como na etapa pós-observação, da restituição, na

⁹Grupo de Trabalho (GT) é o nome dado ao coletivo que atua em conjunto no processo avaliativo. Via de regra é composto por um ou mais avaliadores externos/formadores e pelos profissionais da instituição avaliada que aderem ao processo – professores, coordenadores, demais funcionários. Em cada instituição, pode haver um ou mais grupos de trabalho. A tarefa desse coletivo é deflagrada por uma Plenária Inicial (fase 1), na qual o ou os avaliadores externos apresentam a proposta e o ou os instrumentos avaliativos para possível uso a todo o coletivo da instituição que potencialmente virá a participar do percurso. Nessa ocasião, os interessados ficam com a tarefa de ler e analisar o conteúdo dos instrumentos (que pode ser por intermédio de um questionário - fase 2). Depois se organizam momentos de encontro para todos aqueles que aderiram ao processo, para discussão do ou dos instrumentos, a fim de compreender seu conteúdo e dinâmica de uso (fase 3). Essa é a ocasião da consolidação do GT, que não pode ter mais de 20 participantes, sendo mais interessante que seja menor. Pode ser necessário, em determinada instituição, que se forme mais de um GT. A ação seguinte incorre em um período para observação das turmas/grupamentos de crianças, entrevistas e consulta à documentação. Ao longo de uma semana, cada profissional (avaliador interno) faz a observação correspondente à sua turma e pontua os itens/indicadores presentes no instrumento, incluindo as considerações que julgar necessárias. Em concomitância, o avaliador externo faz o mesmo ao longo da jornada de um dia dessa mesma semana (fase 4). São feitos os encontros de restituição com o ou os GTs. Restituem-se os resultados do processo avaliativo, estimulando reflexões em relação aos itens/indicadores da(s) área(s), avaliadas debatendo sobre os julgamentos discordantes (fase 5). Plenária final. Avaliação do processo vivenciado. E elaboração de um plano de ação (fase 6, ou se for o caso, separando se teria as fases 6, 7 e 8).

qual viabiliza a confrontação das percepções individuais, dos julgamentos da qualidade observada e das pontuações atribuídas.

Compete ao heteroavaliador assegurar uma avaliação que não seja auto-referenciada, a ele cabe o olhar de estranhamento àquela realidade. O papel de avaliador externo é, via de regra, exercido por um profissional da área e especializado nas singularidades e amplitudes do trabalho de cuidar e educar crianças com menos de 6 anos. Nesse sentido, pode ser um pesquisador e professor universitário, um coordenador pedagógico municipal, da mesma rede ou de outra rede municipal, ou de outra instituição da região, não podendo ser o coordenador pedagógico do contexto em avaliação.

Na sequência, ao adentrar as especificidades da restituição, retomam-se mais uma vez os papéis de ambos os tipos de avaliadores participantes do processo, com destaque para a atuação do participante externo, heteroavaliador.

O PAPEL FULCRAL DA RESTITUIÇÃO

A avaliação de contexto, como “promoção a partir do interno”, deriva da referência anterior denominada pelo grupo de Pavia de “modelo de avaliar-restituir-refletir-inovar-avaliar” (BONDIOLI; FERRARI, 2004), no qual o momento da restituição é fundante e crucial para o percurso como um todo.

Becchi (2000, p.205), ao tratar da ética na pesquisa educacional, deu destaque ao ato de restituir como sendo

um retorno àqueles que lhes ofereceram as informações que tornaram possível a investigação. Um retorno de dados elaborados, cheios de significado, organizados, que é uma discussão em conjunto, pesquisadores e provedores, para trocar observações e pontos de vista de maneira transativa; uma dialectização de saberes, abordagens, experiências, modos de ver o mundo, em um forte enriquecimento da pesquisa¹⁰ (Tradução minha).

De modo similar, restituir no percurso avaliativo consolida o processo, designando-o como legítimo, útil, justo, numa perspectiva que preza pela transparência, rigor, precisão, disponibilidade e publicização dos resultados em vista dos impactos que pode suscitar. E ,ainda, corrobora o direito dos envolvidos de conhecer as finalidades, as conclusões e os encaminhamentos decorren-

¹⁰ No original: “ un ridare a coloro che le hanno offerte le informazioni che hanno reso possibile l'indagine. Un ridare dati elaborati, densi di significato, organizzati, che è un discuterli assieme, ricercatori e erogatori, per scambiarsi, in modo transattivo, osservazioni, punti di vista; un dialettizzare saperi, approcci, vissuti, modi di vedere il mondo, in un forte arricchimento della ricerca”.

tes da concretização de uma dada experiência avaliativa. A restituição subentende questões éticas, diplomáticas e políticas do processo de avaliação no seu todo.

Ferrari (2004), no livro organizado por Anna Bondioli e por ela intitulado “*Verso um Modelo di Valutazione Formativa – Ragioni, strumenti, percorsi*”, dedica um capítulo exclusivamente ao tema. Entre as questões tratadas pela autora, denota-se a ideia da complexidade envolta na restituição das informações recolhidas durante o processo avaliativo, entre outras, por demandar uma articulação entre a análise dos resultados e as ações, em confronto com as mudanças de uma realidade em constante movimento e modificação e com a cultura e as expectativas dos diferentes envolvidos e destinatários.

Em termos operativos, na prática, a restituição ocorre em uma ou mais reuniões do Grupo de Trabalho, após a avaliação propriamente dita – realizada por ambos os avaliadores utilizando o instrumento escolhido, na qual os participantes individualmente observam a realidade de cada turma/grupamento dentro da instituição em cotejamento com o descrito no instrumento.

O avaliador externo/formador sustenta e promove a restituição a partir da proposição de discussões e reflexões em Grupo. Para isso, segundo Bondioli e Savio (2015, 2013, 2009) e Savio (2011) ele:

- » solicita a expressão das ideias de cada um, buscando um equilíbrio entre a introspecção e a exteriorização por parte dos integrantes no interior do Grupo;
- » propõe-se a uma escuta atenta e compreensiva dos diferentes participantes, no sentido rogersiano;
- » reapresenta, no Grupo, e aos participantes, individualmente, suas manifestações, no sentido de refratar o foi disposto, repetindo o que foi dito para promover uma visão distanciada, descentrada do que ele (participante) externou;
- » destaca e sintetiza o que foi dito, com intuito de que os participantes aprofundem e ampliem o que referiram;
- » convida os participantes a exemplificarem, com situações do cotidiano educativo, determinada ideia ou questão pedagógica e, também, em sentido inverso, pede para explicitarem os significados das práticas educativas em ideário pedagógico;
- » promove a articulação e o aprofundamento dos conteúdos expressos no Grupo;
- » evidencia questões colocadas, opiniões contrastantes, deduções equivocadas, no sentido de promover reflexões e debates mais aprofundados e provocar e solicitar maior coerência;
- » propõe articulações e conexões entre ideias diferentes que têm uma mesma base conceitual, demandando negociação entre pontos de vista diferentes ou articulados entre si.

O heteroavaliador/formador precisa estabelecer com os participantes do GT uma relação maiêutica. Para Dolci (1996) a perspectiva é de uma maiêutica bidirecional, na qual se acentua o envolvimento recíproco de escuta autêntica, o encontro dialético entre diferenças e semelhanças e a civilidade e participação democráticas alargadas que promovem a emancipação e a justiça social.

Os encontros de restituição pressupõem a apresentação, pelo avaliador externo, da pontuação atribuída por cada avaliador (os internos e os externos) e do patamar de qualidade correspondente. A partir dessa confrontação, o coletivo analisa a diversidade de pontos de vista, os julgamentos, independentemente de serem dissidentes ou concordantes entre si; desencadeando um circuito recursivo que faz com que os participantes adquiram maior consciência acerca do fazer e porque fazem no tocante ao trabalho na educação infantil. De tal modo que acaba por se constituir

um novo patrimônio de dados que não se relaciona apenas às pontuações assinaladas pelos diferentes avaliadores (externos e/ou internos) com base em um instrumento, mas que se desvenda no encontro de discussão e que consiste nas palavras com as quais cada um explica o processo que o levou a atribuir determinadas pontuações, exprimindo uma imagem da realidade educativa que tem, muitas vezes, razões implícitas. Trata-se de temas que chamam a atenção de todos os envolvidos no evento da restituição, a questão dos valores e da negociação dos significados latentes do fazer¹¹ (FERRARI, 2004, p.62 – tradução minha).

É, em face desta compreensão acerca da restituição, que se coloca a oportunidade de ativar trocas culturais entre os envolvidos durante o percurso avaliativo, para aprender no confronto das ideias e percepções divergentes; para explicitar o implícito da comunicação cotidiana através do diálogo que o formador/avaliador evoca e solicita, num processo dialético-dialógico e; para reorientar as escolhas com base numa negociação sinérgica no Grupo, apoiada pelo formador/avaliador. Sordi e Lüdke (2009), sobre a proposição da avaliação institucional participativa referem que:

¹¹No original: " un nuovo patrimonio di dati che non è relativo solo ai punteggi assegnati dai diversi valutatori (esterni e/o interni) sulla base di uno strumento, ma che si dipana nell'incontro di discussione e che consiste nelle parole con cui ciascuno rende conto del processo che lo ha condotto ad assegnare certi punteggi, esprimendo un'immagine della realtà educativa che ha ragioni sovente implicite. Si tratta di temi che riportano all'attenzione di tutti coloro che sono coinvolti nell'evento della restituzione la questione dei valori e della negoziazione dei significati latenti del fare".

Ao orientar a reflexão do grupo sobre alguns indicadores escolares, a avaliação ajuda a recuperar os referenciais de qualidade que estão subjacentes ao processo de significação dos dados, fazendo emergir contradições, hipóteses, vazios, avanços, indispensáveis ao processo de tomada de decisão rumo ao desenvolvimento qualitativo do projeto pedagógico (SORDI; LUDKE, 2009, p. 324).

(...)

Fortalecidos em seus saberes, os professores ganham voz e vez nos processos de qualificação da escola, aprendendo a construir uma rede explicativa para os problemas com que se deparam e a formular também estratégias de superação que lhes façam sentido ou a contestar aquilo que se mostre inconsistente, embora formalmente instituído (SORDI; LUDKE, 2009, p. 326).

A restituição atua nesse fortalecimento, por se caracterizar como um processo permanente de co-construção de significados, conhecimentos e saberes, no sentido deweyano de transação, a fim de pactuar, de se chegar a um acordo, de construir um saber compartilhado. Tal possibilidade coaduna-se com a visão de Tardif (2012) sobre constituir-se um confronto de e com saberes para a construção de novos saberes. Para este autor, fica evidenciado que é por meio “das relações com os pares e, portanto, através do confronto entre os saberes produzidos pela experiência coletiva dos professores, que os saberes experienciais adquirem certa objetividade”, uma vez que as certezas subjetivas são sistematizadas em um discurso “capaz de informar ou formar outros docentes e de fornecer uma resposta aos seus problemas” (TARDIF, 2012, p.52).

Nessa acepção, as trocas durante a restituição no GT, pela atitude maiêutica por parte do formador, permitem maior compreensão e atribuição de objetividade aos diferentes saberes, uma sua retradução, fruto da análise crítica entre os saberes advindos da prática e o conteúdo presente no ou nos instrumentos e mobilizados na restituição. Como consequência, ao rever e refletir sobre o que se faz e como se faz, compreender por que se faz se pode aprimorar não somente as práticas, mas a própria profissão, numa dinâmica de fortalecimento profissional, individual e coletivo.

ALINHAVANDO ALGUNS PONTOS DO QUE FOI AQUI APRESENTADO

Buscar um ampliar a discussão e melhorar o entendimento, assim como, empreender processos de avaliação de contexto precisa constituir-se um compromisso diante da ausência de uma cultura avaliativa com base no paradigma participativo, negocial e formativo. Concorde-se com Rosemberg (2013, p. 52), quando explicita não ser concebível acatar qualquer concepção de avaliação “que não seja ética e iluminadora para tomada de decisões socialmente justas e democráticas.” Em consonância com esse agravo e diante do tratado anteriormente acerca da avaliação de contexto, especialmente da restituição, ressalta-se a potencialidade dessa proposição pelo compromisso com a participação e a autorreflexão colegiada acerca das práticas educativas e das condições estruturais para o trabalho pedagógico e por oferecer oportunidades formativas implicadas com a instituição.

REFERÊNCIAS

BECCHI, Egle. Per un’etica della ricerca educativa: il paradigma della restituzione. In: L. CORRADINI (a cura di). **Pedagogia: ricerca e formazione**. Formello, Roma: SEAM, 2000.

BONDIOLI, Anna. “Promuovere dall’interno”: un’estensione dell’approccio del “valutare, riflettere, restituire”. In: BONDIOLI, Anna; SAVIO, Donatella. (a cura di). **La valutazione di contesto nei servizi per l’infanzia italiani. Riflessioni ed esperienze**. Parma: Junior-Spaggiari, 2015. p. 25-49.

BONDIOLI, Anna. Valutare. In: BONDIOLI, Anna; FERRARI, Monica. (a cura di). **Verso un modello di valutazione formativa**. Edizioni Junior: Azzano San Paolo (Bg), 2004, p. 11-58.

BONDIOLI, Anna; SAVIO, Donatella (a cura di), **La valutazione di contesto nei servizi per l’infanzia italiani**. Riflessioni ed esperienze. Parma: Junior-Spaggiari, 2015.

BONDIOLI, Anna; FERRARI, Monica. (a cura di). **Verso un modello di valutazione formativa**. Ragioni, strumenti e percorsi. Azzano San Paolo (Bg): Edizioni Junior, 2004.

BONDIOLI, Anna; SAVIO, Donatella. Elaborar indicadores de qualidade educativa das instituições de educação infantil: uma pesquisa compartilhada entre Itália e Brasil. In: SOUZA; Gizele; MORO, Catarina; COUTINHO, Angela Scalabrin (orgs.). **Formação da rede em Educação Infantil: avaliação de contexto**. Curitiba: Appris, 2015, p.21-49.

BONDIOLI, Anna; SAVIO, Donatella. **Participação e qualidade em educação da infância**: percursos e compartilhamento reflexivo em contextos educativos. Trad. Luiz Ernani Fritoli. Curitiba: Editora UFPR, 2013.

BONDIOLI, Anna; SAVIO, Donatella. **Uma abordagem reflexiva e dialógica à avaliação de contextos educativos (Parte I)**. Slides (material digitalizado). Curitiba. Não publicado, 2013a.

BONDIOLI, Anna; SAVIO, Donatella. Formare i formatori: un approccio maieutico. In: DOMENICI, G; SEMERARO, R. (a cura di). **Le nuove sfide della ricerca didattica tra saperi, comunità sociali e culture** Roma: Monolite, 2009, p. 467-479.

BORDENAVE, Juan Díaz. **O que é Participação**. São Paulo: Brasiliense, 1992.

BRASIL. **Contribuições para a Política Nacional**: a avaliação em educação infantil a partir da avaliação de contexto. Curitiba: Imprensa/UFPR; Brasília, DF: MEC/SEB/COEDI, 2015.

COUTINHO, Angela Scalabrin; MORO, Catarina; BALDEZ, Ettiene Barbosa. Avaliação de Contexto em Educação Infantil: a participação e o protagonismo docente. **RELADEI**, v. 6, 2017.

CUSSIÁNOVICH, Alejandro; FIGUEROA, Elvira. Participación protagónica: ideología o cambio de paradigma? In: LIEBEL, Manfred; MARTÍNEZ MUÑOZ (coord.). **Infancia y Derechos Humanos** – Hacia una Ciudadanía participante y protagónica. Lima: Ifejant, 2009, p. 83-99.

CUSSIÁNOVICH, Alejandro. Participación ciudadana de la infancia desde el Paradigma del protagonismo. In: II **Congreso Mundial de Infancia Adolescencia “Ciudadanía Desde la Niñez y Adolescencia y Exigibilidad de sus Derechos**. Lima, Peru, 2005.

DAHLBERG, Gunilla; MOSS, Peter; PENCE, Allan. **Qualidade na Educação da Primeira Infância** – Perspectivas pós-modernas. Porto Alegre: Artmed, 2003.

DOLCI, Danilo. **La struttura maieutica e l'evolgerci**. Firenze: La Nuova Italia, 1996.

FERRARI, Monica. Restituire. In: BONDIOLI, Anna; FERRARI, Monica. (a cura di). **Verso un modello di valutazione formativa**. Edizioni Junior: Azzano San Paolo (Bg), 2004, p. 59-91.

GARIBOLDI, Antonio. Co-costruire la valutazione: le linee guida sulla valutazione della qualità in Regione Emilia Romagna. In: BONDIOLI, Anna; SAVIO, Donatella (a cura di), **La valutazione di contesto nei servizi per l'infanzia italiani**. Riflessioni ed esperienze. Parma: Junior-Spaggiari, 2015, p. 165-173.

MORO, Catarina; COUTINHO, Ângela Scalabrin. Avaliação de Contexto como Processo Formativo. **Cadernos de Pesquisa em Educação**. PPG. UFES, v. 20, p. 90-112, 2018. Disponível em: <http://www.periodicos.ufes.br/?journal=educacao&page=article&op=view&path%5B%5D=21333&path%5B%5D=14204>. Acesso em: 14 jun. 2019.

ROSEMBERG, Fúlvia. Políticas de educação infantil e avaliação. **Cadernos de Pesquisa**. [online] vol.43, n.148, pp.44-75, jan./abr. 2013 Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742013000100004&lng=en&nrm=iso. Acesso em: 01 ago. 2019

RUIZ, José Maria. **Cómo Hacer Una Evaluación de Centros Educativos**. Madrid: Narcea, 1996.

SAVIO, Donatella. A brincadeira e a participação da criança: um desafio educativo e seus pontos nodais. In: BONDIOLI, Anna; SAVIO, Donatella. **Participação e qualidade em educação da infância**: percursos e compartilhamento reflexivo em contextos educativos. Trad. Luiz Ernani Fritoli. Curitiba: Editora UFPR, 2013, p. 243-303.

SAVIO, Donatella. **Il gioco e l'identità educativa del nido d'infanzia**. Un percorso di valutazione formativa partecipata nei nidi di Modena. Bergamo: Junior, 2011.

SORDI; Mara Regina Lemes; LUDKE, Menga. Da avaliação da aprendizagem à avaliação institucional: Aprendizagens necessárias. **Avaliação**. Campinas; Sorocaba, SP, v. 14, n. 2, p. 253-266, jul. 2009.

SOUZA, Gizele; MORO, Catarina; COUTINHO, Angela Scalabrin. **Formação da rede em educação infantil**: avaliação de contexto. Curitiba: Appris, 2015.

TARDIF, Maurice. **Saberes docentes e formação profissional**. 14 ed. Petrópolis, RJ: Vozes, 2012.

TOMÁS, Catarina Almeida. **Há muitos mundos no mundo... Direitos das Crianças, cosmopolitismo Infantil e Movimentos Sociais de Crianças** – diálogos entre crianças de Portugal e Brasil. 2006. Tese (Doutorado em Estudos da Criança), Área Sociologia da Infância, Instituto de Estudos da Criança. Braga: Universidade do Minho, 2006.

Recebido em: 17 de dezembro de 2019

Aceito em: 15 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÕES DA EDUCAÇÃO INFANTIL EM MUNICÍPIOS PAULISTAS: LIMITES E POTENCIALIDADES PARA CONTRIBUIR COM A GARANTIA DO DIREITO À EDUCAÇÃO DAS CRIANÇAS PEQUENAS

Cláudia Oliveira Pimenta¹

¹Vínculo profissional e contato/e-mail da autora: Pesquisadora na Fundação Carlos Chagas (FCC); e-mail: cpimenta@fcc.org.br.

Resumo

Este artigo analisa iniciativas de avaliação da educação infantil de dois municípios paulistas, evidenciando suas potencialidades e limites para contribuir com a garantia do direito à educação das crianças, à luz do arcabouço legal e de produção acadêmica que trata da qualidade para essa etapa. As iniciativas se aproximam de indicações dessas referências, mas, em parte, reproduzem aspectos das avaliações e práticas pedagógicas destinadas às demais etapas educacionais, o que implica as finalidades da Educação Infantil.

Palavras-chave: Educação Infantil. Direito à educação. Avaliação da qualidade.

Abstract

This paper analyzes initiatives of the evaluation of early childhood education in two municipalities of São Paulo state, highlighting their potentials and limits to contribute with guaranteeing the right to education children, in the light of the legal framework and academic production that deals with quality for this stage. The initiatives are close to indications of these references, but in part they reproduce aspects of evaluations and pedagogical practices aimed at other educational stages, which implies the finalities of early childhood education.

Keywords: Early Childhood Education. Right to education. Quality evaluation.

INTRODUÇÃO

A educação infantil foi incorporada ao Sistema de Avaliação da Educação Básica (SAEB)² pela Portaria nº 366, de 29 de abril de 2019, tendo sido reafirmada pela Portaria nº 458, de 5 de maio de 2020, após vários anos de debates e embates teórico-metodológicos em torno do tema, atendendo ao arcabouço legal vigente no Brasil desde a promulgação da Constituição Federal de 1988 (BRASIL, 1988), em especial, a Lei 13.005, de 24 de junho de 2014, que estabeleceu o Plano Nacional de Educação (PNE 2014-2024) que prevê, na Meta 1, destinada à essa etapa educacional, a implantação da avaliação da Educação Infantil, a cada dois anos, “com base em parâmetros nacionais de qualidade, a fim de aferir a infraestrutura física, o quadro de pessoal, as condições de gestão, os recursos pedagógicos, a situação de acessibilidade, entre outros indicadores relevantes” (BRASIL, 2014 – Meta 1, estratégia 1.6 do PNE 2014-2024), tendo em vista a universalização da pré-escola e a ampliação do atendimento em creches.

Todavia, vale destacar diferenças entre as duas portarias que, em diferentes graus, carecem de maiores esclarecimentos em relação ao desenho da avaliação destinada à Educação Infantil.

A primeira Portaria, de 2019, estabelecia que, no âmbito do SAEB, as informações sobre a educação infantil deveriam ser coletadas junto aos dirigentes de secretarias estaduais e municipais de educação e aos profissionais que atuam nas instituições públicas de educação infantil, por meio dos questionários contextuais do Sistema, de forma censitária.³ As escolas privadas teriam a prerrogativa de aderir ou não à avaliação. Este documento legal não apresentava informações relativas ao foco da avaliação, ainda que, de acordo com informações divulgadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), no Seminário Internacional Estatísticas Educacionais e Avaliação da Educação Básica⁴, realizado em 2018, as dimensões sobre as quais seriam levantados dados diriam respeito à: condições de acesso e oferta; condições de infraestrutura; insumos e recursos materiais; profissionais da educação infantil; gestão dos sistemas educacionais; gestão das unidades escolares; e acessibilidade. No seminário foi assinalado, ainda, que não era prevista a avaliação das crianças.

² Ver Portaria nº 366, de 29 de abril de 2019, que estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica (SAEB) no ano de 2019. Diário Oficial da União, publicado em 02/05/2019, edição 83, Seção 1, página 47. Órgão Responsável: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

³ Em 2019, a coleta foi realizada em uma amostra de instituições, devendo se tornar censitária apenas em 2021. Ver Portaria nº 366, de 29 de abril de 2019.

⁴ Disponível em: <https://www.youtube.com/watch?v=xgSdJrgKac8>. Acesso em 24 ago. 2020.

A Portaria de 2020, que trata da Política Nacional de Avaliação da Educação Básica, menciona a Educação Infantil uma única vez e nos seguintes termos:

Art. 4º O Saeb é um sistema composto por um conjunto de instrumentos que permite a produção e a disseminação de evidências, estatísticas, avaliações, exames e estudos a respeito da qualidade das etapas que compõem a educação básica, que são: I - a educação infantil; II - o ensino fundamental; e III - o ensino médio (BRASIL, 2020, p. 57).

O documento não explicita nenhuma outra informação sobre a educação infantil e “não demarca atributos concernentes a essa etapa educacional e à sua avaliação no âmbito das políticas públicas de educação” (ANPed, 2020)⁵. A não definição de como a avaliação da educação infantil deve ocorrer, parece contribuir para potencializar disputas teórico-metodológicas já existentes, especialmente no que tange à concepção de qualidade que deve ser assumida pela avaliação da etapa no Brasil. Tais disputas, nas duas últimas décadas, se traduziram em ações desencadeadas pelo governo federal e por governos municipais, bem como por instituições da sociedade civil, com pelo menos duas distintas vertentes: i) uma que privilegia dimensões estruturais, com foco na oferta, nos insumos e processos; ii) outra que toma o desenvolvimento/aprendizagem das crianças, em determinados aspectos e/ou áreas do conhecimento, como expressão da qualidade oferecida (SOUSA, 2014; 2018; PIMENTA, 2017).

Também ações concernentes aos anos iniciais do ensino fundamental, como a aplicação de provas do SAEB para o 2º ano⁶, a reformulação do Pacto Nacional pela Alfabetização na Idade Certa (PNAIC),⁷ que incorporou a pré-escola, a partir de 2017 e a nova Política Nacional de Alfabetização (PNA)⁸, instituída em abril de 2019 e que substituiu o PNAIC, são exemplos de iniciativas com potencial para implicar ações na educação infantil.

No caso do extinto PNAIC, a indicação de ações de formação continuada, com foco na aprendizagem das crianças, para coordenadores pedagógicos e professores que atuam na educação infantil, e o emprego de avaliações voltadas para a alfabetização dos alunos dos anos iniciais do ensino fundamental – ainda que não houvesse menção explícita à aplicação para as crianças da

⁵ Disponível em: <https://anped.org.br/news/manifesto-gt-07-avaliacao-da-educacao-infantil-e-nova-portaria-da-politica-nacional-de>. Acesso em 24 ago.2020.

⁶ O SAEB para o 2º ano era denominado Avaliação Nacional da Alfabetização (ANA), até 2018.

⁷ Portaria n. 826 de 07 de julho de 2017, publicada pelo Ministério da Educação.

⁸ Decreto nº 9.765, de 11 de abril de 2019.

educação infantil – constituíam-se em ações com potencial de induzir instituições e redes educacionais a começarem a efetuar avaliações das crianças, nos moldes do que se vê para as demais etapas da educação básica.

A PNA avança nessa proposição, ao enfatizar, em seu artigo 6º, item I, que “as crianças da primeira infância” compõem o público-alvo da referida política e assinala o “incentivo à elaboração e à validação de instrumentos de avaliação e diagnóstico” (BRASIL; PRESIDÊNCIA DA REPÚBLICA, 2019. Decreto nº 9.765, de 11 de abril de 2019 - Artigo 8º, item XIII), bem como o “incentivo à difusão tempestiva de análises devolutivas de avaliações externas e ao seu uso nos processos de ensino e de aprendizagem” (Idem, 2019 - Artigo 9º, item II).

A aprovação e implementação da Base Nacional Comum Curricular (BNCC) é outra ação que tem sido vista como capaz de estimular a realização de avaliações das crianças e tomar seus resultados como expressão da qualidade oferecida, dado que um de seus propósitos é construir o alinhamento com outras políticas educacionais, dentre elas, a avaliação:

[...] a BNCC integra a política nacional da Educação Básica e vai contribuir para o alinhamento de outras políticas e ações, em âmbito federal, estadual e municipal, referentes à formação de professores, à avaliação, à elaboração de conteúdos educacionais e aos critérios para a oferta de infraestrutura adequada para o pleno desenvolvimento da educação (BRASIL, 2019, p. 8).

Nesse contexto de disputas quanto à abordagem avaliativa a ser assumida pela avaliação da educação infantil, se desenvolveu a pesquisa, da qual decorre o presente artigo, “Avaliações municipais da educação infantil: contribuições para a garantia do direito à educação das crianças brasileiras?”, que analisou proposições de avaliação de 42 municípios paulistas, dado que essa esfera federativa se constitui como a principal responsável pela oferta da etapa, no Brasil.

Estudos e pesquisas anteriores já vinham evidenciando a adesão de municípios a ambas as vertentes avaliativas aqui apresentadas, principalmente da avaliação voltada para o desenvolvimento e/ou aprendizagem das crianças (ADRIÃO et al, 2009; NASCIMENTO, 2012; BAUER et al, 2016; RIBEIRO, 2016). Também estudos publicados no mesmo ano da pesquisa supracitada (BARROS, 2017; CANÇADO, 2017), com a preocupação de investigar, analisar e compreender como avaliações da educação infantil, elaboradas e implementadas por/em municípios, vêm sendo construídas e indicam resultados semelhantes, o que evidencia a preponderância de uma concepção de qualidade atrelada ao currículo.

Acompanhando essa tendência na pesquisa por nós desenvolvida, os resultados obtidos evidenciaram que a maior parte dos municípios analisados utiliza avaliações cujo enfoque recai sobre o desenvolvimento/aprendizagem das crianças. Entretanto, em 18 deles, observou-se a concomitância com outros focos, como: infraestrutura física e de materiais dos estabelecimentos educacionais; e/ou organização do ambiente educativo; e/ou trabalho pedagógico ou ações/proposta pedagógica; e/ou profissionais da educação; e/ou ações da Secretaria Municipal de Educação.

Para efeito deste artigo, são apresentados os desenhos de avaliação de dois municípios, que congregam características como as acima anunciadas, com o objetivo de evidenciar potencialidades e limites de suas abordagens avaliativas para contribuir com a garantia do direito à educação, pública e com qualidade, para as crianças de zero a cinco anos. Esses municípios são Birigui e Taubaté, localizados no interior do estado de São Paulo. A seguir, estão registrados os procedimentos metodológicos da coleta de dados do estudo, a caracterização social e educacional dos municípios citados, a caracterização e análise de seus desenhos avaliativos e considerações finais acerca do objeto em tela.

PROCEDIMENTOS METODOLÓGICOS DA COLETA DE DADOS

O levantamento dos dados ocorreu entre os meses de novembro de 2015 e fevereiro de 2016, via e-mail enviado às secretarias municipais de educação de 102 municípios paulistas que declararam, em pesquisa conduzida pela Fundação Carlos Chagas (FCC) e pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)⁹, em 2014, possuir avaliação da educação infantil.

O e-mail enviado a estes 102 municípios paulistas, continha três questões que versavam sobre : i) quem eram os formuladores da avaliação da educação infantil; ii) qual o foco dessas avaliações; iii) quais os instrumentos/procedimentos utilizados para avaliar a etapa. Deste total de municípios, 42 responderam a essas questões e, parte deles forneceu, também, informações sobre usos dos resultados das avaliações, materiais e documentos orientadores de suas práticas educacionais, instrumentos e/ou procedimentos avaliativos e matérias jornalísticas sobre as avaliações.

⁹ A pesquisa, conduzida pela FCC e pelo Inep, mapeou propostas de avaliação da educação básica, desenvolvidas por municípios brasileiros, no ano de 2014. A autora recebeu autorização de ambas as instituições para utilização dos dados da pesquisa.

Por meio de busca nos *sites* das secretarias e das prefeituras, foram coletados documentos complementares, como, por exemplo, os Planos Municipais de Educação (PME) aprovados a partir do estabelecimento do PNE 2014-2024, que prediz a avaliação da educação infantil como uma das estratégias para a universalização da pré-escola e ampliação da oferta em creches, como informado na introdução deste texto.

Além dessa documentação, foram coletadas informações em bancos de dados nacionais com vistas à dimensionar a responsabilidade dos municípios com a educação básica, quais sejam: indicadores populacionais, educacionais e de desenvolvimento humano disponíveis nos *sites* do Instituto Brasileiro de Geografia e Estatística (IBGE), do Inep e do Programa das Nações Unidas para o Desenvolvimento (PNUD).

As referências utilizadas para a consecução das análises dos desenhos avaliativos foram compostas por: produções acadêmicas nacionais que tratam da qualidade da educação infantil; legislação e documentos norteadores da educação infantil, em âmbito nacional; e, ainda, por experiências nacionais e estrangeiras de avaliação, voltadas para essa etapa.

O exame dos documentos e informações relativas aos municípios estudados, bem como das referências utilizadas para analisar os desenhos avaliativos por eles propostos, foi efetuado por meio de análise documental e de conteúdo (BARDIN, 1977). No primeiro caso, para selecionar documentos e informações de modo a compor um quadro-síntese da situação da educação infantil, nesses municípios, e de sua avaliação. No caso das referências de análise, para elucidar movimentos e tendências na configuração das avaliações voltadas para a educação infantil, possíveis contribuições para a garantia do direito à educação das crianças pequenas e aspectos de qualidade que atendem a esse direito.

No Quadro 1 estão registrados os direitos e aspectos de qualidade mapeados na pesquisa, à exceção do conteúdo extraído das experiências estrangeiras de avaliação.

Quadro 1 – Direitos das crianças e aspectos de qualidade na/da educação infantil, extraídos da literatura analisada

Direito das crianças	Aspectos de Qualidade
<p>O acesso à creche, para as crianças de zero a três anos de idade;</p> <p>A vaga na pré-escola, garantida por legislação que estabelece tanto a universalização, como a obrigatoriedade do atendimento às crianças de quatro e cinco anos;</p> <p>Aprender por meio do jogo e da brincadeira;</p> <p>Ter seu contexto social e cultural respeitado;</p> <p>Interagir com outras crianças, com os adultos e com o meio ambiente;</p> <p>Desenvolver-se integralmente, em seus aspectos físico, psicológico, intelectual e social;</p> <p>Participar, junto com suas famílias, das tomadas de decisão quanto ao seu processo de desenvolvimento e de aprendizagem;</p> <p>Ter seu desenvolvimento registrado e documentado pela instituição de educação infantil, ao longo do tempo;</p> <p>Ser cuidada;</p> <p>Ser alimentada.</p>	<p>Existência de ambientes físicos adequados, com espaços externos e internos bem equipados;</p> <p>Garantia de colaboração entre os entes federados para investimentos na etapa;</p> <p>Garantia de formação específica e condições de trabalho adequadas para os profissionais da educação;</p> <p>Disponibilização de quantidade e qualidade de recursos materiais para as instituições;</p> <p>Melhoria dos processos de gestão administrativa e pedagógica das instituições;</p> <p>Elaboração de um currículo que prime pelo desenvolvimento integral das crianças;</p> <p>Desenvolvimento de práticas pedagógicas que considerem a criança como sujeito e que levem em conta as diretrizes existentes para a educação infantil;</p> <p>Estímulo à escuta das vozes e experiências trazidas pelas crianças;</p> <p>Garantia de proporção crianças/professor adequada;</p> <p>Estímulo à participação das famílias na vida da instituição;</p> <p>Realização de processos de avaliação e autoavaliação participativos das instituições;</p> <p>Atenção às noções de qualidade educacional dos profissionais das instituições;</p> <p>Atenção aos processos de avaliação da aprendizagem e do desenvolvimento das crianças, cuidando para que sejam formativos, não classificatórios e sem fins de promoção;</p> <p>Integração entre o educar e o cuidar, tanto na creche quanto na pré-escola;</p> <p>Adequação do financiamento da educação infantil;</p> <p>Investimento na construção de instituições educacionais pela rede pública direta;</p> <p>Atendimento aos critérios de conveniência indicados pela legislação;</p> <p>Garantia de condições de acessibilidade, física e material, para crianças/profissionais portadores de deficiências;</p> <p>Garantia de alimentação escolar para todas as crianças;</p> <p>Elaboração e monitoramento do planejamento escolar, de forma reflexiva e sistemática;</p> <p>Realização de processos de avaliação e autoavaliação dos órgãos intermediários e centrais da política educacional.</p>

Fonte: Pimenta (2017, p. 609-611).

CARACTERÍSTICAS SOCIAIS E EDUCACIONAIS DOS MUNICÍPIOS DE BIRIGUI E TAUBATÉ

Os municípios de Birigui e Taubaté foram considerados, em 2010, pelo Índice de Desenvolvimento Humano Municipal (IDHM), medido pelo PNUD, municípios com desenvolvimento alto ou muito alto, respectivamente. Em ambos, a variável que mais contribuiu para o resultado obtido foi a longevidade, seguida da renda e, por fim, da educação. No Quadro 2 estão detalhadas as informações.

Quadro 2 – Resultados obtidos no IDHM 2010, por Birigui e Taubaté

Município	IDHM	Longevidade	Renda	Educação
Birigui	0,780	0,869	0,743	0,734
Taubaté	0,800	0,883	0,778	0,746

Fonte: Elaborado pela autora.

Em relação à população, no mesmo ano, Birigui possuía 108.728 habitantes e Taubaté, 235.884. Já em 2015, a população de cada um deles foi estimada em 118.352 e 302.331 habitantes, respectivamente, de acordo com dados do IBGE.

Segundo o Censo Escolar de 2015, 24.313¹⁰ alunos estavam matriculados em escolas sediadas em Birigui, sendo 10.800 (44,42%) na rede municipal de educação. Em Taubaté, havia 78.252¹¹ alunos matriculados em escolas do município, no referido ano. Destes, 45.428 (58,05%) estavam sob a responsabilidade da rede municipal de educação (V. tabela 1).

Tabela 1 – Número de matrículas por etapa e/ou modalidade de ensino, nas redes municipais de educação de Birigui e Taubaté

Etapas e modalidade	Birigui	Taubaté
Educação infantil – creche	2.334	6.649
Educação infantil – pré-escola	2.306	7.519
Anos iniciais EF	6.083	16.553
Anos finais EF	-	13.357
Modalidade EJA	77	285
Ensino médio	-	715
Educação profissional	-	350
Total	10.800	45.428

Fonte: INEP/Sinopse Estatística da Educação Básica, 2015.

¹⁰ Estão incluídas, nesse total, as matrículas regulares da educação infantil e as matrículas da educação especial, dos alunos que frequentam as classes comuns.

¹¹ Idem nota 6.

Ainda que o número de estudantes, sob responsabilidade da rede municipal de educação de Taubaté, seja quatro vezes maior que o de Birigui, ambos atendem a um grande contingente de alunos, em termos percentuais, quando comparados com o total de alunos matriculados em escolas sediadas no município, em que se incluem escolas pertencentes a outras dependências administrativas (estadual, federal e privada).

O Documento Preparatório do Plano Municipal de Educação de Taubaté, elaborado em 2015, registra que os estabelecimentos de educação infantil contam com a seguinte proporção de crianças por profissional da educação, atuante nas turmas:

Berçário I: Não há professor para esta faixa etária, apenas ADI (Auxiliar de Desenvolvimento Infantil), com média de três profissionais por turma de 16 alunos (5,3 alunos por profissional).

Berçário II: Um professor e uma ADI por turma de 18 alunos, em média (9 alunos por profissional).

Educação Infantil: 14,7 alunos por profissional, a partir dos níveis Maternal e Jardim (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015, p. 134).

O Censo Escolar 2015 registra que Birigui possuía 490 docentes atuantes na educação infantil, sendo 335 em creches e 165 na pré-escola. Já em Taubaté, havia 698, com 373 na creche e 427 na pré-escola¹² (V. tabela 2).

Tabela 2 – Número de docentes atuantes na educação infantil em 2015, Birigui e Taubaté

Município	Rede municipal		Escolas privadas		Total
	Creche	Pré-escola	Creche	Pré-escola	
Birigui	283	135	58	30	490
Taubaté	214	259	162	168	698

Fonte: INEP/Sinopse Estatística da Educação Básica, 2015.

¹² Os docentes são contados uma única vez em cada município. Portanto, o total não representa a soma das etapas de ensino, pois o mesmo docente pode atuar em mais de uma unidade de agregação. (Observação adaptada da Nota 5 da planilha 2.6 da Sinopse do Censo Escolar 2015).

No caso de Taubaté, o documento supracitado assinala que a rede municipal de educação contava, no referido ano, com 349 docentes trabalhando em regime estatutário¹³, 40 substitutos em regime estatutário e 102 contratados via Consolidação das Leis Trabalhistas (CLT). O documento também registra que 339 professores, atuantes na educação infantil, eram licenciados em 2015, e 219 possuíam pós-graduação *Lato Sensu* (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015).

Quanto ao número de estabelecimentos educacionais de educação infantil¹⁴, o Censo Escolar de 2015 indica a existência de 39 sediados no município de Birigui e 128 em Taubaté (V. tabela 3).

Tabela 3 – Número de estabelecimentos de educação infantil em 2015, Birigui e Taubaté

Município	Rede municipal		Escolas privadas		Total de estabelecimentos
	Creche	Pré-escola	Creche	Pré-escola	
Birigui	19	21	13	09	39
Taubaté	69	49	71	52	128

Fonte: INEP/Sinopse Estatística da Educação Básica, 2015.

No caso de Birigui, o *site* da Secretaria de Educação registrava que a rede municipal possuía 32 estabelecimentos educacionais que atendiam a educação infantil. Destes, nove pertenciam à rede direta, e seis eram conveniados ao poder público municipal. Todos atendiam crianças de zero a cinco anos de idade, em período integral. Havia, ainda, uma Escola Municipal de Educação Infantil (EMEI) destinada apenas à pré-escola e que atendia crianças desde os três anos de idade até cinco anos e, ainda, 16 Escolas Municipais (EM) que atendiam a educação infantil e o ensino fundamental – anos iniciais (com crianças dos quatro aos 10 anos). Em duas dessas unidades, os alunos eram atendidos em período integral.¹⁵

Em Taubaté, de acordo com o Documento Preparatório do PME, a rede municipal mantém 78 estabelecimentos educacionais que oferecem educação infantil, sendo cinco conveniados e 73 pertencentes à rede direta (62 destinados apenas à etapa e 11 que atendem, também, o ensino fundamental). De acordo com o documento, as instituições privadas que oferecem educação

¹³ Os trabalhadores, em regime estatutário, são concursados.

¹⁴ Informa-se que cada estabelecimento pode atender a mais de uma etapa ou nível de ensino.

¹⁵ Disponível em: <http://www.birigui.sp.gov.br/educacao/site/reparticoes/reparticoes.php>. Acesso em 23/ mar. 2016.

infantil totalizam 46. Tanto as instituições conveniadas como os 62 estabelecimentos da rede direta oferecem educação em tempo integral (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015).

As informações concernentes aos estabelecimentos educacionais, em ambos os municípios, denotam diferentes arranjos para a oferta da educação infantil: atendimento na rede direta e em instituições conveniadas ao poder público; estabelecimentos educacionais específicos para o atendimento à etapa e estabelecimentos que atendem, conjuntamente, a educação infantil e o ensino fundamental.

No caso de Birigui, há estabelecimentos que atendem as crianças em período integral e, outras, em período parcial. Em Taubaté, a rede direta e as instituições conveniadas atendem as crianças em período integral. Em relação às faixas etárias atendidas, em Birigui observa-se a matrícula de crianças de três anos na pré-escola (idade que diz respeito à creche).

Merece atenção, especificamente no caso de Birigui, o fato de crianças de três anos estarem matriculadas na pré-escola, ainda que, de acordo com o PME do município, seja apenas em uma unidade. De acordo com a legislação brasileira, crianças de zero a três anos de idade devem frequentar a creche e não a pré-escola, assim como crianças de quatro e cinco anos devem estar matriculadas na pré-escola e não no ensino fundamental, como indicado pela Resolução nº 2, de 09 de outubro de 2018, do Conselho Nacional de Educação (CNE):

Art. 2º **A data de corte etário vigente em todo o território nacional**, para todas as redes e instituições de ensino, públicas e privadas, **para matrícula inicial na Educação Infantil aos 4 (quatro) anos de idade**, e no Ensino Fundamental aos 6 (seis) anos de idade, é aquela definida pelas Diretrizes Curriculares Nacionais, ou seja, respectivamente, aos 4 (quatro) e aos 6 (seis) **anos completos ou a completar até 31 de março do ano em que se realiza a matrícula.**

[...]

Art. 3º [...]

§ 2º **É obrigatória a matrícula na pré-escola**, segunda etapa da Educação Infantil e primeira etapa da obrigatoriedade assegurada pelo inciso I do art. 208 da Constituição Federal, **de crianças que completam 4 (quatro) anos até o dia 31 de março do ano em que ocorrer a matrícula inicial.**

§ 3º **As crianças que completam 4 (quatro) anos de idade após o dia 31 de março devem ser matriculadas em creches**, primeira etapa da Educação Infantil (CONSELHO NACIONAL DE EDUCAÇÃO – grifo nosso).

O aspecto acima mencionado, relativo ao atendimento da demanda para creches, é um aspecto que diz respeito à oferta da educação infantil e, sobre o qual, a política educacional precisa estar atenta.

No que tange à qualidade da oferta, de acordo com o que prediz o Plano Nacional de Educação e os vários documentos que norteiam a educação infantil no Brasil, faz-se necessário atender às especificidades da etapa. Turmas de educação infantil que funcionam em escolas onde é ofertado o ensino fundamental, por exemplo, necessitam ter garantidos infraestrutura física e de materiais, recursos pedagógicos e professores com formação adequados para o atendimento das crianças de zero a cinco anos.

Avaliar as condições da oferta de educação infantil na rede direta, independentemente se as turmas de educação infantil funcionam em espaço concomitante com o ensino fundamental ou não, bem como em instituições conveniadas, é de suma importância. De acordo com o documento “Educação Infantil: subsídios para a construção de uma sistemática de avaliação”, que se constituiu, em conjunto com outras produções, como inspiração para a formulação da perspectiva de avaliação da educação infantil, assumida pelo SAEB,

A avaliação dessas condições [de oferta], de modo sistemático, é condição inicial para alterar o quadro de precariedade que vem sendo evidenciado pelos diagnósticos disponíveis sobre a educação infantil, cabendo especial atenção para o modo como vem se dando o atendimento, não apenas em estabelecimentos públicos, como também em estabelecimentos conveniados com o Setor Público (BRASIL, 2012, p. 17).

Tomando em conta as características sociais e educacionais dos municípios de Birigui e Taubaté, sobretudo as relativas à educação infantil, no próximo item, são apresentadas a caracterização e análise de suas propostas de avaliação para essa etapa educacional.

CARACTERIZAÇÃO E ANÁLISE DAS INICIATIVAS DE AVALIAÇÃO DE BIRIGUI E TAUBATÉ

Birigui

O Plano Municipal de Educação (PME) de Birigui foi aprovado pela Lei nº 6.064, de 11 de agosto de 2015. Em relação à avaliação da educação infantil, esta define estratégia inspirada no PNE 2014-2024, mas sem detalhar os parâmetros nacionais de qualidade a serem considerados: “1.3. Implantar ou readequar em caso de necessidade, até o segundo ano de vigência deste PME, avaliação da educação infantil com base em parâmetros nacionais de qualidade e/ou outros indicadores relevantes” (PREFEITURA MUNICIPAL DE BIRIGUI; CÂMARA MUNICIPAL DE BIRIGUI, 2015, p. 4).

Outra estratégia, correlata à avaliação, também é prevista no plano e diz respeito ao acompanhamento do desenvolvimento do trabalho efetuado na educação infantil, com vistas a apoiar o ingresso das crianças na etapa seguinte:

1.7. Criar mecanismo de articulação entre a educação infantil e as séries iniciais do ensino fundamental, identificando as deficiências e necessidades no desenvolvimento do trabalho na educação infantil e seus resultados, subsidiando o ingresso do aluno para o ensino fundamental (PREFEITURA MUNICIPAL DE BIRIGUI; CÂMARA MUNICIPAL DE BIRIGUI, 2015, p. 5).

O respondente da pesquisa em Birigui assinalou que o município possuía “uma política de avaliação da educação infantil prevista em lei” (RESPONDENTE DE BIRIGUI, 2016), como resultado dos trabalhos de uma comissão responsável pela elaboração do estatuto e do plano de carreira do magistério biriguiense.

A elaboração da “política de avaliação” envolveu, segundo ele, professores da educação infantil e do ensino fundamental, coordenadores e diretores escolares e um representante da Secretaria Municipal de Educação, que realizaram estudos coletivos e promoveram ajustes, ao longo dos anos, “mediante portarias e instruções normativas expedidas pela Secretaria Municipal de Educação” (RESPONDENTE DE BIRIGUI, 2016). Os instrumentos de avaliação foram adaptados com base na percepção das escolas quanto à sua “efetividade”.

No primeiro contato da pesquisa com o município, o respondente informou que a avaliação ocorria anualmente, em todos os estabelecimentos de educação infantil, sob a coordenação dos técnicos da secretaria, com apoio dos profissionais das escolas e afirmou que tanto as instituições quanto as crianças eram avaliadas em dois eixos: dimensão processual (aprendizagem escolar e ambiente educativo) e dimensão institucional (avaliação dos pais ou responsáveis).

No que tange à *aprendizagem escolar*, a avaliação da educação infantil tinha como foco o nível de desenvolvimento das habilidades das crianças, tanto da creche como da pré-escola, nas seguintes áreas: linguagem oral e escrita; identidade e autonomia das crianças; e movimento. No caso da pré-escola, avaliava-se, ainda, a matemática. Em dezembro de 2016, fez-se necessário solicitar esclarecimentos ao respondente quanto ao formato da avaliação. Em sua resposta, informou que o aspecto *aprendizagem escolar* não seria mais avaliado, na educação infantil, assinalando que, desde agosto de 2016, a avaliação passou a contemplar apenas o “ambiente educativo”, na *dimensão processual* e manteve a aplicação de questionários aos pais para avaliar a “qualidade do processo educativo”, na *dimensão institucional*.

No que concerne ao *ambiente educativo*, o respondente afirmou que se avaliavam as interações e o estímulo à autonomia das crianças e o espaço de aprendizagem, observando se esse espaço “vai além da sala de aula, valorizando-se as rotinas e a organização do trabalho pedagógico que considera a criança como única”. A Instrução Normativa SME nº 002, de 02 de agosto de 2016 explicita o entendimento da Secretaria sobre o processo de avaliação do aspecto supracitado:

I - Ambiente Educativo: avaliação do ambiente educativo por turma/sala considerando-o como espaço de educação, aprendizagem e vivência de valores, no qual os indivíduos se socializam, brincam e experimentam a convivência com a diversidade humana. Neste eixo considerar-se-á a adequação da proposta e dos conteúdos, os agrupamentos, a organização do ambiente, a adequação dos recursos materiais e a interação como práticas fundamentais para a construção de um clima favorável para uma educação de qualidade, considerando a criança como única (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 2 - Mimeo).

Para avaliar a *dimensão institucional*, Birigui utiliza questionários que são aplicados aos pais, com o objetivo de captar “percepções” em relação à “qualidade do trabalho desenvolvido pelos profissionais que atuam diretamente na turma de seus filhos” (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 3). O respondente detalhou essa informação:

[Os questionários] avaliam o que os pais/responsáveis pensam em termos de qualidade do trabalho realizado pelo professor: como é o acolhimento do aluno na chegada e saída da escola; [se] o trabalho do professor contribuiu para que o aluno se desenvolvesse ao longo do ano (RESPONDENTE DE BIRIGUI, 2016).

A Instrução Normativa trata da avaliação da educação infantil de Birigui como componente do processo de avaliação e progressão na carreira dos profissionais docentes e não docentes que atuam na etapa, como registrado nos itens I e II da Instrução:

I. **Da fundamentação legal:** O processo de avaliação da Educação Infantil, concebido tendo em vista as disposições da Lei Complementar nº 32/2.010 e Portaria SME nº 27/2.011¹⁶ e alterações, compõe o sistema de Progressão Funcional Não-Acadêmica dos Profissionais do Quadro do Magistério e de Apoio Educacional do Município de Birigui e tem por princípios fundamentais à valorização dos profissionais da educação e a melhoria da qualidade do ensino público.

II. **Das avaliações da Educação Infantil:** A avaliação da rede municipal de ensino será instrumento relevante que permitirá aferir os níveis de qualidade do trabalho educativo da educação infantil, no que diz respeito às dimensões processual e institucional, tendo como referência as competências, objetivos e/ou habilidades constantes nos planos de ensino da rede municipal e Referenciais Curriculares Nacionais da Educação Infantil (RCNEIs)(SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 1 - mimeo).

Os resultados da avaliação, em conjunto com o Índice de Desenvolvimento da Educação Básica (IDEB) municipal, produzem um indicador para a educação infantil denominado Índice de Desenvolvimento da Educação Infantil (IDEI). Segundo o respondente, o IDEI influencia na progressão da carreira de todos os profissionais da educação infantil da rede municipal, conforme estabelece o Estatuto e Plano de Carreira do Magistério Público e dos Profissionais de Apoio Educacional do Município de Birigui, estabelecido pela Lei Complementar nº 32, de 17 de setembro de 2010, como se segue:

Artigo 104: A progressão funcional pela via não acadêmica dependerá da contagem de pontos dos fatores abaixo descritos:

[...]

III - Avaliação de resultado dos docentes:

[...]

¹⁶ Solicitamos a Portaria SME 27/2011 ao respondente da pesquisa, no município de Birigui. O documento não nos foi enviado, mas de acordo com o respondente, a referida portaria "regula a evolução funcional de várias categorias, repetindo o que está na parte de progressão funcional da Lei Complementar 32/2010", citada neste texto.

b) Aos docentes que atuam na Educação Infantil aplicar-se-ão os indicadores individuais de avaliação da Educação Infantil (IDEI), bem como da avaliação da unidade escolar ou os indicadores de avaliação externa municipais, estaduais e/ou nacionais.

[...]

V- Avaliação de resultado dos Profissionais de Apoio Educacional:

Às Babás aplicar-se-ão os indicadores individuais de avaliação da Educação Infantil (IDEI), bem como da avaliação da unidade escolar ou os indicadores de avaliação externa municipais, estaduais e/ou nacionais (PREFEITURA MUNICIPAL DE BIRIGUI. CÂMARA MUNICIPAL DE BIRIGUI, 2010, p. 32).

Em suma, os resultados da avaliação são utilizados no processo de progressão funcional dos profissionais da educação que atuam nos estabelecimentos educacionais e na Secretaria. A avaliação é efetuada “para as turmas do Berçário à Pré-escola, sob a regência de Babás, Educadores de Creche e Professores de Educação Infantil”, que recebem uma pontuação cuja “valoração máxima” é de “100 (cem) pontos” (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 1 – Mimeo). A metodologia utilizada no processo avaliativo e a pontuação aferida em cada dimensão/eixo, para as turmas regidas por estes profissionais, estão registradas no Quadro 3, retirado da Instrução Normativa SME 02/2016.

Quadro 3 – Metodologia e pontuação das dimensões/eixos da avaliação da educação infantil em Birigui. 2016

Dimensões	Eixos	Metodologia de avaliação	Pontuação	Pontuação final
Dimensão Processual (DP)	Ambiente Educativo	Preenchimento de ficha avaliativa do ambiente educativo por profissional indicado pela Secretaria de Educação	0 a 35 pontos	DP + DI
		Preenchimento de ficha avaliativa do ambiente educativo por gestor da unidade	0 a 20 pontos	
		Preenchimento de ficha avaliativa do ambiente educativo pelo(s) profissional (is) da turma	0 a 15 pontos	
Dimensão Institucional (DI)	Avaliação dos Pais / Responsáveis	Aplicação de questionário aos pais ou responsáveis por profissional indicado pela Secretaria de Educação, visando à coleta de informações sobre a qualidade do trabalho desenvolvido pelos profissionais que atuam diretamente na turma	0 a 30 pontos	

Fonte: Secretaria Municipal de Educação de Birigui (2016, p. 2 - Mimeo).

Na *dimensão processual*, os profissionais da escola que atuam nas turmas (Babás, Educadores de Creche e Professores de Educação Infantil) ou na gestão escolar (diretores e coordenadores) e a equipe técnica da Secretaria são avaliados em cinco indicadores: “adequação da proposta e dos conteúdos; agrupamentos; organização do ambiente; adequação dos recursos materiais; Interação”, de acordo com a Instrução Normativa 02/2016 (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 3 - Mimeo).

Depreende-se da análise da Instrução Normativa que, para cada indicador, é estabelecida uma pontuação, conforme o nível de enquadramento funcional do profissional, sendo que, no Nível I, a pontuação mais alta de cada grupo de profissionais (equipe técnica da Secretaria, gestores escolares e profissionais da turma), multiplicada pela quantidade de indicadores (cinco), resulta na maior pontuação indicada no Quadro 3. No Quadro 4, a seguir, estão registradas as pontuações para cada grupo de profissionais e seus respectivos níveis de enquadramento funcional.

Quadro 4 – Critérios de avaliação em cada nível de enquadramento funcional, na SME e no estabelecimento educacional. Birigui

Nível de enquadramento funcional	Critérios de Avaliação/Pontuação		
	Equipe técnica da SME	Gestores Escolares	Profissionais da turma
I	7,00 – 6,00	4,00	3,00
II	5,00 – 4,00	3,00	2,00
III	3,00 – 2,00	2,00	1,00
IV	1,00 – 0,00	1,00 – 0,00	0,50 – 0,00

Fonte: Secretaria Municipal de Educação de Birigui (2016, p. 3 - Mimeo).

O Estatuto e Plano de Carreira do Magistério Público e dos Profissionais de Apoio Educacional do Município de Birigui definem “Nível” como sendo o “enquadramento vertical automático do servidor, no ato de sua nomeação ou reenquadramento, e dispensados quaisquer interstícios de tempo, nos níveis correspondentes a sua formação acadêmica” (PREFEITURA DO MUNICÍPIO DE BIRIGUI; CÂMARA MUNICIPAL DE BIRIGUI, 2010, p. 3). O artigo 101 da referida lei, que trata da progressão funcional pela via acadêmica, apresenta os critérios para enquadramento em cada um dos níveis mencionados, como se segue:

- I - Curso de graduação, correspondente à licenciatura plena, na área de educação, excetuando-se o de pré-requisito para ingresso na carreira: 4(quatro) níveis;

II - Curso de pós-graduação em nível de especialização “lato sensu”, na área da educação, com duração mínima de 360 (trezentos e sessenta) horas: 4 (quatro) níveis;

III - Curso de pós-graduação “stricto sensu” em nível de mestrado, na área da educação: 5 (cinco) níveis;

IV - Curso de pós-graduação “stricto sensu” em nível de doutorado, na área da educação: 6 (seis) níveis (PREFEITURA MUNICIPAL DE BIRIGUI; CÂMARA MUNICIPAL DE BIRIGUI, 2010, p. 30).

A ausência dos profissionais no dia da realização da avaliação pode implicar a pontuação dos profissionais na *dimensão processual* (eixo *ambiente educativo*), como indica a Instrução Normativa SME 02/2016:

a. Nas turmas regidas por mais de 01 (um) profissional, ocorrendo a ausência de um deles no dia de avaliação por motivo de atestado médico, licença gestante, licença prêmio ou convocações obrigatórias do poder judiciário, sua pontuação será a mesma obtida pelos profissionais da turma.

b. Nas turmas regidas por apenas 01 (um) profissional, ocorrendo sua ausência no dia da avaliação por motivo de atestado médico, licença gestante, licença prêmio ou convocações obrigatórias do poder judiciário, será atribuída a média obtida pela unidade escolar.

c. Nas demais ausências por quaisquer outros motivos, o profissional deixará de ter atribuída a respectiva pontuação no eixo Ambiente Educativo. (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016, p. 3 - Mimeo).

Em relação à *dimensão institucional*, a Instrução orienta como os estabelecimentos educacionais devem organizar a aplicação e a recolha dos questionários dirigidos aos pais ou responsáveis das crianças: a) desenvolver estratégias que garantam o envio dos instrumentos para as famílias de cada turma e a devolução à equipe da Secretaria, nas datas estabelecidas; b) manter espaços organizados no interior da unidade, para que as famílias entreguem os questionários preenchidos; c) receber apenas os questionários entregues pelos próprios pais ou responsáveis, que devem assinar a lista de presença da turma (caso o documento seja entregue pela criança ou pessoa menor de idade e a unidade escolar aceitar esse procedimento, o questionário deve ser anulado) (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, 2016 - Mimeo).

O questionário contém cinco questões de múltipla escolha, iguais para todas as turmas. Em cada questão, os pais ou responsáveis, segundo a Instrução Normativa, devem assinalar ótimo, bom, regular ou ruim. Depreende-se que as referidas questões são elaboradas como afirmativas para as quais as famílias realizam as indicações acima referidas. Os apontamentos recebem uma pontuação, a depender da resposta dada: ótimo (6 pontos); bom (4 pontos); regular (2 pontos); e ruim (0 ponto).

Além da pontuação nas questões, as turmas recebem pontos pela quantidade de questionários respondidos, como se segue:

Fica exigido o retorno mínimo por turma de pelo menos 75% (setenta e cinco por cento) dos questionários respondidos pelos pais e/ou responsáveis para fins de apuração da nota plena obtida. Para a obtenção da pontuação das turmas, fixada em no máximo 30 (trinta) pontos, serão consideradas, além das pontuações obtidas nos questionários, as seguintes ponderações quanto às taxas de devolução dos questionários pelos pais e/ou responsáveis:

- a. Devolução igual ou superior a 75%: não haverá desconto;
- b. Devolução entre 65% e 74%: desconto de 1,00 (um) ponto na média obtida pela turma;
- c. Devolução entre 55 a 64%: desconto de 2,00 (dois) pontos na média obtida pela turma;
- d. Devolução abaixo de 55%: desconto de 3,00 (três) pontos na média obtida pela turma (SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI, p. 4 - Mimeo).

Por fim, a Instrução orienta que os pais ou responsáveis com filhos matriculados em diferentes turmas da educação infantil devem responder um questionário por turma. Para tanto, o gestor da unidade escolar deve incumbir-se de enviar a quantidade necessária de questionários para cada família.

O IDEI, ao qual nos referimos anteriormente, não é mencionado na Instrução Normativa enviada pelo respondente, no segundo contato da pesquisa com o município. Foi solicitado esclarecimento em relação a essa não menção ao índice, bem como sobre a composição do indicador, mas não se obteve resposta.

Em síntese, a iniciativa de avaliação da educação infantil de Birigui tem como finalidade avaliar o trabalho educativo realizado nos estabelecimentos educacionais e, por conseguinte, o desempenho dos profissionais que atuam na etapa.

A avaliação possui duas dimensões: *dimensão processual* que avalia o ambiente educativo (adequação da proposta e dos conteúdos, os agrupamentos, a organização do ambiente, a adequação dos recursos materiais e a interação), por meio de uma ficha avaliativa preenchida pelos técnicos da secretarias, e de uma ficha preenchida pelos profissionais dos estabelecimentos educacionais; e a *dimensão institucional*, onde as famílias respondem um questionário sobre a qualidade do trabalho desenvolvido pelos profissionais que atuam diretamente na turma de seus filhos.

Tendo sido construída no âmbito dos debates em torno da elaboração do estatuto e plano de carreira dos profissionais do magistério, os resultados da avaliação, em conjunto com o resultado do IDEI municipal, produzem um indicador para a educação infantil que é utilizado para determinar a progressão na carreira, pela via não acadêmica, de professores, babás, educadores de creche e equipe técnica da Secretaria Municipal de Educação. Não foram identificados, nos depoimentos e documentos analisados, menção a outros tipos de uso da avaliação da etapa, seja pela Secretaria, seja pelos estabelecimentos educacionais.

Taubaté

O Plano Municipal de Educação (PME) de Taubaté, aprovado pela Lei Complementar nº 392, de 15 de julho de 2016, no que tange à avaliação da educação infantil, baseia-se no que prediz o PNE e acrescenta especificidades locais. O conteúdo sobre a avaliação da etapa está exposto na estratégia 1.6 da Meta 1, relacionada à educação infantil:

Implantar, até o segundo ano de vigência deste PME, avaliação da educação infantil para escolas públicas e privadas, a ser realizada a cada 2 (dois) anos, com base em parâmetros nacionais de qualidade, a fim de aferir a infraestrutura física, o quadro de pessoal, as condições de gestão, os recursos pedagógicos, a situação de acessibilidade, entre outros indicadores relevantes; criar um protocolo de manutenção de qualidade para autorização/alvará de funcionamento de escolas particulares e públicas, prevendo a vistoria periódica dos espaços físicos (PREFEITURA MUNICIPAL DE TAUBATÉ; CÂMARA MUNICIPAL DE TAUBATÉ, 2016, p. 4).

No Documento Preparatório do PME de Taubaté, já citado neste texto, consta a informação de que o município utilizou os Indicadores da Qualidade na Educação Infantil (BRASIL, 2009)¹⁷, para aferir a qualidade das instituições de educação infantil nos anos de 2013 e 2014. Segundo o documento, “participaram desse processo: 446 professores, 338 funcionários e 1.835 membros da comunidade” (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015, p. 139).

De acordo com informações prestadas pela respondente da pesquisa, no município, a adoção do uso dos *Indicadores*, em Taubaté, tem como objetivo “acompanhar e supervisionar o trabalho pedagógico” (RESPONDENTE DE TAUBATÉ, 2016). Em 2013, a equipe de supervisão da Secretaria preparou encontros de formação para que gestores e professores coordenadores pudessem conhecer e estudar o material:

Os gestores e professores coordenadores, durante a formação, vivenciaram uma simulação dos procedimentos a serem adotados na escola, no dia da aplicação dos *Indicadores* junto à comunidade e receberam um kit com todas as orientações e o passo a passo, desde a reunião com a equipe escolar para o planejamento da aplicação e desenvolvimento do roteiro com as questões, até sugestões de elaboração dos planos de ações, cujos indicadores fossem sinalizados com as cores amarelo e vermelho; além da apresentação das boas práticas cujos indicadores fossem sinalizados com a cor verde (RESPONDENTE DE TAUBATÉ, 2016).

A respondente assinalou que, após esses estudos, docentes e gestores da educação infantil responderam ao roteiro de questões proposto pelos *Indicadores*. Em 2014, houve, segundo ela, a participação de “toda a comunidade escolar: docentes, funcionários, pais de alunos, agentes da comunidade, gestores, professores coordenadores” (RESPONDENTE DE TAUBATÉ, 2016). A equipe de supervisão da SME acompanhou o desenvolvimento do trabalho “e auxiliou no esclarecimento de dúvidas que surgiram ao longo do processo” (RESPONDENTE DE TAUBATÉ, 2016). Após a realização da autoavaliação institucional, o relatório individual de cada estabelecimento foi encaminhado para essa equipe, que “pôde comparar os dados dos anos de 2013 e 2014” (RESPONDENTE DE TAUBATÉ, 2016). A partir daí, foram elaborados os Planos de Ação de cada estabelecimento educacional.

No ano de 2015, segundo a respondente, foi realizado o “monitoramento da aplicação do plano de ação, com vistas à superação dos problemas identificados durante a aplicação dos *Indicadores*, com um cronograma a ser seguido

¹⁷ Os Indicadores da Qualidade na Educação Infantil foram publicados pelo MEC em 2009 e constituem-se em uma proposta de autoavaliação institucional, com foco no ambiente educativo.

prevendo as reuniões com a comunidade, para devolutiva” (RESPONDENTE DE TAUBATÉ, 2016). Para 2016, havia a previsão de uma “nova aplicação do roteiro, com quadro síntese para comparação dos resultados e avaliação das ações” (RESPONDENTE DE TAUBATÉ, 2016).

O Documento Preparatório do PME de Taubaté indicava a existência de diferentes movimentos referentes à avaliação educacional, no município. A Secretaria Municipal de Educação implantou, à época de sua publicação, 2015, um “Sistema de Avaliação Interno, como forma de acompanhamento do processo educacional desenvolvido nas escolas. Todos os níveis e modalidades são avaliados, e os resultados se tornam base para o planejamento de ações” (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015, p. 58), o que supõe a inclusão da educação infantil. O documento acrescenta:

A avaliação interna que se faz na rede deve ser vista como um instrumento orientador das políticas educacionais, sobretudo para a melhoria da qualidade da educação. Nesse sentido, a ampla divulgação dos resultados obtidos (globais e locais) para a comunidade é fundamental (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015, p. 59).

Em 2014, a Secretaria aplicou “um teste diagnóstico” às crianças, “para que servisse de ponto de partida para as futuras ações formativas junto aos professores”. Os resultados desse diagnóstico mostraram “a situação dos alunos em 63 escolas, quanto às hipóteses que formulam sobre a leitura e escrita” (PREFEITURA MUNICIPAL DE TAUBATÉ, 2015, p. 143).

A Secretaria, em conjunto com professores representantes da creche e da pré-escola, “elaborou dois instrumentos para o acompanhamento do desenvolvimento dos alunos: a ficha de acompanhamento individual do aluno (bimestral) e as atividades avaliativas comuns (semestral)” (RESPONDENTE DE TAUBATÉ, 2016).

As avaliações realizadas por meio desses instrumentos têm como propósito “o alcance de uma educação de qualidade, favorecendo o máximo possível o desenvolvimento das crianças de 0 a 5 anos e a promoção da melhoria do atendimento, atentando-se também, para a elaboração de ações que contemplem o aperfeiçoamento docente” (RESPONDENTE DE TAUBATÉ, 2016).

A Ficha de Acompanhamento Individual do Aluno tem por base “os Referenciais Curriculares Nacionais para a Educação Infantil, alguns textos voltados para o tema, diferentes modelos de pautas de observação e fichas já existentes em outras redes e sistemas de ensino” (RESPONDENTE DE TAUBATÉ, 2016).

Na referida ficha, “são apresentados os ‘conceitos’ e habilidades mínimas a serem adquiridos, de acordo com o nível de desenvolvimento da criança em relação à turma que está inserida: Berçário I (4 meses a 2 anos, completados a partir de 1º de julho); Berçário II (2 anos); Maternal I (3 anos); Maternal II (4 anos); Jardim (5 anos)”. (RESPONDENTE DE TAUBATÉ, 2016)

Os professores registram, na ficha, “o desempenho e conquistas dos alunos, a cada bimestre”. Após esse processo, “o professor coordenador [...] faz uma síntese do desempenho dos alunos e da turma, identificando as principais necessidades, e propondo ao professor as intervenções pedagógicas mais pertinentes e favoráveis para os avanços dos alunos e melhoria dos resultados de aprendizagem” (RESPONDENTE DE TAUBATÉ, 2016).

A Ficha está estruturada com espaço para: i) identificação da unidade educacional, da professora e da criança; ii) listas de afirmativas quanto ao desenvolvimento infantil, em diversos aspectos, onde a professora deve assinalar a situação do desenvolvimento da criança; iii) espaço para controle de frequência da criança; iv) espaço para que a professora anote observações realizadas por ela, ao longo do bimestre. No Quadro 5, estão registrados os aspectos que são avaliados e os níveis de desenvolvimento das crianças que devem ser apontados pelos professores, na referida Ficha.

Quadro 5 – Aspectos avaliados e níveis de desenvolvimento das crianças, constantes na Ficha de Acompanhamento Individual do Aluno. Taubaté

FAIXA ETÁRIA	ASPECTOS AVALIADOS	ASSINALAÇÕES DO PROFESSOR
Berçário I e II	Linguagem e comunicação Desenvolvimento cognitivo Desenvolvimento socioafetivo Motricidade fina Motricidade ampla	S – Sim N – Não EA – Em aquisição
Maternal I e II Jardim	Linguagem oral e escrita Matemática Movimento Música e artes visuais Natureza e sociedade Identidade e autonomia	RB – Realizou bem RA – Realizou com ajuda EA – Em aquisição

Fonte: Prefeitura Municipal de Taubaté. Secretaria de Educação (s/d – mimeo).

As afirmativas relativas ao aspecto Matemática, tanto para o Maternal II quanto para o Jardim, estão registradas no Quadro 6, de modo a ilustrar o que é considerado na avaliação das crianças. Observam-se pequenas diferenças entre o que é avaliado para as crianças de quatro anos e para as de cinco.

Quadro 6 – Afirmativas do aspecto matemática na ficha de acompanhamento individual do aluno. Maternal I e Jardim. Taubaté

MATERNAL I	JARDIM
Usa contagem oral em diferentes situações, na sequência correta	Usa contagem oral em diferentes situações, na sequência correta.
Percebe e nomeia cores diversas.	Reconhece e nomeia cores diversas.
Percebe a diferença entre número e letras.	Percebe a diferença entre número e letras.
Faz contagem usando materiais concretos até o número [a professora deve preencher]	Faz contagem usando materiais concretos.
-	Identifica numa sequência o número que vem antes, o que fica entre e o que vem após um número indicado.
Nomeia corretamente os números: 1º/2º Bimestre: de 0 a 10.	Nomeia corretamente os números: 1º/2º Bimestre: de 0 a 30.
Nomeia corretamente os números: 3º/4º Bimestre: de 10 a [a professora deve preencher]	Nomeia corretamente os números: 3º/4º Bimestre: até 50.
Faz escrita correta dos números: 1º/2º Bimestre: de 0 a 10 3º/4º Bimestre: de 10 a [a professora deve preencher]	Faz escrita correta dos números: 1º/2º Bimestre: de 0 a 30 3º/4º Bimestre: até 50.
Utiliza noções simples de cálculo mental como ferramenta para resolver problemas.	Utiliza noções de cálculo mental como ferramenta para resolver problemas.
Utiliza procedimentos para comparar grandezas e medidas em diferentes situações.	Utiliza procedimentos para comparar grandezas e medidas em diferentes situações.
Reconhece e nomeia formas geométricas.	Reconhece e nomeia formas geométricas.

Fonte: Prefeitura Municipal de Taubaté. Secretaria de Educação (s/d – mimeo).

O uso dessa Ficha possibilita “ao professor conhecer, de maneira individual, os discentes que constituem sua classe, reconhecendo suas habilidades, capacidades e dificuldades, atentando-se ao desenvolvimento da aprendizagem, considerando seus avanços desde o ingresso deste na instituição de ensino até o término do ano letivo, realizando as intervenções necessárias ao longo do processo” (RESPONDENTE DE TAUBATÉ, 2016).

Já as Atividades Avaliativas Comuns são utilizadas apenas com as crianças do Maternal II e Jardim (quatro e cinco anos, respectivamente). Tais atividades “têm como objetivo básico, servir como instrumento de estudo, pesquisa e acompanhamento do trabalho pedagógico durante o semestre, de acordo com o currículo mínimo estabelecido para cada um dos níveis, constituindo uma consciência de rede e atuando para a melhoria dos resultados de aprendizagens dos alunos” (RESPONDENTE DE TAUBATÉ, 2016).

De acordo com a respondente, o instrumento “não tem caráter seletivo, classificatório ou de promoção”. Os resultados obtidos pelas crianças, nessas atividades, são colocados em “uma planilha” que “fornece um panorama de cada

turma, cada escola e da rede de ensino, subsidiando os estudos e discussões para uma efetiva intervenção pedagógica com vistas à melhoria dos resultados de aprendizagem” (RESPONDENTE DE TAUBATÉ, 2016).

De forma sintética, em Taubaté há duas experiências de avaliação da educação infantil. O município utiliza os Indicadores da Qualidade na Educação Infantil, desde 2013, para promover a autoavaliação dos estabelecimentos educacionais, tendo contado, em 2014, com a participação de toda a comunidade escolar (profissionais da educação docentes e não docentes, famílias, gestores, entre outros). Os Planos de Ação, que sistematizam os resultados dos processos avaliativos, efetuados por meio dos *Indicadores* e que propõem ações de melhoria, são acompanhados e supervisionados pela Secretaria Municipal de Educação.

No que tange ao desenvolvimento/aprendizagem das crianças, a rede dispõe de instrumentos elaborados pela Secretaria, em conjunto com professores: i) a Ficha de Acompanhamento Individual, elaborada com base nos Referenciais Curriculares Nacionais para a Educação Infantil – destinada a todas as crianças – e as Atividades Avaliativas Comuns, utilizadas apenas com as crianças de quatro e cinco anos, no intuito de realizar estudos, pesquisas e acompanhamento do trabalho pedagógico. Os resultados da avaliação, realizada por meio de ambos os instrumentos, são utilizados para detectar os avanços e as dificuldades das crianças, para redimensionar a prática pedagógica e o planejamento docente.

Possibilidades e limites das iniciativas de avaliação para contribuir com a garantia do direito à educação das crianças de zero a cinco anos

Tendo em conta o arcabouço legal e produção acadêmica relativa ao que se compreende como direito à educação infantil bem como os aspectos de qualidade que devem ser relevados para garantir esse direito à todas as crianças de zero a cinco anos de idade, cuja síntese está exposta no item 2 deste artigo, passamos agora às análises das iniciativas de avaliação dos municípios de Birigui e Taubaté, voltadas para essa etapa da educação básica.

Como já assinalado e problematizado neste texto, perspectivas de avaliação com foco no desenvolvimento/aprendizagem das crianças não é uma perspectiva assumida pelos ditames legais existentes no país para a educação infantil. Ao contrário, a avaliação do desenvolvimento/aprendizagem das crianças tem sido reafirmada como não classificatória e sem fins de promoção para etapas seguintes, além de se constituir como de responsabilidade dos profissionais que atuam diretamente com as crianças, nos estabelecimentos

educacionais e, não dos órgãos intermediários e centrais da política educacional. O que é estabelecido para todas as redes é a “expedição de documentação que permita atestar os processos de desenvolvimento e aprendizagem da criança”, de acordo com o Item V do Artigo 31 da Lei de Diretrizes e Bases da Educação Nacional nº 9394/96. (BRASIL, 2017).

Nestes termos, a retirada da iniciativa de avaliação da aprendizagem, proposta por Birigui, parece indicar o atendimento do município aos preceitos dessa legislação.

No caso de Taubaté, as informações prestadas pela Secretaria indicam que a avaliação do desenvolvimento/aprendizagem das crianças é efetuada por meio de instrumentos comuns à toda a rede, formulados no âmbito do órgão central, com a participação de professores da creche e da pré-escola. A afirmação é de que esta avaliação não possui fins classificatórios e de promoção. Ao contrário, tem como finalidade identificar os avanços e as dificuldades das crianças para redimensionar a prática pedagógica e o planejamento docente, sendo que as ações decorrentes da avaliação devem ser articuladas pelos coordenadores pedagógicos dos estabelecimentos educacionais, junto aos professores.

Ainda que a avaliação com esse foco não seja tomada como expressão única e exclusiva da qualidade da educação infantil ofertada em Taubaté e que orientações, advindas da Secretaria, de como realizar a avaliação das crianças, sejam bem-vindas, – dado que, comumente, são os órgãos centrais e intermediários da política educacional que mais têm contato com o debate acerca das especificidades das diferentes etapas educacionais e com a legislação que orienta as práticas educativas – é preciso atentar para como essas orientações estão sendo constituídas e com quais recortes.

A avaliação tem poder de induzir práticas e ações. Ao priorizar determinadas áreas do conhecimento, atrelando seus resultados à elaboração de ações sistemáticas por parte dos profissionais da escola para melhoria destes resultados, a tendência é que outros aspectos pertinentes às especificidades da educação infantil sejam colocados de lado (SOUSA, 2014), como por exemplo, “o desenvolvimento de práticas pedagógicas que considerem a criança como sujeito e que leve em conta as diretrizes existentes para a educação infantil; o estímulo à escuta das vozes e experiências trazidas pelas crianças; a integração entre o educar e o cuidar” (PIMENTA, 2017, p. 610-611).

Não obstante a existência de uma proposição de avaliação com foco na criança, Taubaté também desenvolve perspectiva de autoavaliação institucional participativa dos estabelecimentos educacionais de educação infantil com o uso dos Indicadores da Qualidade na Educação Infantil, que enfoca o ambiente educativo, com sete diferentes dimensões, quais sejam: planejamento institucional, multiplicidade de experiências e linguagens; interações; promoção da saúde; espaços, materiais e mobiliários; formação e condições de trabalho das professoras e demais profissionais; e cooperação e troca com as famílias e participação na rede de proteção social (BRASIL, 2009).

Os resultados do processo avaliativo devem subsidiar a elaboração de um Plano de Ação. Vale destacar que a adesão aos Indicadores é prerrogativa da escola, de acordo com a metodologia proposta no material. No entanto, muitas redes têm utilizado este material para avaliar a educação infantil, como uma iniciativa sistêmica já há algum tempo (BRASIL, 2011; PIMENTA, 2017).

Em Birigui, a avaliação da educação infantil possui duas dimensões:

Dimensão processual, com foco no ambiente educativo: adequação da proposta e dos conteúdos, os agrupamentos, a organização do ambiente, a adequação dos recursos materiais e a interação. Os instrumentos utilizados para avaliar esta dimensão são fichas preenchidas pelos técnicos da Secretaria, pelos gestores escolares e pelos professores (V. Quadro 3);

Dimensão institucional, com foco na qualidade do processo educativo, mais precisamente, no trabalho desenvolvido pelos profissionais que atuam diretamente com as crianças. Para esta dimensão, são aplicados questionários aos pais ou responsáveis das crianças de cada turma, para captar suas percepções em relação a essa questão.

A avaliação de Birigui, ao menos em relação aos focos da avaliação, parece atender, em alguma medida, aspectos de qualidade extraídos da literatura examinada¹⁸ em nossa pesquisa, especialmente em relação aos recursos materiais, adequação da proposta pedagógica às especificidades da educação infantil e às práticas educativas.

Também em Taubaté, a utilização dos Indicadores da Qualidade na Educação Infantil, faz supor o exame de aspectos de qualidade importantes para a garantia do direito à educação das crianças, como a adequação dos ambientes físicos e a qualidade de seus equipamentos; as condições de acessibilidade, física e material, para crianças/profissionais portadores de deficiências, os re-

¹⁸ Ver Quadro 1 do artigo.

ursos materiais, os processos de gestão administrativa e pedagógica das instituições, as noções de qualidade educacional dos profissionais das instituições, a integração entre o educar e cuidar, a formação dos professores, as práticas pedagógicas, a proporção crianças/professor, o estímulo para a participação das famílias, entre outros (PIMENTA, 2017).

Também é considerado como um aspecto de qualidade para a educação infantil, a realização de processos de avaliação e autoavaliação participativos das instituições. Nesse sentido, os *Indicadores* correspondem a essa expectativa.

Ainda que a metodologia dos *Indicadores* se restrinja a avaliar as instituições, o processo de autoavaliação pode demandar ações a serem desenvolvidas pelos órgãos intermediários e centrais da política educacional, como, por exemplo, investir em recursos materiais suficientes para serem utilizados com e pelas crianças ou reformar os parques e bancos de areia das instituições. As ações para atender a esses resultados, muito provavelmente, deverão partir desses órgãos.

De acordo com as informações prestadas pela Secretaria de Taubaté, os resultados obtidos na autoavaliação institucional participativa são utilizados na proposição de ações de melhoria dos aspectos e dimensões considerados mais deficientes. No entanto, não foi informado quais seriam essas ações, para quais dimensões e quem deveria desenvolvê-las.

Em Birigui, os resultados da avaliação da educação infantil, em conjunto com o IDEB obtido pelo município, compõem um indicador que é utilizado para fins de progressão na carreira dos profissionais dos estabelecimentos educacionais e da Secretaria. Ainda que as informações prestadas afirmem que essa decisão tenha sido tomada em um processo que contou com a participação de profissionais das diferentes instâncias, ela se configura como um uso não indicado pela literatura examinada. Ao contrário, este é um uso que remete ao que se tem visto em avaliações direcionadas para as demais etapas da educação básica, com potencial para desencadear processos de competição entre os profissionais da educação (SOUSA, 2009), ao invés de promover ações para a melhoria da qualidade da educação infantil, com respeito às finalidades e especificidades da etapa e, por conseguinte, aos direitos das crianças (DIDONET, 2006; BRASIL, 2012).

CONSIDERAÇÕES FINAIS

As iniciativas de avaliação da educação infantil dos municípios de Birigui e Taubaté possuem perspectivas que se voltam para a infraestrutura física e de materiais, ambiente educativo, trabalho pedagógico, dentre outros aspectos considerados fundamentais nos documentos legais que fundamentam a noção de qualidade a ser assumida na educação infantil.

Contudo, no caso de Birigui, os resultados da avaliação são utilizados de forma mais aproximada aos usos efetuados por avaliações voltadas para o ensino fundamental e médio, promovendo o pagamento de incentivos aos profissionais da escola e da Secretaria.

Em Taubaté, além da iniciativa que enfoca as dimensões/aspectos acima mencionados, há também uma avaliação que se volta para o desenvolvimento/aprendizagem das crianças. Ainda que essa perspectiva não seja utilizada para determinar os rumos da política educacional relativa à educação infantil e, sim, pelos profissionais da escola, com o objetivo de elaborar intervenções concernentes ao processo de aprendizagem das crianças e às práticas educativas, há que se atentar para que esse movimento não promova ações que se distanciem das finalidades e especificidades da educação infantil.

Reitera-se que os municípios, a exemplo de iniciativas gestadas pelo governo federal e por instituições da sociedade civil, estão percorrendo diferentes caminhos para a construção da avaliação da educação infantil, adotando, de um lado, referências nacionais que dizem respeito especificamente a essa etapa educacional e, de outro, àquelas concernentes às demais etapas da educação básica.

Esse movimento implica a necessidade de aprofundamento dos debates em torno dos contornos teórico-metodológicos que vêm sendo adotados pelos municípios para avaliar a educação infantil, problematizando sobre quais desses contornos possuem maior potencialidade para contribuir com a garantia do direito à educação pública e com qualidade para as crianças de zero a cinco anos de idade.

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.

ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO DA EDUCAÇÃO. GT 07 - Educação de Crianças de 0 a 6 anos. **Manifesto GT 07 - Avaliação da Educação Infantil e a nova portaria da Política Nacional de Avaliação da Educação Básica**: indefinições e riscos. ANPED, 2020. Disponível em: <https://anped.org.br/news/manifesto-gt-07-avaliacao-da-educacao-infantil-e-nova-portaria-da-politica-nacional-de>. Acesso em 24 ago. 2020.

BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 1977.

BARROS, Emília Cristina Ferreira. **Políticas de Avaliação e o Prêmio CREI Nota 10 no Município de João Pessoa**: avaliar, controlar e responsabilizar. 2017. Tese (Doutorado em Educação) – Centro de Educação, Universidade Federal da Paraíba, João Pessoa, 2017.

BAUER, Adriana; HORTA NETO, João Luiz; SOUSA, Sandra Zákia. **Avaliação e Gestão Educacional em Municípios Brasileiros: mapeamento e caracterização de iniciativas em curso**. Relatório Final. São Paulo: FCC/INEP, 2016 (mimeo).

BRASIL. Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação - PNE e dá outras providências. **Diário Oficial** (DOU nº 120-A, Edição Extra, quinta-feira, 26 de junho de 2014, Seção 1, p. 1/7). Brasília, DF, 2014.

BRASIL. Congresso Nacional. **Constituição da República Federativa do Brasil**. Brasília: Senado Federal, 1988.

BRASIL. Ministério da Educação. Portaria nº 458, de 5 de maio de 2020. Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica. **Diário Oficial da União**, Brasília, 2020.

BRASIL. Ministério da Educação. Portaria nº 826, de 7 de julho de 2017. Dispõe sobre o Pacto Nacional pela Alfabetização na Idade Certa – PNAIC, suas ações, diretrizes gerais e a ação de formação no âmbito do Programa Novo Mais Educação – PNME. **Diário Oficial da União**, Brasília, 2017.

BRASIL. Ministério da Educação. Instituto Nacional de estudos e Pesquisas Educacionais. Portaria nº 366, de 29 de abril de 2019. Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica (SAEB) no ano de 2019. **Diário Oficial da União**, Brasília, 2019.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. **Base Nacional Comum Curricular**. Brasília: MEC/SEB, 2019.

_____. **Educação Infantil**: subsídios para construção de uma sistemática de avaliação. Relatório Síntese. Brasília: MEC/SEB/COEDI, 2012.

_____. **Monitoramento do uso dos Indicadores da Qualidade na Educação Infantil**. Relatório Final. MEC/SEB/COEDI. 2011.

_____. **Indicadores da Qualidade na Educação Infantil**. Brasília: MEC/SEB, 2009.

BRASIL. Presidência da República. **Decreto nº 9.765, de 11 de abril de 2019**. Institui a Política Nacional de Alfabetização.

BRASIL. Senado Federal. LDB: **Lei de Diretrizes e Bases da Educação Nacional**. – Brasília: Senado Federal, Coordenação de Edições Técnicas, 2017.

CANÇADO, Natalia Francine Costa. **Avaliação na Educação Infantil e Participação**: desafios para a gestão. 2017. 131 f. Dissertação (Mestrado em Educação) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2017.

CONSELHO NACIONAL DE EDUCAÇÃO; Câmara de Educação Básica. Resolução nº 2, de 09 de outubro de 2018. Define Diretrizes Operacionais complementares para a matrícula inicial de crianças na Educação Infantil e no Ensino Fundamental, respectivamente, aos 4 (quatro) e aos 6 (seis) anos de idade. Resolução CNE/CEB 2/2018. **Diário Oficial da União**, Brasília, 10 de outubro de 2018, Seção 1, p. 10.

DIDONET, V. Coerência entre educação e finalidades da educação infantil. **Pátio Educação Infantil**, v. 6, n. 10, 2006.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 366, de 29 de abril de 2019. Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica (SAEB) no ano de 2019. **Diário Oficial da União**, publicado em 02/05/2019, ed. 83, Seção 1, p. 47. Órgão Responsável: Inep.

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**, 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?** 2017.670p. Tese (Doutorado em Educação), Faculdade de Educação da Universidade de São Paulo, São Paulo, 2017.

PREFEITURA MUNICIPAL DE BIRIGUI. **Lei nº 6.064, de 11 de agosto de 2015.** Dispõe sobre o Plano Municipal de Educação e dá outras providências. Birigui: 2015.

_____. Câmara Municipal de Birigui. **Lei Complementar nº 32, de 17 de setembro de 2010.** Dispõe sobre o Estatuto e Plano de Carreira do Magistério Público e dos Profissionais de Apoio Educacional do Município de Birigui e dá Outras Providências. Birigui: 2010.

SECRETARIA MUNICIPAL DE EDUCAÇÃO DE BIRIGUI. **Instrução Normativa SME nº. 002/de 02 de agosto de 2016**, 2016. Mimeografado.

PREFEITURA MUNICIPAL DE TAUBATÉ. **Documento Preparatório do Plano Municipal de Educação 2015-2025** – Anexo Único. Taubaté: 2015.

_____. Câmara Municipal de Taubaté. **Lei Complementar nº 392, de 15 de julho de 2016.** Institui o Plano Municipal de Educação e dá outras providências. Taubaté: 2016.

_____. Secretaria Municipal de Educação. **Educação Infantil** - Ficha de Acompanhamento. Taubaté: s/d. Mimeografado.

_____. **Atividades de Linguagem** – Maternal II. Taubaté: s/d. Mimeografado.

RIBEIRO, Bruna. **Avaliação da aprendizagem das crianças:** insumos para o debate. Relatório Técnico. Projeto UNESCO BRZ/1041. Diretoria de Currículos e Educação Integral. Coordenação Geral de Educação Infantil, SEB/MEC. São Paulo, 2016.

SOUSA, Sandra Zákia. Avaliação e Gestão da educação Básica. In: DOURADO, Luiz. **Políticas e gestão da educação no Brasil**: novos marcos regulatórios. São Paulo: Xamã, 2009, p. 31-45.

_____. Avaliação da Educação Infantil: propostas em debate no Brasil. **Interações**, Lisboa, v. 10, n. 32, p. 68-88, mês? 2014.

_____. Avaliação da Educação Infantil: questões controversas e suas implicações educacionais e sociais. **Educação PUC-Campinas**, Campinas, v. 23, n. 1, p. 65-78, jan./abr. 2018.

Recebido em: 17 de dezembro de 2019

Aceito em: 15 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÃO DE CONTEXTO NA EDUCAÇÃO INFANTIL: UMA ESCOLHA PARA PENSAR A MELHORIA DAS PRÁTICAS EDUCATIVAS

Patrícia Corsino¹

Jordanna Castelo Branco²

¹ Professora Associada da Faculdade de Educação e do Programa de Pós-Graduação em Educação da Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil. E-mail: corsinopat@gmail.com

² Técnica em assuntos educacionais da Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil. Email: jordanna.branco@gmail.com

Resumo

Este texto tem como objetivo analisar parte dos resultados de um processo de avaliação de contexto, que focalizou práticas educativas de oralidade, leitura e escrita, em uma escola de educação infantil da rede pública municipal de ensino da cidade do Rio de Janeiro. Para desenvolver a avaliação de contexto na instituição, foi elaborado um instrumento, e desenvolvida a metodologia própria desta avaliação com base na perspectiva de Bondioli e Savio (2013, 2015). Após percorrer todas as fases da avaliação de contexto, os resultados evidenciaram, pelo envolvimento da equipe e compromisso coletivo de busca da melhoria educativa da instituição, a potência formativa da proposta avaliativa.

Palavras-chave: Avaliação de contexto. Educação Infantil. Melhoria da qualidade.

Abstract

This text aims to analyze part of the results of a context evaluation process, which focuses on oral, reading and writing educational practices, in a kindergarten of the municipal public school of Rio de Janeiro. To develop a context assessment in the institution, an instrument was developed and a methodology of its own based on Bondioli and Savio (2013, 2015). After going through all the phases of context evaluation, the results evidenced by the team involvement and collective commitment to the educational improvement of the institution, a formative power of the evaluation proposal.

Keywords: Context evaluation. Early Childhood Education. Quality improvement.

INTRODUÇÃO

Ou isto ou aquilo

Cecilia Meireles

Ou se tem chuva e não se tem sol,
ou se tem sol e não se tem chuva!

Ou se calça a luva e não se põe o anel,
ou se põe o anel e não se calça a luva!

Quem sobe nos ares não fica no chão,
quem fica no chão não sobe nos ares.

É uma grande pena que não se possa
estar ao mesmo tempo nos dois lugares!

Ou guardo o dinheiro e não compro o doce,
ou compro o doce e gasto o dinheiro.

Ou isto ou aquilo: ou isto ou aquilo...
e vivo escolhendo o dia inteiro!

Não sei se brinco, não sei se estudo,
se saio correndo ou fico tranquilo.

Mas não consegui entender ainda
qual é melhor: se é isto ou aquilo.

Este poema de Cecilia Meireles, de destinação geralmente infantil, trata de questões existenciais que dizem respeito às escolhas que fazemos. Como nem sempre podemos ficar com isso e aquilo, há que se optar e, para tal, é necessário avaliar qual das opções é a mais adequada ou favorável aos objetivos pretendidos. A avaliação é parte corriqueira do dia a dia, mas ao formalizá-la e sistematizá-la, se evidenciam as intenções e opções teórico-metodológicas primeiro de quem a elaborou e, depois, de quem tratou seus resultados, pois estes atores nem sempre coincidem. Trazemos o poema para abrir o diálogo sobre avaliação de contexto na educação infantil como uma alegoria. “Ou isto ou aquilo” apresenta um eu que não sabe “qual é melhor: se isto ou aquilo”, um eu que tem dúvidas mais do que certezas, evidenciando como não é simples fazer escolhas. Ao avaliar, estamos sempre envolvidos entre isso e

aquilo. Sejam avaliações corriqueiras, individuais, de pequenos grupos, sejam institucionais, coletivas, externas, de larga escala. Todas elas exigem escolhas de questões, dimensões, metodologias.

Nossa escolha pela avaliação de contexto na educação infantil se deu pelo seu potencial de reflexão e busca coletiva de melhoria da qualidade educativa. O contexto é aqui compreendido em suas múltiplas dimensões e implicações. Um contexto institucional relacional tanto em relação aos contextos político, social, cultural, quanto das condições materiais do espaço e dos recursos da escola, das pessoas que ali convivem, das inúmeras situações que emergem das interações entre adultos, entre adultos e crianças e das crianças entre si.

Para Bondioli (2014), a avaliação das instituições educacionais tem uma natureza política. Os critérios e os parâmetros com os quais se julga a qualidade do objeto educativo – programas, instituições, ações, sistemas – são frutos de processos de negociação realizados por um grupo. São confiáveis, à medida que se revelam consensuais, refletem aspirações e intenções compartilhadas. Isso não desconsidera o rigor metodológico, nem a possibilidade ou a necessidade de avaliação, mas evidencia que o tipo de avaliação a ser feita depende do processo de negociação acerca da definição de qualidade e do seu compartilhamento. Para a negociação da qualidade, é necessário que haja um movimento dialógico entre os participantes do processo educativo, o que envolve a negociação dos valores dos sujeitos participantes, a fim de estabelecer o que é imprescindível e/ou ideal a ser compartilhado pelo grupo. Trata-se de um processo dialógico em que coletivamente se estabelecem critérios e parâmetros de qualidade com objetivo de avaliar o contexto onde atuam.

A avaliação é tida como uma negociação que ocorre com a participação dos sujeitos que fazem parte da comunidade escolar em articulação com o avaliador externo também entendido como facilitador ou formador. Um processo que alia autoavaliação e avaliação externa que acontece por meio de reflexões que são frutos de confrontos de diferentes pontos de vista e que visam a uma construção coletiva de metas e estratégias para melhoria da escola. Em poucas palavras, no cerne da avaliação de contexto, está a participação e a promoção da participação dos envolvidos no contexto educativo.

Bondioli e Savio (2013, 2015) tomam como ponto de partida as ideias de Guba e Lincon (1989) e consideram a avaliação de contexto como de quarta geração. Para os autores citados, que estudaram as diferentes formas de avaliação ao longo do tempo, a quarta geração de avaliação tem como elemento chave o diálogo, a negociação de pontos de vista e, nesse processo, o avaliador atua como mediador. Trata-se de uma avaliação que busca romper com a he-

gemonia do paradigma positivista, das medições quantitativas, da desconsideração do contexto, da eliminação de caminhos alternativos para se pensar o objeto da avaliação, da não-responsabilização moral e ética do avaliador pelo que emerge da avaliação ou pelo uso dos seus resultados, entre outros. É uma avaliação responsiva, na qual as reivindicações, preocupações e questões dos grupos de interesse são a base para determinar as informações necessárias. A avaliação de contexto, como um processo democrático, dialógico e formativo, que visa à melhoria das práticas educativas, se vale de uma metodologia participativa que articula dois eixos: i) “a negociação da qualidade” – a qualidade como um conceito polissêmico e contextual, implica em concepções e valores, e exige uma negociação entre os integrantes do contexto a ser avaliado; ii) “a promoção a partir do interno” – ao refletir sobre o contexto educativo nas suas articulações e complexidades, a proposta alia autoavaliação com avaliação externa, num processo dialógico de confronto de pontos de vista, com o objetivo de sistematizar hipótese e projetos de melhoria.

Este texto tem como objetivo analisar parte dos resultados de uma avaliação de contexto, que focalizou práticas educativas de oralidade, leitura e escrita, em uma escola de educação infantil da rede pública municipal de ensino da cidade do Rio de Janeiro. Para desenvolver a avaliação de contexto na instituição, foi elaborado um instrumento, e desenvolvidas as fases da metodologia próprias desta avaliação com base na perspectiva de Bondioli e Savio (2013, 2015).

O presente texto está organizado em três partes. Na primeira, apresentamos a metodologia de avaliação de contexto e os procedimentos metodológicos por ela requeridos ao longo do seu desenvolvimento. Na segunda parte, trazemos como esses procedimentos foram operacionalizados no campo da pesquisa, um Espaço de Desenvolvimento Infantil (EDI) da rede pública municipal do Rio de Janeiro. Também, discutimos os resultados parciais da pesquisa. Na terceira e última parte, após percorrer todas as fases da avaliação de contexto, os resultados evidenciaram, pelo envolvimento da equipe e compromisso coletivo de busca da melhoria educativa da instituição, a potência formativa, reflexiva e mobilizadora da proposta avaliativa.

A METODOLOGIA DE AVALIAÇÃO DE CONTEXTO

De acordo com Bondioli e Savio (2015), a metodologia de avaliação de contexto exige que sejam realizados alguns passos sucessivos, caracterizados pela reflexão e pela negociação entre os sujeitos envolvidos no percurso avaliativo, com o acompanhamento do pesquisador/formador. Os passos podem ser

realizados em um ou mais momentos, o que varia de acordo com as exigências e as particularidades educativas em questão.

O primeiro passo consiste em que um dos sujeitos que faça parte da instituição solicite ou se interesse pela realização de um percurso avaliativo de contexto. Na avaliação como “promoção a partir do interno”, a comunidade escolar negocia internamente quais de seus membros participarão de modo ativo do grupo de trabalho, que pode ser composto por até 20 participantes. Por isso, “É importante que tenha uma referência comum, sejam parte da mesma comunidade escolar ou região ou cidade ou estado, o que garante que o percurso de reflexão e o debate se refira a mesma identidade educativa” (BONDIOLLI; SAVIO, 2015, p. 29).

No segundo passo, o grupo de trabalho especifica e negocia os objetivos pelos quais a avaliação é realizada e o aspecto a ser avaliado. O foco é avaliar a própria instituição educativa, a fim de refletir sobre “o que se faz, como faz e por que faz” e depois definir e realizar intervenções que visem a melhoria.

O terceiro passo envolve a escolha do instrumento avaliativo a ser usado e/ou sua análise crítica. É um momento em que cada participante compartilha no grupo de trabalho seus acordo e desacordo em relação ao instrumento e explicita seus motivos. No caso da pesquisa em análise, o instrumento focalizou as linguagens oral e escrita em três dimensões: i) materiais e organização do ambiente educativo; ii) ação/papel do professor; iii) planejamento/ intencionalidade pedagógica. Embora o tema não tenha sido uma proposta do grupo, ele foi bem acolhido, pois este é um campo bastante disputado na pré-escola pelas concepções controversas que coexistem nas políticas e nas práticas das escolas públicas municipais da cidade do Rio de Janeiro.

No quarto passo é realizada a avaliação do contexto educativo utilizando o instrumento já discutido e criticado pelo grupo. É a única etapa de todo o processo que é individual, pois, nela, cada participante, pesquisador/formador e professores da escola realiza a observação em um período comum; comparam o que foi observado com as questões do instrumento e pontuam. Este é o momento da avaliação externa e da autoavaliação ainda em separado.

O quinto passo ocorre quando a auto avaliação e a avaliação externa se confrontam: as pontuações são discutidas e as discrepâncias, justificadas. A partir da análise das avaliações realizadas pelo formador e pelos professores envolvidos, emergem os “pontos fracos” e/ou “fortes” e as eventuais diferenças de avaliação a serem compartilhadas e confrontadas. Neste momento, emerge a

identidade educativa da instituição, e os pontos levantados são organizados para serem discutidos e negociados na restituição ao coletivo da instituição.

O penúltimo passo implica, “com base no resultado da discussão dos dados avaliativos, que o grupo negocie e tome decisões acerca das ações futuras” (BONDIOLLI; SAVIO, 2015, p. 31), definir metas e estratégias com vistas à melhoria.

O último passo consiste em avaliar o processo avaliativo. O pesquisador/formador retoma os passos anteriores, procurando evidenciar o seu objetivo e as passagens que o caracterizaram. Para isso, pede que cada um dos participantes exprima livremente a sua avaliação. Em seguida, o grupo debate e reflete sobre o sentido do processo avaliativo e delinea coletivamente uma avaliação do instrumento avaliativo utilizado ao longo do percurso.

Cada um dos passos compõe diferentes fases da metodologia (BONDIOLLI; SAVIO, 2015, 2013; MEC/SEB, 2015), como pode ser visto no Quadro 1, a seguir.

Quadro 1 – Fases previstas na metodologia de avaliação de contexto

Fase I	Plenária Inicial: apresentação do processo de avaliação de contexto com a participação dos membros da instituição.
Fase II	Discussão e negociação do objetivo da avaliação. Escolha do instrumento avaliativo a ser usado.
Fase III	Análise crítica do instrumento de avaliação escolhido, através da discussão coletiva do instrumento e questionário individual de meta-avaliação.
Fase IV	Realização de observação em turma, entrevista com professores e equipe de direção (articulador pedagógico, diretor e vice-diretor) e consulta a documentação da escola.
Fase V	Encontros de restituição da observação e da pontuação atribuída. Nessa fase, são restituídos ao grupo os resultados do processo avaliativo, provocando reflexões em relação aos itens das áreas avaliadas, debatendo sobre as avaliações discordantes, aprofundando as razões dos diferentes pontos de vista.
Fase VI	Elaboração coletiva de um plano de ação para a melhoria da escola, com base nas discussões em torno da avaliação.
Fase VII	Plenária Final: avaliação do processo avaliativo pelos participantes.

Fonte: elaboração a partir de Souza, Moro e Coutinho (org) (2015, p. 29-31).

A METODOLOGIA EM AÇÃO

A metodologia de avaliação de contexto foi desenvolvida ao longo de dois anos. No primeiro ano, foi feita uma revisão bibliográfica sobre avaliação de contexto e sobre leitura e escrita na educação infantil e um período de seis meses de doutorado sanduíche na Universidade de Pavia. No segundo ano, todo processo de entrada no campo e a pesquisa de campo propriamente dita. Nesse período, foi realizado o planejamento da pesquisa e o desenvolvimento das fases da metodologia de contexto mencionadas anteriormente, de acordo com as possibilidades da escola campo da pesquisa.

Nos meses de abril e julho, realizamos o planejamento da pesquisa. No seu primeiro momento, realizou-se o mapeamento de possíveis escolas interessadas em participar da pesquisa entre as escolas da rede pública municipal do Rio de Janeiro participantes do Complexo de Formação de Professores – CFP da UFRJ³. Chegamos a um Espaço de Desenvolvimento Infantil (EDI) que atende crianças de 3 a 5 anos, que foi muito receptivo à pesquisa. Também foi submetido o projeto da pesquisa ao comitê de ética da SME-Rio para avaliação. Em seguida, foi realizada revisão do Instrumento de Avaliação de Contexto das Práticas Educativas de Oralidade, Leitura e Escrita, elaborado em parceria com as pesquisadoras italianas Ana Bondioli e Donatella Savio da Universidade de Pavia⁴.

Foram consultados documentos oficiais municipais e nacionais relacionados à leitura e escrita na educação infantil. Tais consultas tinham por objetivo observar o alinhamento dos documentos locais e nacionais com o instrumento da pesquisa.

O documento da SME-Rio que orienta as práticas educativas dos EDI são as “Orientações Curriculares para a Educação Infantil”. Estas orientações mencionam como principais linguagens da pré-escola o brincar e o falar, ler e escrever de modo a compreender a brincadeira e a livre expressão das crianças como articuladas, fundamentais e presentes nas suas interações entre elas e delas com os adultos (RIO DE JANEIRO, 2010). O documento aponta que

³ Este trata-se de um novo modelo institucional para a formação de professores, concebida como formação profissional e universitária. O CFP funciona como um terceiro espaço, no qual integram as escolas de educação básica e as diferentes licenciaturas da UFRJ, professores e licenciandos. Esse espaço articula a universidade e a escola básica em prol da formação de professores de modo a reconhecer a diferenciação e a convergência de papéis, singulares e distintos, sem que os saberes de um se sobreponham aos do outro. Isto é, entende-se que as contribuições para a formação sejam pensadas em comum, compartilhadas horizontalmente com todos os envolvidos.

⁴ Parceria estabelecida anteriormente pela participação de uma das autoras num projeto de pesquisa Interinstitucional de Avaliação de Contexto que contou com o convênio com a Universidade de Pavia-Itália, e que teve continuidade no doutorado sanduíche realizado pela outra autora.

a fala, a escrita e a leitura estão articuladas e que isto se dá por meio das interações sociais entre adultos e crianças, e delas entre elas, à medida que nas suas inter-relações são construídos e partilhados significados por meio da experiência de situações reais e imaginadas presentes no cotidiano da pré-escola.

No que tange à ação do professor em relação às crianças, para o documento “a meta é garantir a presença de experiências que sejam importantes para a alegria, o desenvolvimento e o crescimento das crianças, fortalecendo, assim a possibilidade de sucesso escolar na sua trajetória educacional” (RIO DE JANEIRO, 2010, p. 31). Para tal, são estabelecidos objetivos e habilidades a serem alcançadas pelas crianças para cada área do conhecimento, dentre elas, as linguagens oral e escrita. Não faremos aqui uma análise mais detalhada deste documento, mas observamos que tais pressupostos dialogam, mesmo que parcialmente, com o que determina as Diretrizes Curriculares Nacionais para a Educação Infantil – DCNEI (CNE/SEB, Resolução n. 5/2009) documento de cunho mandatário que serviu de base para a elaboração do instrumento de Avaliação de Contexto das Práticas Educativas de Oralidade, Leitura e Escrita. Desse modo, os documentos oficiais dialogavam parcialmente entre si e, conseqüentemente, com o instrumento, o que possibilitou o seu uso na avaliação de contexto.

A entrada no campo compreendeu a negociação com os professores e com a equipe de gestão da escola, quando conseguimos realizar duas reuniões. A escola, localizada na Zona Sul da cidade do Rio de Janeiro, atende, em horário integral, cerca de 120 crianças de 3 a 5 anos de idade, oriundas tanto de classe média, moradoras do entorno da escola, quanto das classes populares, moradoras de comunidades, especialmente, da Rocinha. Em 2018, a equipe da escola era composta por dez professoras de turma, uma professora de arte, um professor de educação física, duas auxiliares de turma, uma merendeira, uma professora articuladora (coordenadora pedagógica), uma diretora e uma vice-diretora.

As fases da pesquisa foram organizadas junto com os professores e a direção da escola. No primeiro momento, apresentamos a pesquisa à equipe da escola numa plenária inicial, e aderiram a proposta, voluntariamente, sete das dez professoras de turma, que constituíram o grupo de trabalho. A fim de que os encontros de discussão do instrumento fossem possíveis, os professores de Educação Física e Artes se propuseram a realizar atividades com as turmas das professoras participantes da pesquisa. Além disso, foi também disponibilizado para a discussão individual da restituição das observações e avalia-

ções o horário de encontro das professoras com a coordenação pedagógica, evidenciando o compromisso da escola com o processo avaliativo. A direção reservou, para a realização da plenária final, o dia do Centro de Estudo Integral, momento em que não teria o atendimento às crianças, sendo o único dia viável no calendário para reunir toda a equipe.

É interessante notar que esse processo de organização da escola para a realização da pesquisa deu início ao movimento de mobilização dos professores em direção à participação. Esta é compreendida não apenas como ação dos diferentes sujeitos que atuam na escola, mas também como um processo dialógico entre eles atravessado pela expressão e pela escuta, de modo a aceitar os diferentes pontos de vista, sem necessariamente concordar, com o objetivo de construção coletiva (SAVIO, 2018). Assim, pode-se dizer que os professores das diferentes áreas de atuação na escola, ao procurarem juntos uma forma de se organizarem para possibilitar a realização da pesquisa, realizam o movimento de colaboração e atuação para a construção coletiva. Isso fica evidente ao serem realizadas propostas de colaboração nesse sentido.

Diante da necessidade de adequar as fases previstas na metodologia com as possibilidades do campo de pesquisa, realizamos: a) a plenária final, que consistiu na apresentação dos pontos levantados na restituição organizados em “pontos fortes” e “pontos a serem melhorados” e na reflexão a respeito deles; e b) a elaboração coletiva de um plano de ação com o objetivo de melhoria dos pontos levantados, processo no qual o formador/pesquisador atuou como mediador.

No Quadro 2, a seguir, apresentamos o processo de avaliação de contexto realizado na escola.

Quadro 2 – Desenvolvimento das fases da metodologia de avaliação de contexto.

Fases da metodologia	Tempo aproximado	Ação
Fase I - Apresentação do instrumento	1 encontro com todos os professores e equipe de gestão Plenária de apresentação	Apresentação do processo de pesquisa. Esta etapa se constituiu num momento de participação coletiva da instituição.
Fase II- Apresentação do instrumento	1 encontro com todos os professores e equipe de gestão Plenária de apresentação	Apresentação do instrumento, distribuição do questionário de meta avaliação do instrumento, assinatura do Termo de Consentimento Livre e Esclarecido da pesquisa e composição do grupo de trabalho e organização do cronograma de desenvolvimento das fases seguinte da pesquisa.
Fase III – Discussão e meta-avaliação do instrumento	2 encontros- com o grupo de trabalho	Encontros de discussão sobre o instrumento, cada um com cerca de 2 horas, com as sete professoras da escola participantes da pesquisa. Entrega e recolhimento do questionário de meta avaliação.
Fase IV - Observação, avaliação e confronto	4 semanas	Observação, na turma de cada uma das 7 professoras participantes da pesquisa, por 3 dias em cada turma. Ao fim desse período, ocorreu a avaliação da pesquisadora e a auto avaliação de cada professora observada. No quarto dia, compararam as pontuações atribuídas por cada uma delas a partir do que foi observado. Também foram realizadas entrevistas e consultas à documentação da escola.
Fase V - Restituição e avaliação do processo	2 encontros com todos os professores e equipe de gestão Plenárias de restituição	Restituição ao grupo dos resultados do processo avaliativo, estimulando reflexões em relação aos itens da(s) área(s) avaliada(s), debatendo sobre as avaliações discordantes, aprofundando as razões dos diferentes pontos de vista. Elaboração coletiva de um plano de ação para a melhoria da escola, processo no qual o formador/pesquisador atua como mediador. Avaliação coletiva do processo avaliativo realizado.

Fonte: Elaborada pelos autores a partir dos dados da pesquisa (BRANCO, 2019).

O DESAFIO DA PROMOÇÃO A PARTIR DO INTERNO

Após a discussão do instrumento, foram realizadas observações em cada turma e conferidas as pontuações sobre os itens do instrumento por cada uma das sete professoras da pesquisa, como autoavaliação, e pela pesquisadora, como olhar externo. Foram realizados encontros individuais da pesquisadora com cada professora para confrontação das avaliações e reflexão sobre os diferentes pontos de vista. A partir das pontuações e das eventuais diferenças de avaliação, foram destacados pontos fortes das práticas educativas desenvolvidas na escola no campo da oralidade, leitura e escrita, e pontos fracos, ou seja, pontos que precisariam de maior atenção da equipe. O objetivo foi que os pontos fortes das práticas, ainda que não fossem comuns a todas as

professoras, fossem tomados como pontos de partida para a melhoria das ações educativas no campo da oralidade, leitura e escrita na educação infantil e que os pontos fracos/ a serem melhorados destacados nas práticas fossem entendidos como fragilidades a serem superadas coletivamente.

Cabe ressaltar que, como o instrumento apresenta princípios baseados nas Diretrizes Curriculares Nacionais para a Educação Infantil – DCNEI (CNE/CEB, Resolução nº 5 / 2009), considera as crianças como centro do processo educativo e as interações e as brincadeiras como eixos das práticas educativas, a linguagem ganha um lugar central nas práticas pedagógicas. No instrumento, foram elencados alguns princípios que serviram de base para as perguntas/ pontos a serem avaliados, são eles:

- A linguagem é valorizada por suas funções constituinte, comunicativa, representativa, imaginativa e criativa. As práticas de oralidade, leitura e escrita se desenvolvem em situações significativas, reais e imaginárias, ou seja, em situações interativas, nas quais as crianças participam ativamente e que são consolidadas como práticas sociais, enquanto necessárias no processo dialógico entre os interlocutores;
- A leitura, para que seja significativa para as crianças, depende da performance do leitor adulto – que empenha toda a sua corporeidade, dando voz ao texto – como também da qualidade do texto, da ilustração e da maneira pela qual o tema é tratado – da sua capacidade de surpreender e comover o leitor, de fazê-lo refletir sobre si mesmo, sobre o outro e sobre o mundo. Na educação infantil, não cabe ler para mera memorização de letras, sílabas, palavras soltas. A leitura é lugar de ampliação do mundo simbólico, de conhecimentos, da fabulação, do uso da língua, do encontro com a alteridade constitutiva;
- A escrita, para ser significativa, está relacionada às atividades de narração e descrição do mundo e surge como necessidade de registro e de apoio à memória. Na turma, exerce uma importante função de construção da história do grupo, de registro das narrativas de experiências, descobertas, conquistas, histórias, passeios e situações vividas e de tudo que for relevante para a turma. Os registros podem ser feitos de diferentes formas, mas, especialmente, por escrito, podem ser impressos em diferentes suportes e apresentar variados gêneros discursivos. Não cabe, na educação infantil, qualquer forma de exercício repetitivo e cópias.

Foram apontados pontos fortes e fracos/a serem melhorados com base nas observações das turmas e nos confrontos de pontos de vista. Na primeira reunião de restituição, a equipe tomou conhecimento do processo e refletiu sobre os pontos elencados de forma bastante envolvida. Perguntas foram levantadas, a partir da apresentação em PowerPoint-PPT que trazia os pontos fortes e a serem melhoradas com fotografias e falas das professoras. Ao final, foram tirados alguns pontos e, no segundo encontro, o grupo chegou aos principais desafios a serem enfrentados, são eles:

- A equipe da escola se constituir enquanto grupo, aprendendo a lidar com as diferenças entre seus membros, mas buscando princípios comuns;
- Necessidade de ter mais momentos de troca de experiências e reflexão do grupo;
- Reflexão coletiva sobre como as crianças têm sido vistas pelos adultos, em especial as crianças tidas como fora do padrão. Considerar, nesse processo, as questões: as crianças de cada grupo são responsáveis de toda a equipe? As crianças são consideradas parte da escola, independente das suas dificuldades?
- Pensar estratégias de escuta das crianças com o objetivo de planejar e propor projetos a partir das suas contribuições e/ou interesses. A partir disso, flexibilizar o planejamento de modo que dialogue com os interesses e contribuições das crianças e, com base nele, organizar a rotina;
- Organizar os espaços e tempos da escola sem deixar de reconhecer a autonomia das crianças, ou seja, propiciar que elas tenham diferentes materiais a sua disposição e que haja momentos na rotina que sejam usados livremente, sem direcionamento direto do adulto;
- Refletir sobre quais são as práticas de oralidade, leitura e escrita próprias da Educação Infantil e o quais o grupo considera como práticas de qualidade nesse sentido.

A partir das reflexões suscitadas ao longo do processo de avaliação de contexto, esta síntese dos desafios a serem enfrentados pelo grupo passou a fazer parte das discussões realizadas pela equipe da escola nos Centros de Estudos de 2019 mensais, que foram acompanhados por pesquisadoras do Grupo de Pesquisa Infância, Linguagem e Educação (GILE) da UFRJ. Este processo formativo do coletivo da escola, com vistas a melhoria da qualidade educativa, está tendo continuidade em 2020, quando a pandemia Covid-19 tem suscitado novos desafios.

CONCLUSÃO

Os resultados da pesquisa evidenciaram alguns pontos que consideramos básicos para se pensar a avaliação de contexto. Em primeiro lugar, destacamos a potência da metodologia para o processo de reflexão tanto individual quanto coletivo. O conhecimento prévio e a discussão de cada questão do instrumento que serviu de parâmetro para o processo avaliativo foi um importante momento, não apenas de familiaridade com os itens da avaliação, mas também de formação dos professores, já que provocaram discussões, trocas, esclarecimentos e desejo de conhecer mais sobre os temas propostos. Foram duas reuniões bastante produtivas. Na primeira, embora as professoras

estivessem ainda se soltando, foi possível criar um espaço de confiança com as pesquisadoras que se estreitou na reunião seguinte. Ficou clara a necessidade de o grupo pensar coletivamente as práticas. Outro ponto importante da metodologia foi o confronto das pontuações. Não foi simples refletir sobre as discrepâncias entre as notas atribuídas pela pesquisadora e as atribuídas pelas professoras, pois, na escola, como em muitas, não há uma cultura de avaliação na perspectiva formativa. Em segundo lugar, destacamos o próprio instrumento, porque, embora ele tenha sido focalizado na linguagem verbal (oralidade, leitura e escrita), permitiu pensar muitas questões que perpassam a linguagem, as interações, as concepções de infância e de educação infantil, como se observa nos desafios que foram elencados pela equipe da escola na plenária de restituição.

Percebemos, neste trabalho de campo, a importância de se abrir espaços interlocutivos nas escolas nos quais as professoras tenham de fato seu lugar de fala garantido, onde a participação como princípio garanta a quebra de hierarquias, tão fortemente instituídas, no âmbito da Secretaria Municipal de Educação do Rio de Janeiro. Foi a escuta dos pontos fortes e fracos, dos comentários de cada professora, na presença da equipe de gestão, que possibilitou à equipe pensar coletivamente os desafios e a buscar formas de superá-los. A negociação da qualidade foi se dando neste processo de pensar junto e propor junto. Por fim, destacamos a importância de as redes municipais de educação pensarem em formas de instituir a avaliação de contexto nessa perspectiva formativa, dialógica e, conseqüentemente, uma busca efetiva, por ser pactuada coletivamente, de melhoria da oferta da educação infantil. Para isso, há que se fazer escolhas já que nem sempre dá para conciliar isto e aquilo.

REFERÊNCIAS

BONDIOLI, Anna; SAVIO, Donatella. Elaborar indicadores de qualidade educativa das instituições de Educação Infantil: uma pesquisa compartilhada entre Itália e Brasil. In: SOUSA, Gisele; MORO, Catarina; COUTINHO, Angela Scalabrín. **Formação da rede em Educação Infantil: avaliação de contexto**. Curitiba: UFPR, 2015.

BONDIOLI, Anna; SAVIO, Donatella (orgs). **Participação e qualidade em Educação da Infância: percursos de compartilhamento reflexivo em contextos educativos**. Curitiba: UFPR, 2013.

BRANCO, Jordanna Castelo. **Avaliação de contextos de práticas de oralidade, leitura e escrita na pré-escola: possibilidades e desafios**. Tese de doutorado. Faculdade de Educação, Universidade Federal do Rio de Janeiro, 2019.

GABRIEL, Carmen Tereza. Complexo de Formação de Professores: uma experiência (inter)institucional em curso. *Profesorado - Revista de currículum y formación del profesorado*, v. 23, n. 3 (jul.-set. 2019), p. 189-209. Disponível em: <https://185.79.129.77/index.php/profesorado/article/view/74500/45295>

Acesso em: 23 de ago. 2020.

GUBA, E.; LINCOLN, Y. **Fourth generation evaluation**. Califórnia-USA: Sage, 1989.

MEC/SEB/COEDI. **Contribuições para a política nacional: a avaliação em Educação Infantil a partir da avaliação de contexto**. Curitiba: Imprensa/UFPR; Brasília: Ministério da Educação. Secretaria da Educação Básica. Coordenação de Educação Infantil, 2015.

NÓVOA, Antonio. **Um Novo Modelo Institucional para a Formação de Professores na Universidade Federal do Rio De Janeiro**. Universidade Federal do Rio de Janeiro, 2017.

RIO DE JANEIRO. Secretaria Municipal de Educação/Gerência Especial de Educação Infantil. **Orientações Curriculares para a Educação Infantil**. Rio de Janeiro, 2010. Disponível em: <http://www.rio.rj.gov.br/dlstatic/10112/4246634/4104953/orientacoesCurricularesEdInfantil.pdf> . Acesso em: 23 ago. 2020.

SAVIO, Donatella. Promover a partir de dentro: uma abordagem reflexiva e participativa da avaliação de contextos educativos1. **Pro-Posições**, Campinas, v. 29, n. 2, p. 72-92, ago. 2018. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-73072018000200072&lng=pt&nrm=iso. Acesso em: 24 nov. 2019.

Recebido em: 24 de novembro de 2019

Aceito em: 20 de junho de 2020

Publicado em: 30 de junho de 2020

INFRAESTRUTURA ESCOLAR E SATISFAÇÃO PROFISSIONAL: PERCEPÇÃO DE PROFESSORES DA EDUCAÇÃO BÁSICA BRASILEIRA

Lívia Maria Fraga Vieira¹

Edmilson Antônio Pereira Junior²

¹ Professora Associada da Graduação e Pós-graduação em Educação da Faculdade de Educação da Universidade Federal de Minas Gerais, Belo Horizonte, MG. E-mail: liviafraga59@gmail.com

² Estatístico, pós-doutorando do Programa de Pós-graduação em Educação da Faculdade de Educação da Universidade Federal de Minas Gerais, Belo Horizonte, MG. E-mail: edmilsonpj@yahoo.com.br

Resumo

Analisa-se a percepção de professores da educação básica sobre as condições da infraestrutura escolar e a satisfação profissional, ressaltando especificidades da educação infantil (EI), com indicadores desenvolvidos a partir de resultados da pesquisa Trabalho Docente na Educação Básica no Brasil – Fase II, coordenada pelo Grupo de Estudos sobre Política Educacional e Trabalho Docente, da Universidade Federal de Minas Gerais (UFMG), de 2015. Mesmo com remuneração e carreiras mais desvantajosas, as docentes da EI apresentaram índice de respostas mais elevado em relação à satisfação profissional.

Palavras-chave: Educação infantil. Satisfação profissional. Infraestrutura escolar. Condições da unidade educacional. Condições da sala de aula.

Abstract

The perception of basic education teachers about the conditions of school infrastructure and job satisfaction is analyzed, highlighting specificities of early childhood education (EI), with indicators developed from the results of the research Work in Basic Education in Brazil - Phase II, coordinated by the Study Group on Educational Policy and Teaching Work, from the Federal University of Minas Gerais (UFMG), 2015. Even with inequalities in pay and careers in relation to other stages of basic education, EI teachers showed a higher response rate related to job satisfaction.

Keywords: Early childhood education. Professional satisfaction. School infrastructure. Conditions of the educational unit. Classroom conditions.

INTRODUÇÃO

O Plano Nacional de Educação, aprovado pela Lei n. 13.005/2014, estabeleceu a criação de um Sistema Nacional de Avaliação da Educação Básica como “fonte de informação para a avaliação da qualidade da educação básica e para a orientação das políticas públicas desse nível de ensino” (BRASIL, 2014, p. 2). A cada dois anos, o sistema de avaliação deveria produzir dois tipos de indicadores, “estimados por etapa, estabelecimento de ensino, rede escolar, unidade da Federação e em nível agregado nacional”:

I - indicadores de rendimento escolar, referentes ao desempenho dos (as) estudantes apurado em exames nacionais de avaliação, com participação de pelo menos 80% (oitenta por cento) dos (as) alunos (as) de cada ano escolar periodicamente avaliado em cada escola, e aos dados pertinentes apurados pelo censo escolar da educação básica;

II - indicadores de avaliação institucional, relativos a características como o perfil do alunado e do corpo dos (as) profissionais da educação, as relações entre dimensão do corpo docente, do corpo técnico e do corpo discente, a infraestrutura das escolas, os recursos pedagógicos disponíveis e os processos da gestão, entre outras relevantes (BRASIL, 2014, p. 2).

Uma das fontes para a elaboração dessa avaliação institucional é o Censo Escolar, por abarcar informações abrangentes sobre diferentes aspectos da oferta escolar, incluindo itens sobre infraestrutura, materiais pedagógicos, equipamentos e docentes. Embora a produção de indicadores relativos ao rendimento escolar tenha sofrido alterações desde a sua criação no início dos anos 1990, e tenha avançado substancialmente no Brasil sobretudo após a proposição do Índice de Desenvolvimento da Educação Básica – o IDEB³, a avaliação institucional pouco caminhou para se firmar com indicadores que permitissem averiguar as condições da oferta educacional para quase 50 milhões de estudantes da educação básica pública e privada, nas etapas e modalidades de ensino.

³ O IDEB foi criado em 2007, sendo considerado o mais proeminente indicador de qualidade da educação básica. Refere-se ao ensino fundamental e médio e toma como referência os resultados da Prova Brasil, aplicada a cada dois anos de forma censitária aos estudantes dos 5º e 9º anos do ensino fundamental e 3º ano do ensino médio, e as informações sobre a aprovação e a evasão coletadas junto às escolas pelo Censo Escolar.

Concebe-se que a infraestrutura das escolas/creches⁴ é um dos elementos que compõem o conceito de condições de trabalho dos docentes⁵. É também um dos critérios a serem considerados para a garantia de bem-estar de profissionais, crianças, jovens e adultos no exercício do seu trabalho e das suas aprendizagens e interações nesses ambientes.

Em recente artigo em que apresenta a proposta de uma escala de infraestrutura escolar, que foi construída utilizando como ferramenta a Teoria de Resposta ao Item e baseando-se em informações referentes às escolas obtidas no Censo Escolar da Educação Básica de 2011, Soares Neto et alii. (2013) fizeram referência a algumas pesquisas que discutem as condições materiais de escolas brasileiras e também trazem alguns estudos internacionais. De modo geral, os estudos brasileiros mostram a relevância da infraestrutura das escolas para o aprendizado dos alunos. Apontam que este é um dos aspectos em que se pode ressaltar as desigualdades da oferta educacional no Brasil e os baixos resultados nos testes de avaliação de larga escala, associados a esta variável.⁶

Levantamento preliminar de literatura brasileira evidencia também a existência de poucas referências de estudos que relacionam infraestrutura escolar e satisfação profissional. Entre os trabalhos, pode-se incluir o de Batista e Odelius (1999) e o de Rebolo e Bueno (2014). Esses últimos analisaram os fatores de satisfação no trabalho e as estratégias que geram e mantêm o bem-estar docente, para compreender o modo como os professores podem construir a felicidade no trabalho. O estudo, baseado em modelo teórico específico para analisar respostas de 250 professores de escolas públicas (questionários e grupos focais), permitiu identificar o grau de satisfação dos professores em múltiplos aspectos do trabalho docente, a relação de tais aspectos com a autopercepção de felicidade e as estratégias de enfrenta-

⁴ Neste trabalho, utilizaremos doravante o termo "unidade educacional" para nos referirmos às escolas, creches e pré-escolas. Isso se justifica pelo fato de creches e pré-escolas não serem consideradas "escolas" no sentido estrito do termo e apresentarem estrutura física e de materiais, proposta pedagógica e curricular que as diferenciam do modelo e forma escolar difundidos mundialmente.

⁵ Adotamos, neste trabalho, a definição de condições de trabalho docente sintetizada por Oliveira e Assunção (2010). Na sua síntese, as condições de trabalho docente compreendem o contexto sócio-histórico no qual se situam as relações de trabalho; passa pelo sistema escolar – municipal, estadual, federal, privado –, pela estrutura física da unidade educacional, pelas normas que organizam a dinâmica das interações na instituição, pelos recursos disponíveis para a realização das atividades, pelas condições de emprego – cargo, função, vínculo de contratação, carga horária de trabalho, salário, plano de carreira, formação continuada, passa também pela experiência relacional entre os sujeitos docentes e com os sujeitos discentes, considerando-se ainda as percepções e os efeitos objetivos e subjetivos de todo esse conjunto de fatores e processos.

⁶ Estudo publicado em 2019, pela UNESCO-Brasil, desenvolveu indicadores de infraestrutura escolar para o ensino fundamental, trazendo contribuições para o desenvolvimento de parâmetros para a avaliação institucional da educação básica no Brasil (UNESCO, 2019).

mento utilizadas em situações de insatisfação e conflito na escola. As análises indicam que o bem-estar docente é o resultado positivo da avaliação cognitiva e afetiva que o professor faz de si próprio e das condições existentes para a realização de sua atividade laboral, um processo que, para ser construído, exige esforços por parte do professor e condições de trabalho que compensem esses investimentos.

Em escala nacional, a principal fonte de dados existente sobre as condições de infraestrutura e equipamentos das unidades educacionais brasileiras é o Censo Escolar. Trata-se de um levantamento anual de dados estatísticos educacionais, de abrangência nacional, coordenado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), órgão do Ministério da Educação responsável pelas informações educacionais e pelo sistema nacional de avaliação da educação básica e superior. O Censo Escolar coleta informações sobre “escolas”, “turmas”, “matrículas” e “docentes” da educação básica regular, além da educação de jovens e adultos, da educação especial e da educação profissional.

Apesar da abrangência nacional das informações, estas se limitam às características mais gerais de uma unidade educacional, como, por exemplo, tipo de abastecimentos de água e de energia elétrica; esgoto sanitário e destinação do lixo; se há ou não determinadas dependências, como salas de diretor e de professores, laboratórios de ciência e de informática; quantidade de determinados equipamentos, como aparelhos de televisão e computadores; e existência ou não de alguns materiais de atendimento à diversidade socio-cultural, quilombola e indígena. Os dados mais específicos da primeira etapa da educação básica são: “parque infantil”, “berçário” e “banheiro adequado à educação infantil”. Entretanto, não há informações a respeito da qualidade desses espaços e instalações. É levantado somente o tipo de determinados serviços e se há ou não certas dependências e equipamentos. Assim, não podemos saber sobre a sua adequação, conservação e qualidade para o desenvolvimento do projeto pedagógico das unidades educacionais.

Este trabalho se baseia em resultados da pesquisa Trabalho Docente na Educação Básica no Brasil (TDEBB) – Fase II, coordenada pelo Grupo de Estudos sobre Política Educacional e Trabalho Docente (GESTRADO), da Universidade Federal de Minas Gerais (UFMG), Brasil, finalizada no ano de 2015 (GESTRADO, 2015). A coleta de dados ocorreu na primeira fase da

pesquisa⁷, realizada no período 2009-2010 (OLIVEIRA; VIEIRA, 2012). Refere-se às respostas de professores da educação básica, nas quais ressaltamos algumas especificidades relacionadas às docentes da educação infantil.

Dados desta pesquisa foram utilizados por Oliveira e Pereira Junior (2016) para analisar as múltiplas associações existentes entre seis construtos referentes aos contextos escolares dos professores da educação básica, entre os quais encontram-se as condições da sala de aula, as condições da unidade educacional e a satisfação profissional. Embora tenham sido exploradas as associações entre tais construtos, inexistiu qualquer tipo de análise que explorasse as especificidades de cada etapa da educação básica, conforme almeja o presente estudo, em que se busca destacar a situação vivenciada na educação infantil.

No presente estudo, buscamos analisar a percepção de professores da educação básica⁸ sobre as condições da infraestrutura escolar e a sua satisfação profissional, ressaltando as especificidades relacionadas às professoras da educação infantil (EI). A infraestrutura escolar é entendida como o local de trabalho, abrangendo três dimensões: as Condições da sala de aula, as Condições da unidade educacional e a Quantidade de alunos por turma. A satisfação profissional é analisada como sendo o sentimento de realização dos professores em relação ao desenvolvimento de suas atividades docentes e às perspectivas direcionadas ao futuro profissional. A cada uma das três dimensões da infraestrutura escolar, juntamente com a satisfação profissional, foi desenvolvido um indicador, de forma a permitir analisar empiricamente a situação verificada nas etapas da educação básica, enfatizando a educação infantil.

O trabalho está organizado da seguinte forma: na seção 1, apresenta-se a metodologia de análise do banco de dados da pesquisa TDEBB e de construção

⁷ Trata-se da pesquisa Trabalho Docente na Educação Básica no Brasil (TDEBB) - Fase I, que consistiu em um survey realizado nos anos de 2009 e 2010 e que entrevistou 8.795 sujeitos docentes da educação básica de sete estados brasileiros: Espírito Santo, Goiás, Minas Gerais, Pará, Paraná, Rio Grande do Norte e Santa Catarina. A pesquisa TDEBB (Fases I e II) teve como objetivo analisar o trabalho docente nas duas dimensões constitutivas, identificando seus atores, o que fazem e em que condições se realiza o trabalho nas escolas de educação básica da rede pública e conveniada, tendo como finalidade subsidiar a elaboração de políticas públicas no Brasil. Investigou em que medida as mudanças trazidas pela nova regulação educativa impactam na constituição de identidades e dos perfis de profissionais da educação básica, identificando estratégias desenvolvidas pelos docentes para responder a novas exigências.

⁸ No Brasil, a educação escolar é organizada em dois grandes níveis: educação básica e educação superior. A educação básica é dividida em três etapas: educação infantil, primeira etapa; ensino fundamental, segunda etapa; ensino médio, terceira etapa. A educação infantil é oferecida em creches/guarderías, para crianças na idade de 0 a 3 anos, e em pré-escolas, para crianças de 4 e 5 anos de idade. O ensino fundamental tem duração de nove anos para crianças e jovens de 6 a 14 anos de idade; o ensino médio, com duração de três anos, acolhe jovens de 15 a 17 anos de idade. A obrigatoriedade escolar abrange pessoas na idade de 4 a 17 anos, desde a pré-escola até o ensino médio.

dos indicadores de percepção sobre a infraestrutura escolar e a satisfação profissional entre docentes das três etapas da educação básica. Na seção 2, buscamos interpretar os resultados à luz de algumas informações contextuais da educação básica presentes no Censo Escolar e de literatura acadêmica relacionada, ressaltando que se trata de um trabalho exploratório sobre o tema.

METODOLOGIA DO ESTUDO E INDICADORES DE PERCEPÇÃO DOS PROFESSORES: RESULTADOS

Antes de abordar a metodologia utilizada para desenvolver os indicadores, é necessário delimitar a amostra a ser analisada. O desenvolvimento dos indicadores produzidos na segunda fase da pesquisa TDEBB tomou como base as respostas de 6.684 profissionais, conforme consta no relatório Trabalho Docente na Educação Básica no Brasil (TDEBB) – Fase II (GESTRADO, 2015). Embora tenham sido entrevistados um total de 8.795 sujeitos docentes⁹ na primeira fase da pesquisa, o desenvolvimento dos indicadores considerou as respostas apenas dos professores da educação básica, pois são esses os profissionais responsáveis formais pela(s) turma(s) em que atua(m). Em relação à educação infantil, na categoria de “professores”, foram englobados também os profissionais classificados como “educadores infantis”, pois são eles, assim como os professores, os responsáveis pela realização das atividades junto às crianças nessa etapa da educação básica.

Ao verificar a(s) etapa(s) da educação básica em que atuam, encontramos um total de 804 professores que indicaram trabalhar em mais de uma etapa de ensino. Considerando que o objetivo deste estudo é analisar a especificidade da educação infantil em comparação com as outras etapas da educação básica, foi estabelecida a delimitação adicional de considerar exclusivamente os professores que atuam em uma única etapa da educação básica. Feito isso, a amostra a ser analisada fica compreendida por 1.255 (21,3%) professoras e educadoras infantis que atuam na educação infantil (EI), 3.487 (59,3%) no ensino fundamental (EF) e 1.138 (19,4%) no ensino médio (EM).

⁹Na TDEBB - Fase I e Fase II, partiu-se do entendimento de que na educação básica há uma grande variedade de profissionais e situações de trabalho nas escolas nas suas diferentes etapas e modalidades. Assim, foram considerados sujeitos da pesquisa: professores, educadores infantis, educadores, monitores, estagiários, diretores, coordenadores, supervisores, atendentes, auxiliares etc. Foram designados com o termo de sujeitos docentes. Na pesquisa, não foram incluídos como sujeitos docentes os trabalhadores de apoio administrativo e serviços gerais.

Em relação à metodologia utilizada para desenvolver os indicadores, foi empregada a técnica estatística multivariada denominada Modelagem de Equações Estruturais (MEE). A MEE é uma

técnica multivariada que combina aspectos de análise fatorial e de regressão múltipla que permite ao pesquisador examinar simultaneamente uma série de relações de dependência inter-relacionadas entre as variáveis medidas e construtos latentes (variáveis latentes), bem como entre diversos construtos latentes (HAIR et al., 2009, p. 542).

Nessa etapa, foi utilizada apenas a análise fatorial do tipo confirmatória, que exige do pesquisador a especificação das variáveis associadas a cada construto. Em cada indicador, a ponderação de seus itens ou variáveis foi feita a partir dos coeficientes de regressão padronizados, considerando como variável dependente o item e como independente o construto ou conceito teórico associado.

O nível de qualidade dessas medidas foi avaliado, conforme apontaram Hair *et al.* (2009), a partir do cumprimento de quatro critérios: 1) definição conceitual; 2) dimensionalidade; 3) confiabilidade e 4) validade. A *definição conceitual* – ponto de partida para criar a escala múltipla – especifica a base teórica para a escala múltipla definindo o conceito a ser representado em termos aplicáveis ao contexto da pesquisa” (HAIR et al., 2009,p. 125). A *dimensionalidade* avalia se os itens estão fortemente associados um com o outro e se representam um só conceito (unidimensionais). A *confiabilidade* se refere à avaliação do grau de consistência entre múltiplas medidas de uma variável. Como os itens devem medir o mesmo construto, isto exige que sejam intercorrelacionados. A *validade* caracteriza o grau em que uma escala ou um conjunto de medidas representa com precisão o conceito de interesse.

Os indicadores foram produzidos em uma escala cujo intervalo varia de zero a um, sendo que o valor mínimo representa a pior situação possível e o valor máximo, a melhor. Além de facilitar o entendimento e interpretação dos indicadores e colocar todos os resultados sob uma escala única, favorece a comparação entre eles.

Os indicadores analisados neste estudo referem-se à: 1) Percepção das condições da sala de aula; 2) Percepção das condições da unidade educacional; 3) Satisfação profissional e 4) Quantidade de alunos por turma (média). Tais indicadores foram desenvolvidos na pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015) e a descrição de cada um deles se encontra adiante.

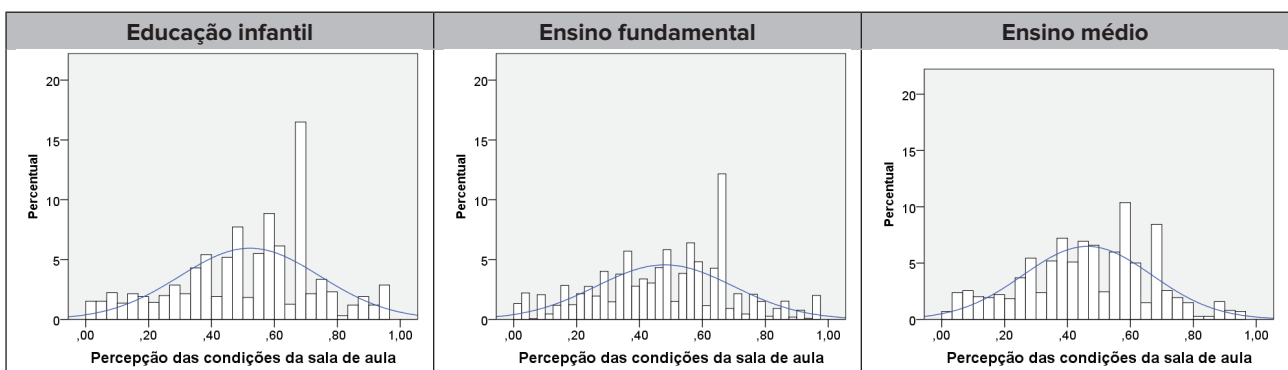
INDICADORES DE PERCEÇÃO DOS PROFESSORES

O indicador de *Condições da sala de aula* exprime o “nível de adequação da sala de aula – local onde, geralmente, os professores passam a maior parte do tempo de trabalho” (GESTRADO, 2015, p. 131). O desenvolvimento da atividade docente necessita de condições apropriadas de aspectos ambientais e estruturais, afetando tanto alunos quanto professores. O indicador é composto por quatro itens: 1) Ventilação; 2) Iluminação; 3) Condições das paredes; e 4) Ruído originado dentro da sala de aula.

A distribuição dos professores de cada etapa da educação básica, de acordo com o indicador de Percepção das condições da sala de aula, é apresentada no Gráfico 1. Em relação aos resultados obtidos, a média do indicador foi equivalente a 0,52 na educação infantil, 0,48 no ensino fundamental e 0,46 no ensino médio.

Servilha, Leal e Hidaka (2010) estudaram a legislação trabalhista brasileira sobre os riscos ocupacionais relacionados à saúde e à voz dos professores e identificaram as queixas dos professores sobre as condições de trabalho. Segundo as autoras, esses profissionais se encontravam expostos a altos níveis de ruído, ao desconforto térmico (existindo salas de aula muito frias no inverno e muito quentes no verão), à pouca ventilação, à presença de umidade, à higiene deficitária e à iluminação deficiente.

Gráfico 1 – Distribuição dos professores de acordo com os resultados do indicador de Percepção das condições da sala de aula



Fonte: Elaboração dos autores, com base nos dados da pesquisa Trabalho Docente na Educação Básica no Brasil – Fase II (GESTRADO, 2015).

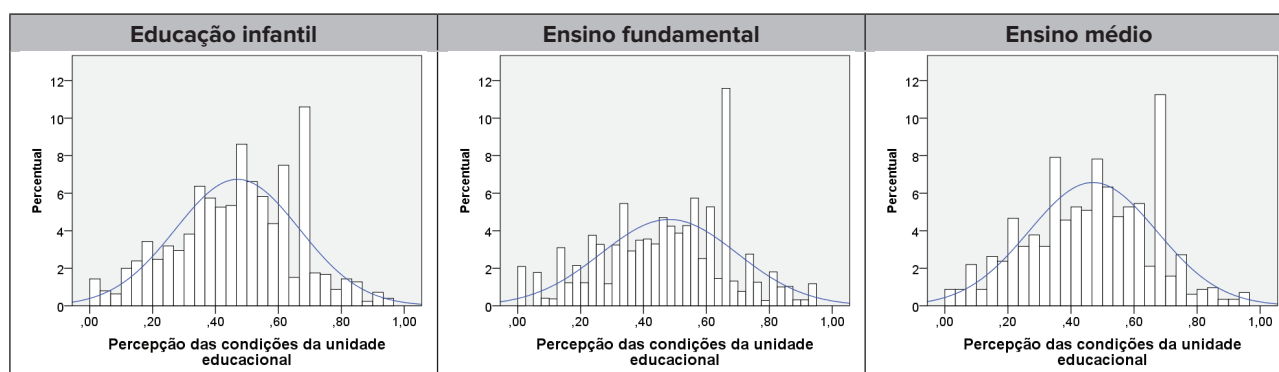
O indicador de *Condições da unidade educacional* representa a “avaliação dos professores sobre aspectos da infraestrutura das escolas” (GESTRADO, 2015, p. 135), pois a adequação dos estabelecimentos e a disponibilidade de materiais e equipamentos favorece o desenvolvimento da atividade docente, afetando tanto alunos quanto professor(es). O indicador contempla a avalia-

ção dos professores em relação a quatro itens: 1) Sala específica de convivência e repouso; 2) Banheiros para funcionários; 3) Equipamentos; e 4) Recursos pedagógicos.

Considerando os resultados obtidos, a média do indicador de Percepção das condições da unidade educacional foi equivalente a 0,47 na educação infantil, 0,48 no ensino fundamental e 0,47 no ensino médio. A distribuição dos professores de cada etapa da educação básica em relação a esse indicador é apresentada no Gráfico 2.

Araújo e Carvalho (2009) analisaram estudos epidemiológicos sobre as condições de trabalho dos professores da rede particular de ensinos fundamental e médio de Salvador (BA), tendo sido constatado que 55,7% das escolas não possuíam local para descanso dos professores.

Gráfico 2 – Distribuição dos professores de acordo com os resultados do indicador de Percepção das condições da unidade educacional



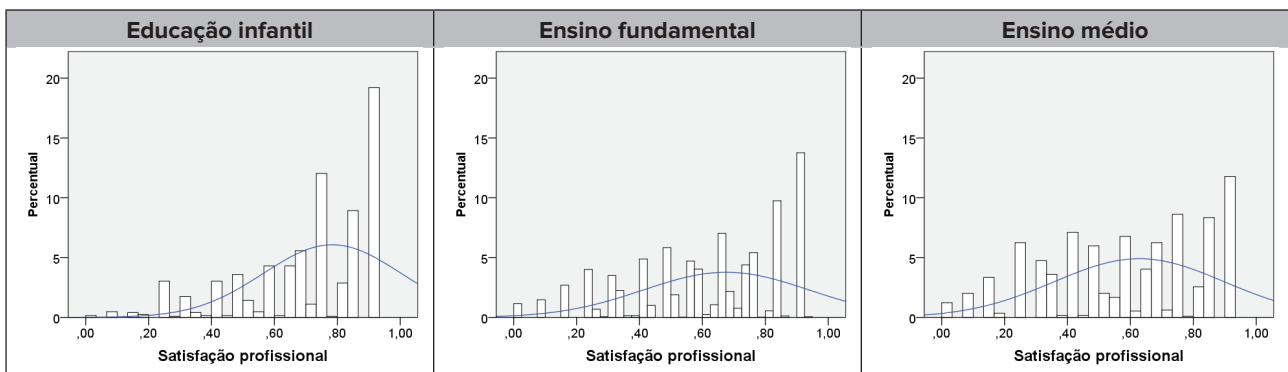
Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015).

O indicador de Satisfação Profissional representa o “nível de realização que os professores de educação básica sentem ao desenvolver suas atividades e as perspectivas em relação ao futuro profissional” (GESTRADO, 2015, p. 143). Além de se relacionar à motivação e atitude desses profissionais diante do trabalho, pode apontar propensão ao abandono da carreira. O indicador é composto por quatro itens: 1) Frustração com o trabalho; 2) Pensa em parar de trabalhar na educação; 3) Trabalhar na educação proporciona grandes satisfações; e 4) Escolheria trabalhar em educação se tivesse que recomeçar a vida profissional.

A distribuição dos professores de cada etapa da educação básica, de acordo com o indicador de Satisfação profissional, é apresentada no Gráfico 3. Em relação aos resultados obtidos, a média do indicador foi equivalente a 0,78 na educação infantil, 0,68 no ensino fundamental e 0,63 no ensino médio.

Os resultados convergiram, de certa forma, ao estudo de Soratto e Olivier-Heckler (1999), que avaliaram a satisfação no trabalho entre professores e constataram que o maior percentual de satisfeitos (91%) se encontrava entre aqueles que lecionavam na pré-escola ou nos anos iniciais do ensino fundamental. Os autores mensuraram a satisfação no trabalho por meio das seguintes questões: 1. Eu me arrependo de ter escolhido esta profissão; 2. Quaisquer que sejam os problemas do meu trabalho, estou satisfeito com a minha escolha; 3. Se eu pudesse, mudaria de emprego; 4. Em geral, estou satisfeito com o meu emprego atual; e 5. Muitas vezes, quando estou trabalhando, sinto que estou perdendo tempo.

Gráfico 3 – Distribuição dos professores de acordo com os resultados do indicador de Satisfação profissional



Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015).

Alternativamente, visando a facilitar o entendimento dos indicadores, seus resultados foram classificados em quatro categorias: 1) *Muito baixo* – que contempla os professores com indicadores que variam de 0,00 a 0,25; 2) *Baixo* – englobando aqueles com indicadores de 0,25 a 0,50; 3) *Alto* – composto pelos professores com indicadores de 0,50 a 0,75; e 4) *Muito alto* – que abrange os resultados acima de 0,75.

A distribuição dos professores pesquisados, de acordo com a classificação nos indicadores analisados e a etapa exclusiva da educação básica em que atuam, é mostrada na Tabela 1.

Tabela 1 – Distribuição dos professores de acordo com a etapa de ensino em que atuam e com as categorias dos indicadores de Percepção das condições da sala de aula, Percepção das condições da unidade educacional e Satisfação profissional

Etapa exclusiva	Muito baixo	Baixo	Alto	Muito alto	Total
Percepção das Condições da sala de aula					
Educação infantil	13,2%	30,4%	43,0%	13,3%	100,0%
Ensino fundamental	16,3%	36,3%	37,7%	9,8%	100,0%
Ensino médio	15,9%	40,2%	36,8%	7,0%	100,0%
Percepção das Condições da unidade educacional					
Educação infantil	14,7%	39,8%	39,2%	6,4%	100,0%
Ensino fundamental	16,1%	35,4%	39,6%	8,9%	100,0%
Ensino médio	15,6%	39,6%	38,9%	5,8%	100,0%
Satisfação profissional					
Educação infantil	3,8%	9,6%	24,0%	62,5%	100,0%
Ensino fundamental	9,3%	18,5%	26,3%	45,9%	100,0%
Ensino médio	12,6%	22,4%	26,2%	38,8%	100,0%

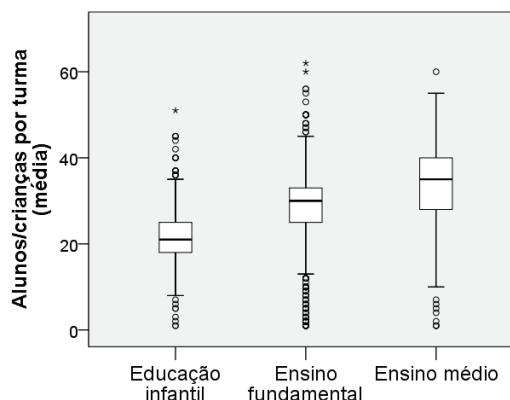
Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015).

O outro indicador analisado é a variável *Alunos por turma*, que mensura a quantidade de alunos que os professores possuem por turma. No caso de possuírem mais de uma turma, o instrumento de coleta de dados captou a média de discentes entre elas. A distribuição dos professores de cada etapa (exclusiva) da educação básica é exibida no Gráfico 4, cuja mediana¹⁰ daqueles que atuam na educação infantil é de 21 alunos por turma, contra 30 do ensino fundamental e 35 do ensino médio.

Situação mais favorável à educação infantil, em relação ao número de alunos por turma, também foi encontrada entre os professores das redes públicas de Pernambuco. Enquanto 37% dos professores da educação infantil possuíam até 20 alunos por turma, no ensino fundamental esse valor foi de 20% e, no ensino médio, 4% (OLIVEIRA; VIEIRA, 2014).

¹⁰ No boxplot, a linha dentro da caixa representa a mediana de um conjunto de dados. Por sua vez, a mediana permite afirmar que 50% dos valores de um conjunto de dados se encontram acima e outros 50% abaixo de seu valor numérico.

Gráfico 4 – *Boxplot* da distribuição dos professores de acordo com a etapa de atendimento exclusiva em que atuam



Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015).

Mais informações sobre os resultados da variável Quantidade de alunos/criança por turma é mostrada na Tabela 2. Os valores máximos de discentes por turma foram referentes a 51 na educação infantil, 62 no ensino fundamental e 60 no ensino médio.

TABELA 2 – Estatísticas descritivas da Quantidade de alunos/crianças por turma (média)

Etapa exclusiva de atendimento	Estatísticas descritivas				
	Mínimo	Máximo	Mediana	Média	Desvio-padrão
Educação infantil	1	51	21	21,4	6,1
Ensino fundamental	1	62	30	28,5	7,3
Ensino médio	1	60	35	32,5	8,2

Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015),1

A análise dos resultados das três etapas da educação básica será feita na seção seguinte, amparada na literatura acadêmica e ressaltando as características educação infantil.

DISCUSSÃO DOS RESULTADOS: ANÁLISES EXPLORATÓRIAS E CONSIDERAÇÕES FINAIS

Ao analisar as condições da sala de aula, verificamos que as professoras da educação infantil as avaliam de forma mais positiva que as professoras das outras etapas da educação básica. A média registrada desse indicador foi de 0,52 entre as docentes da primeira etapa da educação básica, 0,48 do ensino fundamental e 0,46 do ensino médio. Ainda, os percentuais de entrevistados das três etapas, cujos indicadores foram classificados em alto ou muito alto, contabilizam 56,3%, 47,5% e 43,8%, respectivamente. Os resultados apontam que, na percepção dos docentes, existe um maior grau de adequação da sala de aula, para o desenvolvimento das atividades, considerando a avaliação dos docentes sobre os aspectos ambientais e estruturais (ventilação, iluminação, condições das paredes e ruídos).

Sobretudo em relação aos aspectos ambientais investigados, um possível fator que pode influenciá-los é o tamanho das turmas, pois uma quantidade alta de alunos em uma mesma sala pode prejudicar a situação encontrada de ventilação, iluminação e ruídos. Sob esse prisma, a quantidade verificada de alunos/crianças por turma acompanha os resultados do indicador de Percepção das condições da sala de aula, registrando menor média (21,4) na educação infantil e maior (32,5) no ensino médio. Isto é, ocorre a relação inversa entre os dois indicadores, sendo que, à medida que um aumenta, o outro diminui e vice-versa. Cabe ressaltar que não foi realizado cruzamento entre os dois indicadores – tamanho da turma e condições da sala de aula, pois foram analisados separadamente. Mas podemos observar que no EF a criança, geralmente, encontra uma sala de aula com espaço físico maior e uma turma mais numerosa. Além disso, haverá menos adultos à disposição (geralmente, só uma professora por turma). Na pesquisa TDEBB – Fase I, pode-se apurar se os docentes contariam com apoio de pessoal para acompanhamento de seus alunos/crianças. Os respondentes da EI foram os que evidenciaram receber maior apoio: EI – 74,3%, EF – 59,8% e EM – 47,5%. Entre o pessoal de apoio, encontram-se estagiários, pedagogos, outro professor, auxiliares, educadores especiais, coordenadores, entre outros (GESTRADO, 2010). Isso pode nos ajudar a entender a percepção mais positiva dos docentes da EI em relação às condições da sala de aula.

Além disso, o número de alunos por turma foi considerado o segundo aspecto mais importante para melhorar a qualidade do trabalho docente, pelos respondentes das três etapas da educação básica na pesquisa TDEBB – Fase I (GESTRADO, 2010), conforme se mostra na Tabela 3 abaixo. O que pode indi-

car que a percepção das condições da sala de aula seja associada ao número de alunos/crianças por turma, a cargo de um docente.

Tabela 3 – Aspectos mais importantes⁽¹⁾ para melhorar a qualidade do trabalho de acordo com a etapa da educação básica em que atuam os professores

Aspectos	Educação infantil	Ensino fundamental	Ensino médio
Receber melhor remuneração (ter aumento de salário)	75,1%	76,5%	76,6%
Reduzir o número de alunos/crianças por turma	62,1%	63,4%	58,4%
Receber mais capacitação para as atividades que exerce	57,6%	47,8%	39,6%
Aumentar o número de horas destinadas às atividades extraclasse	20,6%	26,8%	41,3%
Ter dedicação exclusiva a uma única unidade educacional	24,5%	34,1%	42,1%
Contar com maior apoio técnico nas suas atividades	33,9%	31,3%	25,0%
Outras	5,3%	2,7%	2,4%

Fonte: Elaboração dos autores, com base nos dados da pesquisa *Trabalho Docente na Educação Básica no Brasil – Fase II* (GESTRADO, 2015).

Nota: (1) São permitidas até três opções de resposta.

Ao ampliar o escopo de análise, passando a considerar toda a unidade educacional, os indicadores de Percepção das condições da unidade educacional apresentam bastante semelhança entre as etapas da educação básica. As médias encontradas desse indicador correspondem a 0,47 na educação infantil e no ensino médio, e 0,48 no ensino fundamental. Dois fatores podem atuar nesses resultados. O primeiro, de forma mais geral, é que mais da metade das unidades de educação básica no Brasil oferece mais de uma etapa de atendimento. Assim, dentro de uma escola, esperaríamos condições semelhantes a todas as etapas por ela ofertadas. O segundo é que as necessidades da educação infantil em relação à estrutura do estabelecimento são bem específicas, quando em comparação ao ensino fundamental e médio. Como já mencionamos, tais estruturas compreendem os parques infantis, os berçários e os banheiros adequados à EI, enquanto o indicador de Percepção das condições da sala de aula contempla sala específica de convivência e repouso, banheiros para funcionários, equipamentos e recursos pedagógicos.

Apresentamos, na Tabela 4 a seguir, algumas informações selecionadas do Censo Escolar (anos 2013 e 2014) sobre disponibilidade recursos – espaços, serviços e equipamentos nas unidades educacionais, que nos ajudam a contextualizar a percepção dos docentes sobre as mesmas.

Tabela 4 – Educação básica: unidades educacionais das redes públicas, segundo recursos disponíveis, Brasil, 2014 (em %)

Recursos	Educação Infantil		Ensino Fundamental	Ensino Médio
	Creche	Pré-escola		
Água Filtrada	93,2	82,3	80,2	89,2
Abastecimento de água				
• Rede pública	76,6	53,6	54,4	88,9
• Poço artesiano	13,4	17,8	17,6	13,0
• Cacimba/cisterna/poço	10,4	15,0	13,7	3,8
• Outros	2,9	7,4	8,2	1,5
• Inexistente	6,3	9,0	7,5	0,4
Esgoto sanitário				
• Rede Pública	45,6	27,7	29,5	59,3
• Fossa	57,6	64,4	60,6	44,2
• Inexistente	4,3	8,0	8,2	0,7
Acesso à energia elétrica	98,5	94,5	93,2	99,9
Acesso à internet	50,7 ¹¹		48,8	
Parque infantil	43,4	24,6	-	-
Banheiro adequado à EI	46,7	24,7	-	-
Banheiro dentro do prédio	90,6	79,8	80,0	97,4
Biblioteca e/ou sala de leitura	14,0	12,3	43,7	87,7
Laboratório de Ciências	-	-	8,1	44,6
Quadra de esportes			32,5	76,3

Fonte: MEC/Inep/DEED - Microdados do Censo Escolar de 2013 e 2014.

Nota: a mesma unidade educacional (estabelecimento) pode possuir mais de um tipo de abastecimento de água ou esgoto sanitário.

O Censo não informou sobre disponibilidade de livros, materiais diversos adequados ao desenvolvimento de atividades pedagógicas concernentes a cada etapa de ensino.

Os dados constantes da Tabela 4 mostram que, em relação à presença de parque infantil, biblioteca e/ou sala de leitura, banheiro específico à EI, a creche mostra condições mais adequadas que a pré-escola. Em relação aos itens abastecimento de água (rede pública), esgoto (rede pública), água filtrada também está em posição mais adequada à creche em relação à pré-escola e ao ensino fundamental. De modo geral, o ensino médio tem melhores percentuais de presença de recursos que o ensino fundamental, que tem situação semelhante à pré-escola. Isso se deve a que mais de 65% da oferta de EI acontece na unidade educacional em que funciona concomitantemente o ensino fundamental. Tais informações podem ajudar a

¹¹ Refere-se ao Censo Escolar do ano de 2013.

explicar a semelhança dos indicadores de Percepção das condições da unidade educacional nas três etapas da educação básica.

O último indicador analisado refere-se à Satisfação Profissional, que abrange sentimentos dos professores em relação ao desenvolvimento de suas atividades e às perspectivas direcionadas ao futuro na profissão. Os resultados indicam um grau mais elevado de satisfação das professoras da EI, cuja média foi de 0,78 e do ensino fundamental e do ensino médio, 0,68 e 0,63, respectivamente. Ao analisar as proporções de docentes classificados na categoria “muito alto”, a educação infantil apresenta 62,5%, o ensino fundamental, 45,9%, e o ensino médio, 38,8%.

O grau elevado de satisfação entre as professoras da educação infantil, comparando-se com aqueles do ensino fundamental e médio, necessita de muitos aprofundamentos. A categoria satisfação profissional é complexa pela sua associação a diferentes variáveis objetivas e subjetivas (MARQUEZE; MORENO, 2005). Embora as docentes da EI, em relação à valorização profissional (considerando-se a remuneração e a carreira – plano de cargos e salários), apresentem situação desigual, pois entre os docentes da educação básica brasileira são aquelas que recebem os menores salários, para jornadas de trabalho e formação equivalentes e possuem as carreiras mais desvantajosas (OLIVEIRA; VIEIRA, 2012), foram aquelas que apresentaram índice de respostas mais elevado em relação à satisfação profissional. Os aspectos relacionais na educação infantil podem ser significativos na determinação da satisfação profissional, pois permeiam de forma intensa o trabalho docente nesta etapa da educação básica (CARVALHO, 2014; FOLLE *et al.*, 2008). Com efeito, na pesquisa TDEBB (Fase I) (GESTRADO, 2010), 81% dos respondentes da EI concordaram com o seguinte enunciado: "a minha relação com meus alunos/crianças é em base afetiva". Os respondentes do EF e EM apresentaram, respectivamente, as seguintes frequências em relação ao mesmo enunciado: 66% e 46%. São também aqueles que mais concordaram sentir que realizam um trabalho que é socialmente valorizado (36% EI, 30% EF e 25% EM). Ao realizar um balanço sobre produção acadêmica recente sobre trabalho docente na EI, Vieira e Oliveira (2013, p.149) verificaram que

os sentimentos de afeto pelas crianças e os sentimentos positivos que as crianças nutrem pelas suas professoras na educação infantil são elementos gratificantes reiteradamente citados pelas pesquisadas, nos estudos analisados, para permanecer e buscar melhores condições de trabalho nas instituições educacionais.

De forma preliminar, nós podemos interpretar esse resultado pelas características pedagógicas da organização da EI, pela presença maciça de mulheres trabalhando em creches e pré-escolas, e pelas bases históricas e culturais da educação de bebês e crianças pequenas, que diferem da escola obrigatória elementar e secundária (ensino médio).

Sobre o tipo de associação entre os indicadores analisados, a satisfação profissional é considerada variável dependente em relação às percepções sobre as condições tanto da sala de aula quanto da unidade educacional (OLIVEIRA; PEREIRA JUNIOR, 2016). Tendo apresentado coeficientes positivos, isso significa que quanto maior a percepção sobre as condições da sala de aula, maior foi a medida de satisfação profissional. O mesmo ocorreu em relação às percepções das condições da unidade escolar. Diante disso, os resultados obtidos neste estudo mostraram que os professores da educação infantil tiveram melhores indicadores das condições da sala de aula e da relação de alunos por turma. Ou seja, o investimento no local onde os professores passam a maior parte do tempo de trabalho possui o potencial de melhorar a satisfação profissional dos professores.

A exploração do banco de dados da pesquisa TDEBB – Fase I e II (GESTRADO, 2010; 2015), sobre a temática aqui tratada, requer o aprofundamento de estudos e pesquisas na literatura nacional e internacional sobre as condições do trabalho docente na educação básica, ressaltando a associação entre a infraestrutura escolar e a satisfação profissional.

REFERÊNCIAS

ARAÚJO, T. M.; CARVALHO, F. M. Condições de trabalho docente e saúde na Bahia: estudos epidemiológicos. **Educação e Sociedade**, v. 30, n. 107, p. 427-449, maio/ago., 2009.

BATISTA, A. S.; ODELIUS, C. C. Infra-estrutura das escolas e burnout dos professores. In: CODO, W. (Ed.). **Educação: carinho e trabalho**. Petrópolis: Vozes, 1999.

BRASIL. Conselho Nacional de Educação. Resolução CEB/CNE n. 05/2009, de 17 de dezembro de 2009. **Dispõe sobre as Diretrizes Curriculares Nacionais da Educação Infantil**. Brasília, 2009.

_____. Ministério da Educação. **Parâmetros Nacionais de Qualidade para a Educação Infantil**. Brasília: MEC/SEB, 2006a.

_____. Ministério da Educação. **Parâmetros Básicos de Infraestrutura para Instituições de Educação Infantil**. Brasília: MEC/SEB, 2006b.

_____. 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, 26 jun. 2014. Edição Extra, Seção 1.

CAMPOS, M. M.; FÜLLGRAF, J.; WIGGERS, V. A qualidade da Educação Infantil: um estudo em seis capitais brasileiras. **Cadernos de Pesquisa**, São Paulo, v. 41, n. 142, p. 20-54, jan./abr., 2011.

CARVALHO, Rodrigo S. O imperativo do afeto na educação infantil: a ordem do discurso de pedagogas em formação. **Educação e Pesquisa**, São Paulo, v. 40, n. 1, p. 231-246, jan./mar., 2014.

FALCIANO, Bruno T.; SANTOS, Edson C.; NUNES, Maria Fernanda R. Infraestrutura escolar: um critério de comparação da qualidade na educação infantil. **Estudos de Avaliação Educacional**, São Paulo, v. 27, n. 66, p. 880-906, set./dez., 2016.

FOLLE, Alexandra; BORGES, Lucélia J.; COQUEIRO, Raildo S.; NASCIMENTO, Juarez V. Nível de (in)satisfação profissional de professores de Educação Física da Educação Infantil. **Motriz**, Rio Claro, v. 14, n. 2, p. 124-134, abr./jun., 2008.

GESTRADO. Grupo de Estudos Sobre Política Educacional e Trabalho Docente. Relatório de pesquisa. **Trabalho docente na Educação Básica no Brasil – Fase II**. Belo Horizonte: UFMG, 2015.

_____. Grupo de Estudos Sobre Política Educacional e Trabalho Docente. Base de dados. **Trabalho docente na Educação Básica no Brasil – 2009**. Belo Horizonte: UFMG, 2010.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise Multivariada de Dados**. 6. ed.. Tradução Adonai Schlup Sant'Anna, Porto Alegre: Bookman, 2009.

MARQUEZE, Elaine C.; MORENO, Cláudia R. C. Satisfação no trabalho: uma breve revisão. São Paulo. **Revista Brasileira de Saúde Ocupacional**, v.30, n. 112, p. 69-79, 2005.

OLIVEIRA, D. A.; PEREIRA JUNIOR, E. A. Indicadores do trabalho docente: múltiplas associações no contexto escolar. **Estudos em Avaliação Educacional**, v. 27, n. 66, p. 852-878, set./dez., 2016.

OLIVEIRA, D. A.; VIEIRA, L. F. (orgs). **Trabalho na educação básica em Pernambuco**. Camaragibe, PE: CCS Gráfica e Editora, 2014.

_____. **O trabalho na educação básica:** a condição docente em sete estados brasileiros. Belo Horizonte: Fino Traço, 2012.

OLIVEIRA, T. G. de. As condições das creches públicas e conveniadas com o poder público no Brasil. **Revista Ibero-americana de Educação**. v. 71, p. 73-62, maio 2016.

RAYNA, S.; BROUGÈRE, G. (sous la direction). **Traditions et innovations dans l'éducation préscolaire:** perspectives internationales. INRP: Paris, 2009.

REBOLO, F.; BUENO, B. O. O bem-estar docente: limites e possibilidades para a felicidade do professor no trabalho. **Acta Scientiarum. Education**, Maringá, v. 36, n. 2, p. 323-331, jul./dez., 2014.

SERVILHA, E. A. M.; LEAL, R. O. F.; HIDAKA, M. T. U. Riscos ocupacionais na legislação trabalhista brasileira: destaque para aqueles relativos à saúde e à voz do professor. **Revista da Sociedade Brasileira de Fonoaudiologia**, v. 15, n. 4, p. 505-513, 2010.

SOARES NETO, Joaquim J. et ali. Uma escala para medir a infraestrutura escolar. **Estudos de Avaliação Educacional**. São Paulo, v. 24, n. 54, p. 78-99, jan./abr., 2013.

SORATTO, L.; OLIVIER-HECKLER, C. Trabalho: atividade humana por excelência. In: CODO, W. (Coord.). **Educação:** carinho e trabalho. Petrópolis, RJ: Vozes – Brasília: Confederação Nacional dos Trabalhadores em Educação: Universidade de Brasília. Laboratório de Psicologia do Trabalho, 1999.

UNESCO. **Qualidade das escolas públicas do ensino fundamental no Brasil**. Brasília: UNESCO, 2019.

VIEIRA, Lúvia F.; OLIVEIRA, Tiago G. As condições do trabalho docente na educação infantil no Brasil: alguns resultados de pesquisa (2002-2012). **Revista Educação em Questão**, Natal, v. 46, n. 32, p. 131-154, maio/ago., 2013.

Recebido em: 17 de dezembro de 2019

Aceito em: 18 de junho de 2020

Publicado em: 30 de junho de 2020

OBRIGATORIEDADE DA PRÉ-ESCOLA EM PERSPECTIVA: AVALIAÇÃO DO ATENDIMENTO NO ESTADO DO RIO DE JANEIRO

Bruno Tovar Falciano¹

Maria Fernanda Rezende Nunes²

¹ Doutorando e mestre em Educação pela Universidade Federal do Rio de Janeiro (Unirio), Rio de Janeiro, RJ, Brasil. E-mail: bruno@tovar.com.br

² Professora Titular-Livre em Educação Infantil da Universidade Federal do Estado do Rio de Janeiro (Unirio), Rio de Janeiro, RJ, Brasil. E-mail: nunes.mariafernandarezende@gmail.com

Resumo

O artigo tem como objetivo avaliar os efeitos decorrentes da obrigatoriedade escolar na Educação Infantil a partir da EC nº 59/2009 (BRASIL, 2009a). Foram adotadas como unidade de estudo as redes municipais de ensino do estado do Rio de Janeiro, buscando observar se a população de 4 e 5 anos de idade foi atendida e como ocorreu este atendimento. Para tanto, como escolha metodológica, optou-se pela construção de uma amostra ancorada na seguinte indagação: municípios com população e PIB semelhantes têm o mesmo grau de expansão do atendimento? A partir de uma amostra de 16 municípios, foi possível conhecer as características do atendimento à pré-escola e suas formas de expansão – ensino público, privado ou conveniado, infraestrutura e, por fim, os impactos da expansão no atendimento à creche. As análises foram elaboradas a partir de dados quantitativos, como os microdados do Censo da Educação Básica do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE).

Palavras-chave: Educação Infantil; obrigatoriedade; avaliação.

Abstract

The article aims to evaluate the effects of compulsory schooling on Early Childhood Education after the constitutional amendment nº 59/2009 (BRASIL, 2009a). The research investigate the municipal education system in the state of Rio de Janeiro and observe if the children at 4 and 5 years old were at school system and how this service was provide. As a methodological choice, the sample were build with the goal to address the following question: do municipalities with similar population and GDP have the same degree of expansion of service? From a sample of 16 municipalities, the characteristics of the preschool service and its forms of expansion - public, private, partnership, infrastructure are described and, finally, the impacts of the expansion in the daycare service are examined. The analyzes were based on quantitative data extracted from the microdata of Basic Education Census of Ministry of Education and the Demographic Census of the Brazilian Institute of Geography and Statistics (IBGE).

Keywords: Early Childhood Education, compulsory education; evaluation

INTRODUÇÃO

Este artigo é parte dos resultados da pesquisa *Expansão da escolaridade obrigatória e políticas municipais de Educação Infantil no estado do Rio de Janeiro*, desenvolvida pelo grupo Educação Infantil e Políticas Públicas/EIPP, que teve como objetivo acompanhar as disposições do governo federal quanto à expansão da escolaridade obrigatória para as crianças de quatro e cinco anos e também levantar o comportamento das matrículas da creche e pré-escola em cada município. Desse modo, foi possível mapear os tipos de expansão do atendimento obrigatório para a população de 4 e 5 anos e o impacto desta medida nas creches municipais.

Trabalhamos com a concepção de que o gestor público está permanentemente imprimindo uma marca na Educação Infantil do seu município, seja por meio da sua capacidade de formulação, argumentação e conhecimento dos significados da Educação Infantil, seja pelo lugar que ocupa no cenário municipal, na construção da política. Além das ideias que orientam o trabalho da gestão, há interesses, disputas, valores e concepções em jogo. Por isso, a capacidade de diálogo, de escuta, de saber lidar com as divergências e com os imprevistos, tão corriqueiros no cenário estadual, torna-se mais do que fundamental. Percebe-se que a política vai se construindo na expressão de vários atores que formulam narrativas, seja pelos discursos, documentos oficiais, deliberações, propostas pedagógicas, planejamentos ou pela prática com as crianças. É sabido que a baixa capacidade estatal e o elevado grau de compromissos burocráticos da administração municipal contribuem, sem dúvida, para o ingresso do setor privado nos processos relativos à oferta da educação pública (TRIPODI, Z. F; SOUSA, 2018). Tem sido esse o caso do estado do Rio de Janeiro?

Os debates educacionais se fizeram presentes na formulação de indagações de pesquisa e, mais especificamente, nos dados a serem apresentados neste artigo. Embora a universalização da pré-escola tenha sido formulada pela EC nº 59/2009, que torna obrigatória a escolaridade de crianças de 4 a 17 anos, foi na Conferência Nacional de Educação em 2010 e na tramitação do último Plano Nacional de Educação para o decênio 2014-2024 (PNE), no Congresso Nacional, que o debate ganhou novos desdobramentos para a política educacional. Partindo do pressuposto de que política pública é “[...] tudo que um governo faz e deixa de fazer, com todos os impactos de suas ações e de suas omissões” (AZEVEDO, 2003, p. 38), a pressão social, além de contribuir para que a obrigatoriedade entrasse na agenda pública, trouxe outros desafios para o atendimento em creche, que, no Rio de Janeiro, vinha sendo ofertado para um contingente pequeno de crianças, e em tempo integral. Estamos remetendo à aprovação da Meta 1 do PNE – “universalizar, até 2016, a

Educação Infantil na pré-escola para as crianças de 4 e 5 anos de idade e ampliar a oferta de Educação Infantil em creches de forma a atender, no mínimo, 50% das crianças de até 3 anos até o final da vigência deste PNE” (BRASIL, 2014).

Além dos problemas quantitativos decorrentes do cumprimento da meta 1 do plano, em artigo intitulado “Infância e crianças de 6 anos: desafios das transições na educação infantil e no ensino fundamental”, Kramer, Nunes e Corsino (2011) alertavam sobre o quanto a obrigatoriedade da pré-escola poderia “[...] levar à compreensão de que é aí que a educação básica começa, reiterando a antiga cisão entre creches e pré-escolas e trazendo de volta à cena o ideário de preparação para o ensino fundamental” (p. 75). Muitas outras preocupações puderam ser adensadas à consecução dessa política, como a de Campos (2013), por exemplo, em artigo que faz menção à revisão das pesquisas sobre o impacto da qualidade da Educação Infantil no desenvolvimento de crianças pequenas, “[...] que apontam para a diminuição da diferença entre alunos com distintas origens sociais e culturais, após a frequência a programas pré-escolares desenvolvidos em diversos países” (CAMPOS, 2013, p. 39). Ao abordar o tema da obrigatoriedade, também não é possível deixar de lado os aspectos relativos à infraestrutura dos estabelecimentos escolares, “[...] já que temos claro que o espaço físico também educa e que ter acesso a uma infraestrutura de qualidade constitui um direito de todas as crianças” (FALCIANO, SANTOS, NUNES, 2016, p. 899).

Este artigo tem como unidade de estudo as redes municipais de ensino do estado do Rio de Janeiro e busca observar se a população de 4 e 5 anos de idade foi atendida e como ocorreu este atendimento. O primeiro item aborda a construção da amostra: municípios com população e PIB semelhantes têm o mesmo grau de expansão do atendimento? O segundo item busca conhecer as características do atendimento à pré-escola e as formas de expansão – ensino público, privado ou conveniado, infraestrutura e, por fim, os impactos da expansão no atendimento à creche.

A CONSTRUÇÃO DA AMOSTRA

A proposta metodológica foi construída a partir de uma indagação: municípios com condições iniciais semelhantes obtêm resultados de expansão da pré-escola também semelhantes? Qualificar o que são condições iniciais semelhantes foi a etapa principal, e também estabelecer um marco inicial de análise.

Em relação às condições iniciais, três categorias foram utilizadas para qualificar algumas das características dos entes federados: a primeira delas se refe-

re ao produto interno bruto *per capita* (PIB *per capita*); a segunda, ao tamanho da população; e a terceira, ao nível de atendimento da pré-escola. O PIB *per capita* foi escolhido como uma forma de expressar a disposição que cada município teria de mobilizar recursos para a educação. A escolha do tamanho da população ajuda a ter um parâmetro da complexidade da gestão municipal, em função do tamanho da sua população. Por fim, o nível de atendimento da pré-escola em relação à população de 4 e 5 anos fornece informações sobre o maior ou menor esforço empreendido, tendo em vista a universalização. Os dados do PIB e o tamanho da população foram obtidos a partir dos dados divulgados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), sendo o PIB *per capita* calculado por meio da divisão do PIB total pela população total do município. Os dados de matrícula foram obtidos pelo sistema InepData do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

O ano de 2010 foi estabelecido como o marco inicial por dois motivos de dimensões diferentes, mas que se conjugam: (i) foi o ano subsequente ao da promulgação da EC nº 59/2009, que estabeleceu a obrigatoriedade de atendimento escolar para a população na faixa etária dos 4 aos 17 anos de idade e (ii) é o ano do último Censo Demográfico do IBGE – fonte oficial e detalhada de informações sobre a população. Ressalte-se que a proporção de atendimento no ano de 2010 foi construída a partir do número de matrículas na pré-escola daquele ano dividido pela população na faixa etária de 4 e 5 anos de idade daquele mesmo ano. Entretanto, vale esclarecer que, para monitorar a obrigatoriedade, isto é, verificar se todas as crianças de 4 e 5 anos de idade estariam efetivamente na escola, precisaríamos conhecer a população de 4 e 5 anos e comparar esta informação com o número de crianças de 4 e 5 anos que estivessem efetivamente na escola. Como não há dados da população para esta faixa etária com abrangência municipal posterior ao ano de 2010, bem como os relatórios do Censo da Educação Básica do Inep (divulgados em sua página eletrônica) não trazem o número de matrículas por idade, optamos por trabalhar com a proporção de atendimento seguindo o conceito de “taxa bruta de atendimento”, utilizado pelo Inep para expressar a relação entre as matrículas de uma determinada etapa escolar e a população com idade compatível com esta mesma etapa¹.

Utilizando os critérios adotados, categorizamos os municípios em condições iniciais semelhantes: oito grupos construídos a partir de quatro faixas populacionais (até 20.000, de 20.001 a 50.000, de 50.001 a 200.000 e acima de 200.000 habitantes) e duas faixas de PIB *per capita* (até R\$12.999 e acima de R\$12.999 por

¹A utilização da taxa bruta pode fazer com que municípios apresentem atendimento acima de 100%. Isso pode ocorrer, pois há municípios que atendem crianças de outras localidades limítrofes, bem como há crianças na pré-escola com idade diferente de 4 e 5 anos.

habitante). As quatro faixas populacionais utilizadas foram semelhantes aos cortes utilizados pelo IBGE, porém, aglutinadas de forma a distribuir proporcionalmente o número de municípios. Já a faixa de PIB *per capita* foi dividida pela mediana do valor do PIB *per capita* dos municípios do estado. Com a tabela construída e organizada nessas oito categorias, classificamos então os municípios de acordo com o seu nível de atendimento. Em seguida, analisamos os seus respectivos resultados nos anos subsequentes, medidos pela variação das matrículas da pré-escola entre os anos de 2010 e 2014². Como resultado, encontramos municípios com as mesmas condições iniciais que apresentavam resultados significativamente diferentes, respondendo, assim, uma das indagações da pesquisa.

Diante desse achado, procuramos, então, conhecer os possíveis caminhos da obrigatoriedade nos municípios desse estado. Para tanto, elegemos uma amostra com dois municípios em cada faixa de PIB *per capita* que tivessem populações equivalentes e estivessem situados em estágios próximos do atendimento à meta de universalização da pré-escola em 2010, mas que um deles tivesse obtido uma maior e o outro uma menor diferença do número de matrículas na pré-escola entre os anos de 2010 e 2014. Dessa forma, a amostra obtida foi composta por 16 municípios apresentados no quadro abaixo.

Quadro 01 – Municípios da amostra de acordo com as suas categorias de classificação.

Categorias			Descritivos em 2010				
PIB per capita	População	Varição	Município	PIB per capita	População	Atendimento Pré-escola	Var. Matrícula vs. 2014
até R\$ 12.999	até 20.000	menor	Cambuci	10.916	14.827	115%	-33%
		maior	São Sebastião do Alto	11.168	8.895	74%	26%
	de 20.001 a 50.000	menor	Vassouras	11.851	34.410	102%	-16%
		maior	Iguaba Grande	10.463	22.851	101%	10%
	de 50.001 a 200.000	menor	Saquarema	12.781	74.234	88%	3%
		maior	Guapimirim	9.563	51.483	52%	62%
acima de 200.000	menor	Belford Roxo	9.568	469.332	48%	19%	
	maior	Nova Iguaçu	11.985	796.257	44%	55%	
acima de R\$ 12.999	até 20.000	menor	Carmo	16.204	17.434	95%	-7%
		maior	Quatis	14.581	12.793	94%	52%
	de 20.001 a 50.000	menor	Santo Antônio de Pádua	13.352	40.589	106%	-11%
		maior	Paraíba do Sul	13.051	41.084	96%	11%
	de 50.001 a 200.000	menor	Barra do Pirai	14.603	94.778	88%	-10%
		maior	Teresópolis	17.005	163.746	73%	12%
acima de 200.000	menor	Niterói	23.092	487.562	115%	0%	
	maior	Duque de Caxias	30.924	855.048	44%	47%	

Fonte: Tabulação própria a partir do Censo Demográfico do IBGE e do Censo da Educação Básica do Inep.

A pergunta a seguir é decorrente da primeira: como municípios que estão próximos da universalização atingiram esse patamar de atendimento?

² Para a composição da amostra, utilizou-se a variação até 2014, pois esse era o último ano disponível do Censo da Educação Básica do Inep, no momento do desenho inicial da pesquisa. Entretanto, como a pesquisa encerrou em 2018, incorporamos nas análises os últimos dados divulgados, que vão até 2017.

OS CAMINHOS DA UNIVERSALIZAÇÃO DA PRÉ-ESCOLA

Inicialmente, foi necessária a construção de um indicador de tendência que fosse capaz de relatar a capacidade de atendimento dos municípios em relação à população de 4 e 5 anos. Esse indicador foi construído a partir da comparação das variações do número de matrículas da pré-escola e do número de nascidos vivos por domicílio materno, divulgado anualmente pelo Departamento de Informática do Sistema Único de Saúde (Datasus). O conceito proposto por esse indicador de tendência é de que, se um município reduzir o seu número de nascidos vivos, ele reduzirá potencialmente a sua população de 4 e 5 anos de idade, e se esse mesmo município apresentar um aumento no número de matrículas, ele ampliará também potencialmente o seu atendimento na pré-escola. De outra forma, se um município aumentar a sua população de 4 e 5 anos de idade ao mesmo tempo em que reduzir o número de matrículas na pré-escola, ele reduzirá potencialmente o seu atendimento³.

O quadro a seguir apresenta os resultados da aplicação dessa metodologia nos municípios da amostra.

Quadro 02 - Aplicação do indicador de tendência do atendimento à pré-escola nos municípios da amostra.

Município	Categorias			Nascidos vivo			Matrículas Pré-escola em 2010			Indicador de tendência
	PIB per capita	População	Var. Matrículas	2010	2016*	2010 vs 2016	2010	2017	2010 vs 2017	
Cambuci	baixo	pequeno	menor	142	134	-6%	413	279	-32%	redução
São Sebastião do Alto	baixo	pequeno	maior	116	101	-13%	185	191	3%	aumento
Vassouras	baixo	médio-pequeno	menor	500	504	1%	982	836	-15%	redução
Iguaba Grande	baixo	médio-pequeno	maior	247	342	38%	594	741	25%	redução
Saquarema	baixo	médio-grande	menor	843	1.258	49%	1.882	2.141	14%	redução
Guapimirim	baixo	médio-grande	maior	672	725	8%	841	1.310	56%	aumento
Belford Roxo	baixo	grande	menor	7.129	7.215	1%	6.741	8.214	22%	aumento
Nova Iguaçu	baixo	grande	maior	11.597	11.842	2%	10.002	15.020	50%	aumento
Carmo	alto	pequeno	menor	209	207	-1%	466	449	-4%	redução
Quatis	alto	pequeno	maior	169	139	-18%	341	315	-8%	aumento
Santo Antônio de Pádua	alto	médio-pequeno	menor	465	461	-1%	1.092	932	-15%	redução
Paraíba do Sul	alto	médio-pequeno	maior	531	541	2%	1.088	1.135	4%	aumento
Barra do Pirai	alto	médio-grande	menor	1.026	1.130	10%	2.223	2.106	-5%	redução
Teresópolis	alto	médio-grande	maior	2.200	2.198	0%	3.241	4.175	29%	aumento
Niterói	alto	grande	menor	5.966	5.903	-1%	11.398	10.415	-9%	redução
Duque de Caxias	alto	grande	maior	12.511	13.583	9%	10.849	17.573	62%	aumento

* Esta variável só estava disponível até do ano de 2016.

Fonte: Tabulação própria com base nos dados do Datasus e do Censo da Educação Básica do Inep.

³ É importante ratificar que este indicador é uma aproximação para se exprimir o atendimento de maneira indireta e não uma projeção do índice de atendimento, em que outros fatores como variação na taxa de ocupação, município de residência das crianças e defasagem idade/série deveriam ser considerados.

No quadro acima, é possível verificar que 9 dos 16 municípios da amostra apresentaram um crescimento do número de nascidos vivos e, consequentemente, da população infantil, apesar de macros movimentos demográficos da população brasileira, demonstrando que a natalidade e o tamanho das famílias vêm reduzindo. Este fato alerta para que o planejamento de vagas para a Educação Infantil seja realizado com uma perspectiva territorial bem específica. Pelo indicador de tendência, que expressa o efeito combinado da variação do número de nascidos vivos com o número de matrículas, verificaremos ainda que metade dos municípios apresentou uma tendência de redução e a outra metade, uma tendência de aumento do atendimento à pré-escola.

Ao cruzar os dados do Quadro 02 com os dados do Quadro 01, é interessante notar que, em geral, os municípios mantiveram a sua tendência, seja ela de aumento ou redução do número de matrículas, tanto no período de 2010 a 2014 quanto no período mais longo, de 2010 a 2017. Em outras palavras, há indícios, nessa amostra, de que a obrigatoriedade de matrícula para a população a partir de 4 anos de idade proposta pela nova legislação, por si só, não alterou o movimento de expansão ou redução de matrículas da pré-escola nesses municípios. Essa hipótese pode ser corroborada pelo Relatório do 2º ciclo de monitoramento do Inep, que aponta essa unidade da federação com o menor progresso na cobertura da pré-escola, em relação aos demais estados da federação (p. 26-27) até o ano de 2016. Nesse mesmo relatório, fica claro ainda que, a despeito até mesmo dos municípios que apresentaram potencial de expansão no indicador de tendência, os patamares de atendimento da pré-escola nesse estado ainda estão muito aquém da universalização.

VARIÁVEIS DE QUALIDADE DA PRÉ-ESCOLA

A partir de algumas variáveis obtidas nos microdados do Censo da Educação Básica do Inep, foi possível compreender as características desse atendimento. O quadro abaixo apresenta a relação entre o número de matrículas da rede pública e da rede privada.

Quadro 03 - Número absoluto e relativo de matrículas da pré-escola nas redes pública e privada dos municípios da amostra nos anos de 2010 e 2017.

Município	Categorias			Matrículas Pública		Matrículas Privada		Proporção Pública	
	PIB per capita	População	Var. Matrículas	2010	2017	2010	2017	2010	2017
Cambuci	baixo	pequeno	menor	392	207	21	72	95%	74%
São Sebastião do Alto	baixo	pequeno	maior	185	191	0	0	100%	100%
Vassouras	baixo	médio-pequeno	menor	730	617	252	219	74%	74%
Iguaba Grande	baixo	médio-pequeno	maior	426	556	168	185	72%	75%
Saquarema	baixo	médio-grande	menor	1.628	1.798	254	343	87%	84%
Guapimirim	baixo	médio-grande	maior	678	1.015	163	295	81%	77%
Belford Roxo	baixo	grande	menor	2.737	4.169	4.004	4.045	41%	51%
Nova Iguaçu	baixo	grande	maior	4.621	6.881	5.381	8.139	46%	46%
Carmo	alto	pequeno	menor	384	333	82	116	82%	74%
Quatis	alto	pequeno	maior	316	287	25	28	93%	91%
Santo Antônio de Pádua	alto	médio-pequeno	menor	934	772	158	160	86%	83%
Paraíba do Sul	alto	médio-pequeno	maior	937	968	151	167	86%	85%
Barra do Pirai	alto	médio-grande	menor	1.827	1.695	396	411	82%	80%
Teresópolis	alto	médio-grande	maior	2.080	3.050	1.161	1.125	64%	73%
Niterói	alto	grande	menor	3.707	4.324	7.691	6.091	33%	42%
Duque de Caxias	alto	grande	maior	5.572	8.908	5.277	8.665	51%	51%

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

A proporção de matrículas atendidas pela rede pública oscila entre 33% a 100%, e é possível perceber sua importante presença no caso dos municípios com uma população de pequeno a médio porte (até 200 mil habitantes) – em torno de 80% do atendimento. Nos municípios com uma população grande (acima de 200 mil habitantes), o atendimento público representa em torno da metade do número de matrículas. A partir desse achado, várias indagações se fazem pertinentes: os municípios menores apresentam baixa atratividade para o setor privado? Os municípios maiores apresentam dificuldades de atendimento para a rede pública? Caso afirmativo, quais dificuldades? Sabe-se que o número de habitantes de um município não é diretamente proporcional à sua extensão territorial e, por isso, outros esforços de atendimento se fazem presentes. Certamente, o comparecimento ou não da rede privada nos ajuda nessa análise.

Ainda em relação ao atendimento da rede pública, podemos afirmar que, em 13 dos 16 municípios pesquisados, a rede pública manteve ou aumentou a sua participação. Analisando especificamente o número absoluto de matrículas, constatamos que a rede pública induziu o movimento total da rede de pré-escolas em 14 dos 16 municípios, isto é, quando a rede pública aumentou o seu número de matrículas, a rede de pré-escolas como um todo também aumentou o seu número de matrículas. De outra maneira, quando houve redução no número de matrículas da rede pública, a rede de pré-escolas como um todo reduziu o seu número de matrículas. Com apenas esses dados, não temos

subsídios para propor qualquer efeito de causalidade, mas parece-nos que há uma relação marcada pela indução da rede pública, que representa em geral 3/4 do número total de matrículas da pré-escola nos municípios da amostra.

O quadro a seguir lança foco sobre a rede de estabelecimentos conveniados com o poder público, que, de acordo com o Censo da Educação Básica do Inep, são apresentados como uma subcategoria da rede privada. Pretendemos, assim, averiguar se, nos municípios da amostra, essa seria uma forma de se chegar à universalização.

Quadro 04 - Número absoluto e relativo de matrículas da pré-escola em estabelecimentos conveniados da rede privada em 2010 e 2017.

Município	Categorias			Nº matrículas		Proporção	
	PIB per capita	População	Var. Matrículas	2010	2017	2010	2017
Cambuci	baixo	pequeno	menor	21	0	100%	0%
São Sebastião do Alto	baixo	pequeno	maior	0	0	0%	0%
Vassouras	baixo	médio-pequeno	menor	41	0	16%	0%
Iguaba Grande	baixo	médio-pequeno	maior	42	0	25%	0%
Saquarema	baixo	médio-grande	menor	63	0	25%	0%
Guapimirim	baixo	médio-grande	maior	0	0	0%	0%
Belford Roxo	baixo	grande	menor	2.159	602	54%	15%
Nova Iguaçu	baixo	grande	maior	1.001	448	19%	6%
Carmo	alto	pequeno	menor	23	0	28%	0%
Quatis	alto	pequeno	maior	0	0	0%	0%
Santo Antônio de Pádua	alto	médio-pequeno	menor	25	30	16%	19%
Paraíba do Sul	alto	médio-pequeno	maior	47	0	31%	0%
Barra do Pirai	alto	médio-grande	menor	0	7	0%	2%
Teresópolis	alto	médio-grande	maior	331	107	29%	10%
Niterói	alto	grande	menor	3.125	1.795	41%	29%
Duque de Caxias	alto	grande	maior	1.372	666	26%	8%

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

Observa-se que os convênios diminuíram em 11 municípios e em outros 3 municípios simplesmente não foram realizados – nem no ano de 2010, nem no ano de 2017. Até mesmo onde os convênios se fizeram mais presentes na amostra, a representatividade deles não passou de 1/3 do total de matrículas da rede privada no ano de 2017. Esses dados fazem crer que as estratégias de atendimento para a pré-escola no Rio de Janeiro não passaram pelo uso de convênios.

Outra estratégia de expansão de que se tem amplo conhecimento na literatura (BRASIL 2006, 2009b; CAMPOS, 2013) diz respeito à utilização de prédios adaptados – os famosos “puxadinhos” –, que garantem a ampliação de vagas, sem que sejam observadas, contudo, as garantias mínimas de qualidade do atendimento.

Quadro 05 - Número absoluto e relativo de matrículas em prédio próprio da rede total de pré-escola nos anos de 2011* e 2017.

Município	Categorias			Nº matrículas		Proporção	
	PIB per capita	População	Var. Matrículas	2011*	2017	2011*	2017
Cambuci	baixo	pequeno	menor	113	207	31%	74%
São Sebastião do Alto	baixo	pequeno	maior	240	170	84%	89%
Vassouras	baixo	médio-pequeno	menor	423	690	45%	83%
Iguaba Grande	baixo	médio-pequeno	maior	574	741	91%	100%
Saquarema	baixo	médio-grande	menor	1.540	2.027	86%	95%
Guapimirim	baixo	médio-grande	maior	673	981	75%	75%
Belford Roxo	baixo	grande	menor	4.166	5.422	62%	66%
Nova Iguaçu	baixo	grande	maior	7.709	11.678	73%	78%
Carmo	alto	pequeno	menor	220	261	52%	58%
Quatis	alto	pequeno	maior	379	315	92%	100%
Santo Antônio de Pádua	alto	médio-pequeno	menor	853	649	88%	70%
Paraíba do Sul	alto	médio-pequeno	maior	454	667	38%	59%
Barra do Pirai	alto	médio-grande	menor	1.685	1.619	78%	77%
Teresópolis	alto	médio-grande	maior	1.471	2.327	49%	56%
Niterói	alto	grande	menor	7.884	7.447	72%	72%
Duque de Caxias	alto	grande	maior	7.562	11.831	61%	67%

* Esta variável só está disponível a partir do ano de 2011.

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

Identificamos que a oferta de matrículas da pré-escola na rede como um todo, seja pública ou privada, vem crescendo em prédios próprios. Foram 12 municípios da amostra que demonstraram um aumento na proporção de matrículas, entre 2011 e 2017, nesse tipo de estabelecimento. Entretanto, ainda em 2017, um a cada três estabelecimentos da rede de pré-escolas não era próprio. Estamos, em certa medida, dizendo que, quando um prédio é próprio, em geral, há mais investimentos de médio e longo prazo, o que possibilita uma maior adequação das instalações à proposta pedagógica e às particularidades inerentes à Educação Infantil. Conforme levantam Falciano, Santos e Nunes (2016), “[...] a infraestrutura dos estabelecimentos escolares é um dos fatores preponderantes na busca da qualidade na educação infantil, já que temos claro que o espaço físico também educa e que ter acesso a uma infraestrutura de qualidade constitui um direito de todas as crianças” (p. 890).

Assim, achamos oportuno conhecer, nos municípios da amostra, os estabelecimentos com banheiro adaptado e parque infantil.

Quadro 6 - Número absoluto e relativo de matrículas da pré-escola em estabelecimentos com banheiro adaptado para Educação Infantil em 2010 e 2017

Município	Categorias			Matrículas		Proporção	
	PIB per capita	População	Var. Matrículas	2010	2017	2010	2017
Cambuci	baixo	pequeno	menor	210	188	51%	67%
São Sebastião do Alto	baixo	pequeno	maior	107	113	58%	59%
Vassouras	baixo	médio-pequeno	menor	280	525	29%	63%
Iguaba Grande	baixo	médio-pequeno	maior	168	555	28%	75%
Saquarema	baixo	médio-grande	menor	845	1.460	45%	68%
Guapimirim	baixo	médio-grande	maior	189	797	22%	61%
Belford Roxo	baixo	grande	menor	4.233	6.713	63%	82%
Nova Iguaçu	baixo	grande	maior	4.709	11.291	47%	75%
Carmo	alto	pequeno	menor	142	215	30%	48%
Quatis	alto	pequeno	maior	197	206	58%	65%
Santo Antônio de Pádua	alto	médio-pequeno	menor	454	559	42%	60%
Paraíba do Sul	alto	médio-pequeno	maior	517	807	48%	71%
Barra do Pirai	alto	médio-grande	menor	1.136	1.512	51%	72%
Teresópolis	alto	médio-grande	maior	1.693	2.137	52%	51%
Niterói	alto	grande	menor	8.963	10.085	79%	97%
Duque de Caxias	alto	grande	maior	6.247	13.243	58%	75%

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

Em média, nos municípios da amostra, tínhamos 57% das matrículas em 2010 em estabelecimentos com banheiro adaptado. Em 2017, esse número aumentou para 77% das matrículas. Houve um aumento real em 14 dos 16 municípios, o que é bastante promissor; entretanto, chama a atenção o fato de que, em 2017, ainda há municípios onde metade das matrículas é atendida em estabelecimentos sem banheiro adaptado. Que comentários poderiam ser feitos? Em busca da universalização, os municípios estão alocando turmas de pré-escola em escolas voltadas para outras etapas da educação básica? Qual a concepção de criança e a identidade desse atendimento? Interessante notar que os espaços também produzem narrativas, e a que está sendo formulada, para os menos avisados, é a de que a criança tem que se adaptar à vida “dos grandes”?

O próximo quadro nos ajuda a pensar sobre as concepções presentes nesses espaços.

Quadro 7 - Número absoluto e relativo de matrículas da pré-escola em estabelecimentos com parque infantil em 2010 e 2017.

Município	Categorias			Matrículas		Proporção	
	PIB per capita	População	Var. Matrículas	2010	2017	2010	2017
Cambuci	baixo	pequeno	menor	242	160	59%	57%
São Sebastião do Alto	baixo	pequeno	maior	107	113	58%	59%
Vassouras	baixo	médio-pequeno	menor	294	295	30%	35%
Iguaba Grande	baixo	médio-pequeno	maior	443	741	75%	100%
Saquarema	baixo	médio-grande	menor	1.055	1.533	56%	72%
Guapimirim	baixo	médio-grande	maior	223	650	27%	50%
Belford Roxo	baixo	grande	menor	4.266	4.250	63%	52%
Nova Iguaçu	baixo	grande	maior	5.294	9.307	53%	62%
Carmo	alto	pequeno	menor	252	315	54%	70%
Quatis	alto	pequeno	maior	190	32	56%	10%
Santo Antônio de Pádua	alto	médio-pequeno	menor	656	479	60%	51%
Paraíba do Sul	alto	médio-pequeno	maior	811	622	75%	55%
Barra do Pirai	alto	médio-grande	menor	1.245	989	56%	47%
Teresópolis	alto	médio-grande	maior	2.234	2.312	69%	55%
Niterói	alto	grande	menor	9.921	9.530	87%	92%
Duque de Caxias	alto	grande	maior	5.899	11.137	54%	63%

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

Apesar de 10 municípios da amostra terem aumentado a disponibilidade de parque infantil, outros 6 reduziram a oferta de parque infantil. O saldo disso foi que, na média desses municípios da amostra, quase não houve mudanças na oferta de parque infantil, uma vez que em 2010 representavam 63% das matrículas e 65% em 2017. Interessante notar que não há indicativos, nesse grupo de municípios, de que o tamanho da população contribua para a disponibilidade de espaço para os parques infantis. Niterói, por exemplo, que é uma cidade urbana com mais de duzentos mil habitantes, possui 90% do atendimento com parques infantis.

Em artigo sobre os resultados do Programa Proinfância – programa do governo federal que tem como meta a garantia do acesso à creche e à pré-escola, com a construção de novos estabelecimentos, a partir de seus projetos-padrão –, Falciano, Nunes e Santos (2019) apontam o papel indutor do estado na melhoria da infraestrutura: “[...] houve ganhos significativos para as redes como um todo nas adequações dos estabelecimentos às necessidades da Educação Infantil” (p. 276). Neste estudo, tínhamos como hipótese a ideia de que a obrigatoriedade de universalização da pré-escola pudesse concentrar todos os esforços de expansão da rede, mas vimos que a creche foi a maior beneficiada pelo programa, garantindo, assim, a sua expansão e o aprimoramento da qualidade.

Assim, neste recorte, construímos a mesma pergunta com a intenção de explicitar os dados referentes à creche nos municípios da amostra, ou seja, diante da necessidade de universalização da pré-escola, em que medida são feitos esforços municipais de expansão da creche?

Quadro 8 - Número absoluto e relativo de matrículas em creche em relação ao total de matrículas da Educação Infantil e crescimento do número de matrículas de creche em 2010 e 2017.

Município	Categorias			Matrículas em creche		Proporção da creche		Matrículas 2017 vs 2010
	PIB per capita	População	Var. Matrículas	2010	2017	2010	2017	
Cambuci	baixo	pequeno	menor	169	276	29%	50%	63%
São Sebastião do Alto	baixo	pequeno	maior	219	229	54%	55%	5%
Vassouras	baixo	médio-pequeno	menor	412	595	30%	42%	44%
Iguaba Grande	baixo	médio-pequeno	maior	347	684	37%	48%	97%
Saquarema	baixo	médio-grande	menor	689	1.668	27%	44%	142%
Guapimirim	baixo	médio-grande	maior	922	785	52%	37%	-15%
Belford Roxo	baixo	grande	menor	3.425	2.187	34%	21%	-36%
Nova Iguaçu	baixo	grande	maior	1.673	2.547	14%	14%	52%
Carmo	alto	pequeno	menor	359	390	44%	46%	9%
Quatis	alto	pequeno	maior	54	364	14%	54%	574%
Santo Antônio de Pádua	alto	médio-pequeno	menor	578	788	35%	46%	36%
Paraíba do Sul	alto	médio-pequeno	maior	859	1.043	44%	48%	21%
Barra do Pirai	alto	médio-grande	menor	1.183	1.478	35%	41%	25%
Teresópolis	alto	médio-grande	maior	1.908	2.712	37%	39%	42%
Niterói	alto	grande	menor	3.489	6.276	23%	38%	80%
Duque de Caxias	alto	grande	maior	2.949	5.299	21%	23%	80%

Fonte: Tabulação própria a partir dos microdados do Censo da Educação Básica do Inep.

Percebe-se que a creche não sofreu efeitos negativos com a obrigatoriedade da pré-escola nos municípios da amostra. Suas matrículas cresceram em 14 dos 16 municípios estudados, sendo que, desses, 7 municípios registraram aumentos superiores a 50% no número de matrículas. E não foi apenas isso que verificamos:

- a Educação Infantil – creche e pré-escola – cresceu em quase todos os municípios, sendo impulsionada justamente pelo crescimento das matrículas em creche;
- as matrículas em creche aumentaram a sua representatividade em relação a pré-escola na Educação Infantil; e
- a creche cresceu até mesmo em 7 dos 16 municípios em que a pré-escola reduziu as suas matrículas.

NOTA FINAL

Ter uma agenda nacional acordada com o conjunto da sociedade civil é fundamental, pois ela induz demandas educacionais, decisões e ações dos atores envolvidos. Entretanto, quanto à universalização da pré-escola nos municípios da amostra objeto desta pesquisa, observa-se que há outras agendas em curso e que as demandas macro não garantem uma correspondência rigorosa às necessidades locais. Os movimentos demográficos da população infantil, por exemplo, são muito diversos e não necessariamente estão afinados com a tendência estadual ou nacional. Vimos que a população infantil cresce no período analisado – de 2010 a 2017 – em vários municípios do Rio de Janeiro.

Constata-se, também, uma desigualdade de origem determinada pelo município de nascimento das crianças, revelando que nem todas têm as mesmas condições de acesso à Educação Infantil. Muitos empreenderam esforços em direção à universalização, enquanto em outros municípios a oferta é muito incipiente, tendo em vista a meta 1 do PNE.

Há indícios de que apenas a obrigatoriedade de matrícula para a população a partir de 4 anos de idade, por si só, não alterou a trajetória de expansão ou redução de matrículas na pré-escola – vimos que os municípios mantiveram a tendência de aumento ou redução do número de matrículas, tanto no período de 2010 a 2014 quanto no de 2010 a 2017.

Outro achado foi a proporção da rede pública no atendimento à pré-escola. Essa rede se fez mais presente nos municípios com até 200 mil habitantes, ficando em torno de 80%, enquanto, nos municípios acima dessa faixa populacional, representou em torno de 50% do atendimento. Verificou-se, também, que a rede pública induziu o aumento ou a redução do número total de matrículas da rede de pré-escolas e que as estratégias de atendimento para a pré-escola contaram com prédios próprios e não foi feito uso de convênios.

Em relação ao impacto da obrigatoriedade da pré-escola na creche, a pesquisa mostrou que, nos municípios da amostra e no período estudado, a creche não sofreu efeitos negativos. Ocorreu exatamente o contrário, pois a Educação Infantil – creche e pré-escola – cresceu em quase todos os municípios, impulsionada pelo crescimento das matrículas em creche, que aumentaram a sua representatividade em relação à pré-escola no âmbito da Educação Infantil, tendo ainda a creche ampliado o seu número de matrículas até mesmo em municípios onde a pré-escola reduziu as suas matrículas.

REFERÊNCIAS

AZEVEDO, S. Políticas públicas: discutindo modelos e alguns problemas de implementação. In: SANTOS JUNIOR, O. A. et al. **Políticas públicas e gestão local**: programa interdisciplinar de capacitação de conselheiros municipais. Rio de Janeiro: Fase, 2003.

BRASIL. Congresso Nacional. **Lei nº. 13.005/2014**. Aprova o Plano Nacional de Educação. Brasília, 2014.

_____. Congresso Nacional. **Lei nº. 12.796/2013**. Altera a Lei nº. 9.394, de 20 de dezembro de 1996. Brasília, 2013.

_____. Congresso Nacional. **Emenda Constitucional nº. 59/2009**. Brasília, 2009a.

_____. MEC. **Indicadores da Qualidade na Educação Infantil**. Brasília: MEC/SEB, 2009b.

_____. MEC. **Parâmetros nacionais de qualidade para a educação infantil**. Brasília: MEC, SEB, 2006. v.1; il.

_____. Congresso Nacional. **Lei nº 10.172/2001**. Aprova o Plano Nacional de Educação e dá outras providências. Brasília, 2001.

_____. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Base de microdados do Censo Escolar da Educação Básica**. Disponível em: <http://portal.inep.gov.br/microdados>. Acesso em: 03 jun. 2018.

_____. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Sistema de consulta a informações e estatísticas educacionais (InepData)**. Disponível em: <http://portal.inep.gov.br/inep-data>. Acesso em: 24 abr. 2015.

_____. Ministério da Saúde. Datasus. **Nascidos Vivos 1994-2016 no Estado do Rio de Janeiro**. Disponível em: <http://datasus.saude.gov.br/informacoes-de-saude/tabnet/estatisticas-vitais>. Acesso em: 03 abr. 2015.

CAMPOS, M. M. Entre as políticas de qualidade e a qualidade das práticas. **Cadernos de Pesquisa**, São Paulo, v. 43, n. 148, p. 22-43, abr. 2013.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Censos demográficos**: características gerais da população. Disponível em: <https://sidra.ibge.gov>. Acesso em: 10 mar. 2015.

_____. **PIB Municipal 2007-2011**. Disponível em:

http://www.ibge.gov.br/home/estatistica/economia/pibmunicipios/2011/default_xls.shtm. Acesso em: 30 mar. 2015.

FALCIANO, B. T.; NUNES, M. F. R.; SANTOS, E. C. Dos. Dez anos do Proinfância: efeitos de uma política de indução. **Estudos em Avaliação Educacional**, São Paulo, SP, v. 30, n. 73, p. 254-278, 2019. ISSN: 0103-6831, DOI: <http://dx.doi.org/10.18222/eae.v30i73.5719>.

FALCIANO, B. T.; SANTOS, E. C. Dos; NUNES, M. F. R. Infraestrutura escolar: um critério de comparação da qualidade na Educação Infantil. **Estudos em Avaliação Educacional**, São Paulo, SP, v. 27, n. 66, p. 880-906, 2016. DOI: dx.doi.org/10.18222/eae.v0ix.4021.

KRAMER, S.; NUNES, M. F. R.; CORSINO, P. Infância e Crianças de 6 anos: desafios das transições na Educação Infantil e Ensino Fundamental. **Educação e Pesquisa** (USP. Impresso), v. 37, p. 69-85, 2011.

TRIPODI, Z. F.; SOUSA, S. Z. Do governo à governança: permeabilidade do estado a lógicas privatizantes na educação. **Cad. Pesqui.**, São Paulo, v. 48, n. 167, p. 228-253, mar. 2018.

Recebido em: 01 de dezembro de 2019

Aceito em: 20 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÕES EM LARGA ESCALA DE LÍNGUA PORTUGUESA: UMA PESQUISA SOBRE A COMPLEXIDADE DOS TEXTOS QUE DÃO SUPORTE A ITENS QUE AVALIAM A LEITURA

Begma Tavares Barbosa¹

Hilda Linhares Micarello²

Rosângela Veiga Ferreira³

¹ Professora titular aposentada da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: begmatb@gmail.com

² Professora associada da Faculdade de Educação da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: hilda.micarello@uab.ufjf.br

³ Professora do Colégio de Aplicação João XXIII, da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: rosangelaveiga.ferreira@ufjf.edu.br

Resumo

Este artigo discute o tema da complexidade dos textos que dão suporte a itens de avaliações em larga escala de Língua Portuguesa, com base em resultados de uma pesquisa no âmbito da qual foram analisados 763 textos utilizados em itens de avaliações em larga escala de Língua Portuguesa. A metodologia da pesquisa consistiu em análise dos itens sustentada pelas teorias da cognição, da linguística textual, e nos parâmetros estatísticos dos itens com base na Teoria da Resposta ao Item, modelo de três parâmetros. Com base na pesquisa realizada, foram definidos quatro níveis de complexidade dos textos. Os resultados apontam a relevância de critérios claros para definir a complexidade dos textos na elaboração de novos itens e para uma descrição mais precisa de padrões de desempenho com base na proficiência em leitura demonstrada pelos estudantes avaliados.

Palavras-chave: Avaliação da leitura. Avaliação em larga escala de Língua Portuguesa. Complexidade dos Textos.

Abstract

This article discusses the complexity of texts that support large scale Portuguese Language assessment items, based on the results of a research in which 763 texts used in large scale Portuguese Language assessment items were analyzed. The research methodology consisted of item analysis supported by theories of cognition, textual linguistics, and the statistical parameters of the items based on the Item Response Theory, a three parameter model. Based on the research, four levels of text complexity were defined. The results point to the relevance of clear criteria for defining the complexity of texts in the elaboration of new items and for a more accurate description of performance standards based on the reading proficiency demonstrated by the students evaluated.

Key words: Reading assessment. Large-scale Portuguese Language Assessment. Text Complexity.

INTRODUÇÃO

A pesquisa, cujos resultados são apresentados neste artigo, foi apoiada e financiada pela Fundação CAEd, da Universidade Federal de Juiz de Fora. O objetivo da pesquisa foi produzir conhecimentos sobre a validade e confiabilidade de itens aplicados em avaliações sistêmicas da leitura aplicadas em municípios e estados da federação e em avaliações nacionais. O recorte da pesquisa tratado no presente artigo se refere ao tema da complexidade dos textos que dão suporte a esses itens.

A avaliação da leitura é um tema desafiador, por ser este um construto⁴ complexo e multidimensional. Kleiman (2019, s/p) afirma que

na compreensão leitora, estão envolvidos: um texto – objeto linguístico e cultural portador de um significado -; um leitor - com saberes, experiências, capacidades e habilidades; e uma situação comunicativa de interação entre leitor e autor via texto escrito, que determina em grande parte o *que* e *como* se compreende.

No caso específico da presente pesquisa, o foco se coloca no texto como objeto linguístico portador de um significado e nos desafios que ele coloca ao leitor para que este seja capaz de compreender o que lê.

O tema da complexidade dos textos que são utilizados como suporte de itens que avaliam a leitura em avaliações de larga escala de Língua Portuguesa vem se apresentando como relevante, na medida em que a consideração dessa complexidade é fundamental para a compreensão do nível de dificuldade das tarefas de leitura propostas pelos itens aplicados com o intuito de aferir a proficiência em leitura dos estudantes, nas diversas etapas da educação básica. Além disso, critérios mais claros para a definição da complexidade dos textos contribuem para maior confiabilidade dos instrumentos de avaliação e, conseqüentemente, para a validade dos testes. De acordo com Primi (2012, p.300), em psicometria o conceito de validade “relaciona-se à questão que investiga se o teste está medindo o construto que se propõe medir”. Nesse sentido, sedo a leitura uma atividade cognitiva multidimensional, uma maior clareza sobre a natureza da atividade de leitura que está sendo requerida pelos itens que compõem os testes terá como conseqüência um melhor ajuste do teste ao que se pretende avaliar em cada etapa da escolarização.

⁴ Em psicometria, o constructo é um atributo intangível, com manifestação variável entre indivíduos, que só pode ser avaliado indiretamente, quando expresso sob a forma de performances físicas, emocionais, atitudinais ou cognitivas. Disponível em <http://portal.inep.gov.br/documents/186968/484421/Relat%C3%B3rio+Saeb+2017/e683ba93-d9ac-4c2c-8f36-10493e99f9b7?version=1.0>. Acesso em: 10 dez. 2019.

Sousa e Hübner (2014) afirmam que a opção por um tipo de tarefa para avaliar a compreensão leitora requer a explicitação da concepção de leitura que fundamenta essa opção. No caso específico de que trata o presente artigo, a concepção de leitura que tem fundamentado não apenas os processos de avaliação desse construto, mas, principalmente, as práticas de ensino é a de que a leitura é uma construção subjetiva de sentidos com base nos significados do texto, ou seja, os leitores atuam sobre o texto a partir de suas vivências em uma determinada cultura. Diante de um texto, os leitores acionam conhecimentos mais ou menos formalizados, como aqueles sobre os textos e sobre a própria língua, adquiridos na escola ou fora dela. Desse modo, o texto não porta um sentido pré-existente à leitura, ou seja, o significado não está no texto, mas se constitui na interação entre o leitor e o texto. Este, por sua vez, oferece um conjunto de pistas que guiam o leitor na tarefa de construção de sentido que é a leitura.

Os Parâmetros Curriculares Nacionais de Língua Portuguesa (BRASIL, 1998, p.36) consagraram uma virada pragmática no ensino da língua materna ao definirem o texto como “unidade básica de ensino”. Essa definição implica em assumir que ensinar a língua é, principalmente, ensinar como identificar as pistas que o texto oferece e a se guiar por essas pistas para produzir sentido para o que se lê.

Mais recentemente, a Base Nacional Comum Curricular (BRASIL, 2017) para o componente Língua Portuguesa manteve a ênfase no texto como unidade de ensino da língua, ao organizar as competências e habilidades do currículo em campos de atuação dos sujeitos na vida social. À noção de campos de atuação está subjacente a assunção de que os sujeitos atuam na sociedade por meio dos textos que nela circulam e que se materializam em gêneros que apresentam características comuns, face ao campo da vida social em que circulam. Nesse sentido, cabe pensar que nas práticas de ensino e, conseqüentemente, de avaliação, uma questão fulcral é definir de que modo progride a abordagem dos textos e quais os percursos cognitivos dos estudantes na interação que estabelecem com eles. Essa interação vai depender da estrutura do texto e do tipo de sinalização que ele oferece ao leitor para construir sentido para o lido. Assim, textos de um mesmo gênero podem apresentar estruturas distintas, o que demandará ao leitor a mobilização de habilidades também distintas.

Em geral, a complexidade dos textos é tratada, nas práticas de ensino e nas avaliações em larga escala, com base na noção de gênero textual. Essa noção tem sua origem nos estudos de Mikhail Bakhtin, filósofo da linguagem,

que afirma: “Evidentemente, cada enunciado particular é individual, mas cada campo de utilização da língua elabora seus *tipos relativamente estáveis de enunciados*, os quais denominamos *gêneros do discurso*.” (BAKHTIN, 2003, p. 262 – grifo do autor). Os textos que têm origem nas diferentes esferas da vida social apresentam, portanto, características comuns, decorrentes da relativa estabilidade dos enunciados que se originam nessas esferas. Graças a tais características comuns, pode-se identificar o gênero textual. Uma carta pessoal, por exemplo, tem características diversas de um conto.

Sem dúvida, a noção de gênero textual é relevante para orientar a progressão das práticas de ensino da leitura, visto que é possível pensar que o ensino deva partir de gêneros que circulem em esferas mais próximas à vida cotidiana dos estudantes, portanto menos complexos, e avançar para gêneros que circulem em esferas mais especializadas, portanto, mais complexos. Também as avaliações em larga escala têm se valido dessa noção, tanto para a construção dos testes quanto para a divulgação de seus resultados. Entretanto, a categoria gênero não é suficiente para revelar, em maior detalhe, como os estudantes interagem com os textos, uma vez que tal processo de interação envolve diferentes habilidades que são mobilizadas pelo leitor não apenas em função do gênero textual com o qual se põe em interação, mas também em função da própria estrutura dos textos e do tipo de tarefa de leitura que é demandada do leitor.

Neste artigo se pretende, com base nos resultados obtidos em uma pesquisa desenvolvida pelo Centro de Avaliação e Políticas Públicas para a Educação (CAEd) da Universidade Federal de Juiz de Fora, apresentar uma classificação da complexidade dos textos em níveis, definidos a partir de algumas categorias, com base numa incursão a um conjunto amplo de textos utilizados como suportes de itens que avaliam a leitura, aplicados a estudantes do 5º ao 12º ano da educação básica, em diferentes estados e municípios da federação e também em avaliações nacionais, no Brasil.

Inicialmente, serão apresentadas algumas considerações sobre como o tema da complexidade dos textos tem sido tratado na divulgação dos resultados do *Programme for International Student Assessment (PISA)*, relatório de 2018, e no relatório do SAEB/ Prova Brasil, por meio da escala de proficiência em Língua Portuguesa.

Em seguida, é apresentada a metodologia e o *corpus* da pesquisa em que se fundamenta o presente artigo, assim como a classificação dos níveis de complexidade do texto a que se chegou com base nos dados da pesquisa realizada.

Finalmente, são tecidas algumas considerações sobre a relevância de haver maior clareza sobre o nível de complexidade do texto para os processos de elaboração dos testes que avaliam a leitura, assim como para a comunicação dos resultados das avaliações às escolas e aos professores.

A COMPLEXIDADE DOS TEXTOS NA COMUNICAÇÃO DE RESULTADOS DE AVALIAÇÕES INTERNACIONAIS E NACIONAIS

No relatório do Sistema de Avaliação da Educação Básica (SAEB) (BRASIL, 2019) são apresentados os resultados das avaliações de Língua Portuguesa e Matemática realizadas, no ano de 2017, no âmbito da Prova Brasil, em todo o território nacional. Nesse relatório, os desempenhos dos estudantes em ambos os componentes curriculares são apresentados por meio de uma escala de proficiência⁵. Com base na escala de proficiência, esses desempenhos são interpretados e descritos, na avaliação de Língua Portuguesa, leitura, em 9 (nove) níveis. Essa forma de apresentação e divulgação de resultados é aquela que vem sendo utilizada pelo INEP ao longo das diferentes edições da Prova Brasil.

No quadro 1, são apresentadas as descrições dos três primeiros níveis de desempenho da escala de proficiência em Língua Portuguesa, tal como descritos no relatório do SAEB (BRASIL, 2019).

⁵ Escala de proficiência: Conjunto de números ordenados obtido pela Teoria de Resposta ao Item (TRI) que representa a medida da proficiência em uma determinada área de conhecimento. No SAEB, em cada área do conhecimento, há uma escala de proficiência única para todos os anos escolares (séries) e todas as edições da avaliação (Ex.: a escala de proficiência em Língua Portuguesa da edição de 2017 é a mesma das edições anteriores e agrega as proficiências dos estudantes que participaram dos testes do 5º e do 9º ano do ensino fundamental e da 3ª série do ensino médio). Cada escala varia de 0 a 500 pontos, com média de 250 e desvio padrão de 50. Disponível em <http://portal.inep.gov.br/documents/186968/484421/Relat%C3%B3rio+Saeb+2017/e683ba93-d9ac-4c2c-8f36-10493e99f9b7?version=1.0> Acesso em: 10 dez. 2019.

Quadro 1 – Níveis de Desempenho em Língua Portuguesa segundo a escala de proficiência da SAEB (fragmento).

NÍVEL	DESCRIÇÃO DO NÍVEL
Nível 1 Desempenho maior ou igual a 125 e menor que 150	Os estudantes provavelmente são capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios. Identificar o tema de um texto. Localizar elementos, como o personagem principal. Estabelecer relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.
Nível 2 Desempenho maior ou igual a 150 e menor que 175	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em contos. Identificar o assunto principal e a personagem principal em reportagem e em fábulas. Reconhecer a finalidade de receitas, manuais e regulamentos. Inferir características de personagens em fábulas. Interpretar linguagem verbal e não verbal em tirinhas.
Nível 3 Desempenho maior ou igual a 175 e menor que 200	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informação explícita em contos e reportagens e em propagandas com ou sem apoio de recursos gráficos. Reconhecer relação de causa e consequência em poemas, contos e tirinhas. Inferir o sentido de palavra, o sentido de expressão ou o assunto em cartas, contos, tirinhas e histórias em quadrinhos, com o apoio de linguagem verbal e não verbal.

Fonte: Relatório SAEB 2017.

A descrição dos demais níveis da escala segue o mesmo padrão da descrição dos três primeiros, trazidos, aqui, a título de exemplificação. Com base nesse exemplo, observa-se que a descrição do desempenho em leitura dos estudantes se faz numa articulação entre a habilidade de leitura avaliada (por exemplo, localizar informação explícita, inferir, reconhecer a finalidade do texto) e o gênero textual utilizado como suporte aos itens que avaliam a habilidade (reportagem, poemas, conto, fábula). A progressão entre os diferentes níveis é dada pelo acréscimo de habilidades àquelas indicadas no nível anterior, entretanto, pela pesquisa que vimos desenvolvendo, constatamos que dessa forma não é possível compreender como se altera a complexidade da tarefa de leitura ou o que torna um gênero mais complexo que outro, do ponto de vista da interação do leitor com o texto. A título de exemplificação tem-se, no nível, o mais elementar, a descrição de que os estudantes, provavelmente, “são capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios.” Já no nível 2, no qual os estudantes teriam uma proficiência maior do que aqueles de nível 1, lê-se que os estudantes provavelmente são capazes de, além de realizar as tarefas de leitura do nível anterior “localizar informações explícitas em contos”. Os contos podem ser textos narrativos curtos. Além disso, não existem informações sobre a posição da informação a ser localizada e nem como ela deve ser recuperada pelo leitor, com base no tipo de situação problema construída no item. Desse modo, quanto à habilidade de localizar informações, não é possível perceber o que distingue um estudante com proficiência compatível com o nível 1 daqueles com proficiência compatível com o nível 2. Esse mesmo problema se repete com relação a outros descritores e na passagem entre outros níveis.

Conclui-se, a partir do exposto, que narrativas, como um conto, por exemplo, podem se materializar em textos com diferentes níveis de complexidade, consequência do tipo de sintaxe utilizada, do léxico empregado, do tema abordado, dentre outras possibilidades. Além disso, no caso da habilidade de localizar informações explícitas, a posição da informação solicitada e, ainda, a existência de maior ou menor quantidade de informações no texto também são fatores a considerar, além do modo como a informação é recuperada na situação problema proposta pelo item: se por meio de uma repetição, de uma quase repetição, ou por uma paráfrase simples, por exemplo. Desse modo, a descrição das habilidades de leitura tal como se apresenta no relatório do SAEB é pouco elucidativa da natureza das habilidades de leitura desenvolvidas por estudantes que se encontram nos diferentes níveis da escala de proficiência.

No relatório do PISA, o desempenho dos estudantes em leitura também é apresentado por meio de 6 (seis) níveis. No quadro 2, são apresentados os três primeiros níveis da escala de proficiência do PISA (OECD, 2019).

Quadro 2 – Níveis de desempenho em *reading literacy proficiency* do PISA (fragmento).

Level 1 B	Readers at level 1b can locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. Texts in level 1b tasks typically provide support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. Level 1b readers can interpret texts by making simple connections between adjacent pieces of information.
Level 1B	Readers at Level 1a can locate one or more independent pieces of explicitly stated information; they can recognise the main theme or author's purpose in a text about a familiar topic, or make a simple connection between information in the text and common, everyday knowledge. Typically, the required information in the text is prominent and there is little, if any, competing information. The student is explicitly directed to consider relevant factors in the task and in the text.
Level 2	Readers at Level 2 can locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. They can recognize the main idea in a text, understand relationships, or construe meaning within a limited part of the text when the information is not prominent and the reader must make low level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflecting tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.

Fonte: Assessment and Analytical Framework, PISA (OECD, 2019, s/p).

No relatório do PISA, é possível encontrar um maior detalhamento das tarefas de leitura que os estudantes que se encontram nos diferentes níveis de proficiência provavelmente podem realizar. Por exemplo, a habilidade de localizar informações é descrita como uma habilidade que pode ser mais ou menos complexa se o texto apresenta muitas informações que concorrem com aque-

la solicitada pelo item, se a informação solicitada é mais ou menos proeminente no texto. Não há referências ao gênero textual, mas a características gerais do texto, como ao fato de abordar um tema mais ou menos familiar. Neste caso, há uma análise mais detalhada da tarefa de leitura, mas tal detalhamento se faz sem uma definição mais precisa da natureza do texto, suas características e microestrutura, que podem tornar essas tarefas de leitura mais ou menos complexas. Entretanto, há um avanço em relação ao relatório do SAEB, no sentido de tornar mais evidentes as tarefas de leitura que os estudantes dos diferentes níveis têm probabilidade de realizar com êxito.

Os estudos desenvolvidos no âmbito da pesquisa de que trata o presente artigo têm mostrado que as habilidades de leitura desenvolvidas pelos estudantes estão fortemente relacionadas à estrutura do texto e ao tipo de sinalização que ele oferece para que os estudantes possam realizar a tarefa de leitura requerida, evidenciando o desenvolvimento de determinadas habilidades. Um poema, por exemplo, pode apresentar uma linguagem próxima àquela utilizada na vida cotidiana, o que exigirá uma proficiência em leitura relativamente menor, para sua compreensão pelo leitor; ou pode utilizar uma linguagem metafórica, o que demandará uma elevada proficiência em leitura para que seja compreendido pelo leitor.

A DEFINIÇÃO DOS NÍVEIS DE COMPLEXIDADE DO TEXTO NA PESQUISA EM AVALIAÇÃO CAED, ÁREA DE LÍNGUA PORTUGUESA

No âmbito da pesquisa em avaliação, área de Língua Portuguesa, foram analisados, ao longo dos anos de 2017 e 2018, quatro descritores que integram a Matriz de Referência para Avaliação do CAEd, quais sejam: D06 - Localizar informação explícita; D07 - Reconhecer o assunto de um texto; D08 - Inferir o sentido de uma palavra ou expressão; D09 - Inferir informação em um texto verbal. Em todos esses descritores, são utilizados como suporte para a construção da situação problema dos itens textos de diferentes gêneros. Constituíram o corpus da pesquisa 763 textos, que deram suporte a itens que avaliaram esses descritores.

As análises empreendidas no âmbito da pesquisa tinham como objetivo a identificação da progressão da dificuldade que cada uma das habilidades apresenta ao longo da escala de proficiência do SAEB. Para compreender tal progressão, a complexidade dos textos utilizados como suporte para a elaboração dos itens que avaliam cada uma das habilidades se mostrou um ele-

mento decisivo. Isso ocorreu porque a leitura pressupõe uma interação entre leitor e texto que são fatores relevantes à natureza do texto e à experiência do leitor. O texto, objeto central da investigação, configura-se como um conjunto de sinalizações que “guiam” o leitor para que ele produza sentidos para o que lê, e o leitor como alguém que atua nesse processo a partir de uma gama de conhecimentos – sobre o mundo, sobre temas específicos do mundo, sobre os textos, sobre a língua.

Para a análise dos descritores da matriz de Referência para Avaliação em Língua Portuguesa, foi realizada, inicialmente, uma explicitação da natureza linguística do descritor com base nas teorias do texto, sobretudo as teorias de base cognitivista (SMITH 1991; KLEIMAN 1989; KOCH 1998) e na linguística do texto (KOCH, 1989; KOCH, TRAVAGLIA 1989).

A TRI – Teoria da Resposta ao Item – e o modelo de três parâmetros foram aqueles que subsidiaram a análise dos resultados das avaliações. Para identificar os níveis de complexidade dos textos que deram suporte aos itens, foram eliminados: itens com parâmetro A menor que 0,01, uma vez que esses itens apresentam fraco poder de discriminação entre estudantes que desenvolveram as habilidades avaliadas e aqueles que não as desenvolveram e, ainda, itens com parâmetro C superior a 0,025, uma vez que esses itens apresentam forte probabilidade de acerto ao acaso. Foram considerados, na análise, o ponto de ancoragem dos itens na escala de proficiência em leitura do SAEB e a distribuição desses itens em intervalos de 25 pontos.

O segundo passo consistiu numa incursão aos itens que avaliam cada um dos descritores, buscando identificar a gradação da complexidade com que tais habilidades se manifestam ao longo da escala, o que ofereceu indícios da proficiência em leitura dos estudantes nas diferentes etapas de escolarização. Tal identificação se deu pela comparação entre textos que davam suporte a itens que ancoravam em pontos próximos da escala de proficiência, buscando identificar aspectos em comum entre esses textos. As definições das habilidades avaliadas, elaboradas na primeira etapa da pesquisa foram, também, fundamentais para a análise dos textos que davam suporte aos itens.

A análise das tarefas propostas pelos itens considerou tanto a complexidade do texto quanto a natureza das sinalizações que ele oferecia para a realização da tarefa. São exemplos de sinalizações de apoio à leitura: a presença, no texto, de sinônimos que auxiliam a tarefa de inferir o sentido de uma palavra; a presença de marcação tópica que facilita ao leitor a tarefa de reconhecer o assunto de um texto; um enunciado parafrástico que permite acessar uma

informação, dentre outras possibilidades. Por outro lado, o nível de complexidade de um texto – definido a partir de critérios como sintaxe, léxico, tema, extensão – tem, também, impacto sobre a complexidade da tarefa.

A classificação decorrente dos critérios definidos no âmbito da pesquisa considera quatro níveis de complexidade dos textos, em cuja definição a sintaxe se apresenta como elemento bastante definidor da complexidade, quando se consideram os dois primeiros níveis e o léxico e o tema se apresentam como bastante determinantes nos níveis três e quatro. Com base nos critérios tamanho, sintaxe, léxico e tema, definiram-se os seguintes níveis de complexidade dos textos a seguir.

Textos de Nível 1: os textos são pouco extensos (até 15 linhas). A **sintaxe** é simples, em que predominam os períodos curtos e estruturas coordenadas, ou ainda em tópicos; os processos coesivos são simples, como aqueles envolvendo retomadas com referente próximo – por referência através de pronomes pessoais do caso reto ou reiteração, por repetição ou sinonímia; a estrutura dos textos é convencional e as frases apresentadas na ordem canônica. O **léxico** é mais próximo do coloquial e adequado ao interlocutor preferencial do texto, normalmente criança. Os **temas** são mais cotidianos, uma vez que temas menos familiares requerem um tratamento que “facilite” a leitura do texto. Os textos mais comuns neste nível são de estrutura canônica, estando presentes as narrativas e os textos expositivos: sobre os primeiros, podemos destacar parlendas, lendas, narrativas da literatura infantil e, ainda, os poemas de estrutura narrativa; entre os segundos, se encontram as “curiosidades científicas” ou “notas científicas”, como definidos pela BNCC.

Textos de Nível 2: os textos são pouco extensos (até 15 linhas). A **sintaxe** é um pouco mais complexa do que aquela que se observa em textos de nível 1, com pequenas alterações na estrutura sintática canônica – orações reduzidas no início dos períodos e intercalações breves, por exemplo; há o predomínio de períodos um pouco mais longos, com duas linhas em média; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos mais usuais. Há textos que apresentam **léxico** mais usual ou aqueles cujo léxico é próprio do discurso formal, mas com estratégias de facilitação do discurso, como aposto depois de um vocábulo pouco usual ou de um termo científico ou técnico. Os **temas** são relacionados ao cotidiano e, quando menos familiares, o texto oferece pistas para a identificação de seu conteúdo. Os textos mais comuns são as narrativas de estrutura canônica e as argumentações breves e de estrutura mais simples.

Textos de Nível 3: os textos apresentam-se em qualquer **tamanho**, ou seja, tanto há textos curtos quanto muito longos entre os classificados nesse nível. A **sintaxe** é complexa e própria do uso formal da língua, com estruturas subordinadas, intercalações mais longas ou impessoalizações, por exemplo. Os períodos são mais longos, com cerca de três linhas; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos menos usuais, assim como retomadas de referentes mais distantes. O **léxico** é próprio do discurso formal, apresentado em textos diversos, de estrutura menos canônica. Os **temas** são pouco familiares e os suportes mais variados, voltados a interlocutores diversos. As narrativas que aparecem nesse nível são os contos, as crônicas, os fragmentos de romance – textos da literatura brasileira. Encontram-se nesse nível textos expositivos ou argumentativos mais longos, como notícias, reportagens e poemas de elaboração metafórica.

Textos de Nível 4: assim como no nível anterior, os textos de nível 4 apresentam-se em qualquer **tamanho**. A **sintaxe** é complexa e própria do uso formal da língua, com estruturas subordinadas, intercalações mais longas e impessoalizações, por exemplo. Os períodos são mais longos, com três linhas em média; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos menos usuais, assim como retomadas de referentes mais distantes. O **léxico** é próprio do discurso formal, especializado e erudito, apresentado em textos diversos, de estrutura menos canônica. O predomínio é de textos com tratamento literário mais elaborado, afetando sua estrutura e seus recursos, e aqueles que tratam de **temas** especializados, sobretudo do campo da ciência e da vida em sociedade. Como exemplo de textos narrativos, destacam-se textos clássicos: os contos, crônicas e fragmentos de romances, narrativas menos focadas na ação e com estratégias mais complexas, como o discurso indireto livre, descrições que capturam o fluxo de consciência da personagem. Nesse nível encontram-se, ainda, resenhas, notícias e reportagens, textos de opinião com vários argumentos e poemas com tratamento literário mais elaborado, com a presença de metáforas, alegorias e imagens poéticas.

Com base nos quatro níveis de complexidade dos textos aqui apresentados, foram analisados os itens referentes aos descritores referidos anteriormente e aplicados em avaliações realizadas pelo CAEd a alunos 5º ao 12º ano de escolarização. Nos quadros 3, 4, 5 e 6, a seguir, é apresentada a distribuição dos itens analisados, relacionados a cada descritor, por intervalos de 25 pontos da escala de proficiência.

Quadro 3 – Distribuição dos itens que avaliam a habilidade de localizar informações explícitas em textos por níveis de complexidade.

D06- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
100-125	1	0,31						
126-150	3	0,95	5	1,58				
151-175	16	5,07	29	9,20			1	0,31
176-200	6	1,90	38	12,06	5	1,58		
201-225	3	0,95	51	16,19	17	5,39	2	0,63
226-250			33	10,47	18	5,71	5	1,58
251-275			22	6,98	13	4,12	4	1,26
276-300			9	2,85	3	0,95	7	2,22
301-325			3	0,95	2	0,63	7	2,22
326-350			1	0,31	2	0,63	4	1,26
351-375							3	0,95
376-400								
Acima 401					1	0,31	1	0,31
Totais	29	9,18	191	60,59	61	19,32	34	10,74
Total Geral	315 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 4 – Distribuição dos itens que avaliam a habilidade de reconhecer o assunto de um texto por níveis de complexidade.

D07- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	3	1,33	1	0,54				
151-175	4	1,78	16	7,10	1	0,54		
176-200	3	1,33	17	7,58	11	4,91		
201-225			20	8,92	21	9,37		
226-250			14	6,25	35	15,62	3	1,33
251-275			5	2,23	14	6,25	16	7,10
276-300			6	2,67	9	4,01	2	0,89
301-325			1	0,54	2	0,89	3	1,33
326-350					7	3,12	6	2,67
351-375					1	0,54	2	0,89
376-400								
Acima 401					1	0,44		
Totais	10	4,44	80	35,83	102	45,25	32	14,21
Total Geral	224 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 5 – Distribuição dos itens que avaliam a habilidade de inferir o sentido de uma palavra ou expressão por níveis de complexidade

D08- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	1	0,65						
151-175	4	2,80	2	1,40				
176-200	7	4,89	12	8,39				
201-225	3	2,09	20	13,98	3	2,09		
226-250	8	5,59	19	13,29	1	0,65		
251-275	1	0,65	19	13,29	8	5,59		
276-300			3	2,09	13	9,09		
301-325			1	0,65	11	7,69		
326-350			2	1,40	2	1,40		
351-375					1	0,65		
376-400			1	0,65	1	0,65		
Acima 401								
Totais	24	16,67	79	55,14	40	27,81		
Total Geral	143 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 6 – Distribuição dos itens que avaliam a habilidade de inferir informação em um texto verbal por níveis de complexidade.

D09- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	1	1,23	1	1,23				
151-175	2	2,46	1	1,23				
176-200	3	3,70	11	13,58				
201-225			5	6,17	2	2,46	1	1,23
226-250	1	1,23	7	8,64	7	8,64	2	2,46
251-275			2	2,46	6	7,40	3	3,70
276-300			1	1,23	12	14,81	3	3,70
301-325					1	1,23	2	2,46
326-350					3	3,70	2	2,46
351-375					1	1,23	1	1,23
376-400								
Acima 401								
Totais	7	8,62	28	34,54	32	39,47	14	17,24
Total Geral	81							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Como é possível observar pela análise dos quadros 3, 4, 5 e 6, os textos de nível 1 são aqueles predominantes em itens que ancoram entre 125 e 200 pontos da escala de proficiência em leitura do SAEB. Os textos de nível 2 predominam como suporte de itens que ancoram entre 151 e 275 pontos da escala de proficiência. Textos de nível 3 aparecem predominantemente como

suporte de itens que ancoram entre 176 e 300 pontos da escala. Finalmente, os textos de nível 4 aparecem como suporte de itens que ancoram acima de 226 pontos da escala, com maior concentração entre itens que ancoram acima de 250 pontos da escala. A comparação entre a distribuição dos textos, de acordo com seu nível de complexidade, entre os diferentes descritores da matriz de referência, indica uma certa regularidade dessa distribuição. Tal regularidade é indicativa de padrões de desempenho em leitura apresentados pelos estudantes que realizaram os testes e que se encontram em diferentes etapas da educação básica. É possível afirmar, por exemplo, com base na descrição dos níveis de complexidade do texto, que estudantes com proficiência entre 125 e 200 pontos leem textos pouco extensos (até 15 linhas), que apresentam sintaxe simples, em que predominam os períodos curtos e estruturas coordenadas, ou em tópicos, com processos coesivos simples, como aqueles envolvendo retomadas com referente próximo, por referência através de pronomes pessoais do caso reto ou reiteração, por repetição ou sinonímia. A estrutura dos textos que esses estudantes leem é convencional, com frases apresentadas na ordem canônica (sujeito, verbo, objeto), o léxico é mais próximo do coloquial e os temas mais cotidianos. Parlendas, lendas, narrativas da literatura infantil, os poemas de estrutura narrativa, as “curiosidades científicas” ou “notas científicas”, como definidos pela BNCC, são os gêneros com os quais os estudantes com essa proficiência leem.

De acordo com a análise realizada, o que distingue os textos de nível 1 dos textos de nível 2 é fundamentalmente a sintaxe. Pequenas alterações da estrutura sintática canônica (como intercalações e antecipações), bem como a extensão dos períodos, podem tornar um texto mais complexo para leitores iniciantes que não possuem fôlego de leitura. O que se observa é que a extensão dos períodos é um fator mais determinante para a complexidade do texto do que, propriamente, a extensão do texto. Por outro lado, textos de nível 1 e 2 caracterizam-se por exibir léxico mais familiar, próximo do uso coloquial da linguagem.

O léxico se apresentou, nesta investigação, como elemento definidor da complexidade de textos de nível 3 e 4. Textos de nível 3 apresentam, normalmente, léxico próprio do discurso formal, implicando em atividade inferencial mais constante. Textos de nível 4 apresentam léxico especializado, presente em textos cuja temática também se especializa, ou seja, textos que abordam questões de áreas de conhecimento específicas: ciência e tecnologia, vida em sociedade etc.

No entanto, a consideração desses traços do léxico (se mais, ou menos, próximo do discurso formal) não deve ser rígida. Em alguns textos, por exemplo,

usos formais da língua são seguidos de recursos de “facilitação do discurso” que permitem lidar com um léxico desconhecido. Textos que tratam de temas especializados e que são dirigidos ao leitor criança ou adolescente costumam exibir várias estratégias para facilitar sua recepção (como expressões explicativas, sinonímia, exemplos etc.). Léxico e tema são, portanto, critérios importantes para a classificação de textos de nível 3 e 4.

A investigação feita até aqui permite propor que a competência de leitura é fortemente determinada, inicialmente, pelo domínio da sintaxe da língua. Os leitores iniciantes vão, aos poucos, apropriando-se da sintaxe da língua, de seu uso formal, a partir do contato com textos cada vez mais complexos do ponto de vista sintático. O amadurecimento desse leitor se consolida à medida em que ele vai ampliando seu repertório de leitura e, conseqüentemente, seu repertório lexical, familiarizando-se com textos de vários campos do conhecimento, dominando sua estrutura, abordando temas diversos cada vez mais especializados. Essa competência se amplia ainda mais quando se compreende o potencial expressivo da linguagem recriada literariamente.

CONSIDERAÇÕES FINAIS

Os resultados da pesquisa aqui apresentada indicam que a complexidades dos textos é um critério importante a ser considerado na descrição de padrões de desempenho em leitura, definidos com base em avaliações de larga escala. Considerar os níveis de complexidade dos textos que dão suporte aos utilizados nessas avaliações permite avançar para além da descrição das tarefas de leitura propostas por esses itens e vislumbrar competências de leitura mais amplas, possivelmente desenvolvidas pelos estudantes que se encontram em determinados intervalos de proficiência.

Os resultados da pesquisa apontam, ainda, que, na definição dos níveis de complexidade que um texto pode apresentar, é necessário considerar um conjunto de características, que se combinam para que um texto seja mais ou menos desafiador ao leitor e que, ao longo do processo de formação do leitor, esses fatores desempenham um papel diverso. Para leitores com pouca proficiência, a sintaxe do texto é um fator bastante definidor da complexidade do texto. Para leitores mais experientes, o tema desempenha um papel mais determinante para que um texto possa ser considerado mais ou menos complexo. Deduz-se, do exposto, que a categoria gênero e a extensão do texto não são suficientes para uma descrição mais consistente das habilidades de leitura desenvolvidas pelos estudantes. Por exemplo, um poema, a despeito de sua extensão, pode se apresentar como de leitura mais ou menos comple-

xa, a depender de sua estrutura: se for um poema narrativo que faça uso de uma linguagem coloquial, com versos de estrutura simples, ele será de leitura mais fácil do que se apresentar uma linguagem metafórica e versos com uma estrutura menos canônica.

Dado o exposto, conclui-se que a descrição de padrões de desempenho em leitura deva considerar uma análise mais criteriosa dos suportes utilizados na construção das situações problema propostas pelos itens, para além da consideração do tipo de habilidade avaliada. Nesse sentido, é importante que, ao se descrever a habilidade, por exemplo, de localizar uma informação explícita em um texto, se considere de que tipo de texto se está falando e do tipo de apoio oferecido por ele para que o leitor realize a tarefa proposta. Se considerarmos que a leitura é um processo de interação entre leitor e texto, é preciso considerar como o texto conduz o leitor nessa interação, além de considerar o que o leitor traz para ela.

Outra consideração importante com base nos dados da pesquisa é a de que uma maior clareza quanto às características dos textos, obtida por meio da definição de critérios claros para a classificação de seus níveis de dificuldade, favorece a produção de itens que contemplem, de forma mais satisfatória, a avaliação de leitura de estudantes que se encontram em etapas específicas de escolarização, considerando o que está previsto na norma curricular. Assim, uma próxima etapa da pesquisa consistirá em cotejar a classificação dos níveis de dificuldade dos textos, produzida com base na incursão aos itens já utilizados em avaliações de larga escala, para estabelecer uma relação entre esses níveis e as diferentes etapas da educação básica.

Destaca-se, ainda, o fato de que a maior clareza quanto ao nível de dificuldade dos textos é um conhecimento relevante para subsidiar a produção de itens que avaliam outras áreas de conhecimento, como as ciências da natureza e as ciências humanas, vez que nessas áreas também são utilizados textos verbais que dão suporte à elaboração dos itens que as avaliam.

REFERÊNCIAS

BAKHTIN, M. **Estética da criação verbal**. São Paulo: Martins Fontes, 2003.

BRASIL. **Base Nacional Comum Curricular (BNCC)**. Educação é a Base. Brasília, MEC/CONSED/UNDIME, 2017. Disponível em: [568http://basenacionalcomum.mec.gov.br/images/BNCC_publicacao.pdf](http://basenacionalcomum.mec.gov.br/images/BNCC_publicacao.pdf). Acesso em: 10 dez. 2019.

BRASIL. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais (PCNs):** terceiro e quarto ciclos do ensino fundamental: Língua Portuguesa. Brasília: MEC/SEF, 1998.

BRASIL. **Relatório SAEB 2017.** Brasília, 2019. Disponível em: http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkWI/document/id/6730262. Acesso em: 10 dez. 2019.

KLEIMAN, A. Compreensão Leitora. In: **Glossário CEALE.** Disponível em: <http://ceale.fae.ufmg.br/app/webroot/glossarioceale/verbetes/compreensao-leitora>. Acesso em: 10 dez. 2019.

_____. **Texto e leitor:** aspectos cognitivos da leitura. Campinas, SP: Pontes, 1989.

KOCH, I. G. V; TRAVAGLIA, C. **Texto e Coerência.** São Paulo: Cortez, 1989.

KOCH, I. G. V. **O texto e a construção dos sentidos.** São Paulo: Contexto, 1998.

_____. **A coesão textual.** São Paulo: Contexto, 1989.

OECD. **PISA 2018.** Assessment and Analytical Framework, PISA, OECD Publishing, Paris, 2019. Disponível em: <https://doi.org/10.1787/b25efab8-em>. Acesso em: 10 dez. 2019.

PRIMI, R. **Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes.** Avaliação psicológica, 2012. Disponível em: <http://pepsic.bvsalud.org/pdf/avp/M1n2/M1n2a15.pdf>. Acesso em: 10 dez. 2019.

SMITH, F. **Compreendendo a leitura:** uma análise psicolinguística da leitura e do aprender a ler. Porto Alegre: Artes Médicas, 1991.

SOUSA, L. B. de; HÜBNER, L.C. Desafios na avaliação da compreensão leitora: demanda cognitiva e leiturabilidade textual. **Revista neuropsicologia latinoamericana**, 2014. Disponível em :https://www.neuropsicolatina.org/index.php/Neuropsicologia_Latinoamericana/article/view/237. Acesso em: 10 dez. 2019.

Recebido em: 10 de dezembro de 2019

Aceito em: 15 de junho de 2020

Publicado em: 30 de junho de 2020

A AVALIAÇÃO EXTERNA E OS NOVOS SUJEITOS DA EDUCAÇÃO



Marcelo Baumann Burgos¹

¹Professor e pesquisador do Departamento de Ciências Sociais da Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil. E-mail: burgos@puc-rio.br

Resumo

O lugar social ocupado pela avaliação externa no funcionamento do sistema educacional vem mudando com o tempo: de instrumento de controle externo e verticalizado do trabalho escolar, se converte em ferramenta fundamental ao cotidiano escolar. Mas essa mudança de lugar não estaria ocorrendo sem que os próprios profissionais da educação se transformassem. Neste artigo, analisamos essa mudança, valorizando, em especial, a forma como ela vem se dando no contexto brasileiro. Sustentamos, ao final, que o processo de decantação da avaliação externa na cultura profissional escolar vem forjando o surgimento de novos sujeitos no mundo da educação, porta-vozes e difusores de uma nova linguagem.

Palavras-chave: Avaliação Externa. Escola. Cultura Profissional. Gestão Escolar. Atividade Docente.

Abstract

The social place occupied by external evaluation in the functioning of the educational system has changed over time: from an external and vertical control instrument of school work, it becomes a fundamental tool for school daily life. But this change of place would not be taking place without the education professionals themselves changing together. In this article, we analyze these changes, valuing, in particular, the way it has been happening in the Brazilian context. Finally, we argue that the process of decantation of external evaluation in the school professional culture has been forging the emergence of new subjects in the world of education, spokespersons and diffusers of the new language.

Keywords: External Evaluation. School. Professional Culture. School Management. Teaching Activity.

INTRODUÇÃO

Com o uso cada vez mais rotineiro da avaliação externa pelas redes educacionais, é inegável que a presença crescente da linguagem dos descritores, das escalas de proficiência e da Teoria da Resposta ao Item, entre outras categorias desse novo idioma, não apenas interfere na forma como o currículo é customizado, a fim de se fazer valer a melhorias dos resultados nos testes, mas também no modo como a escola é tocada pela linguagem de uma gestão orientada para resultados. Por isso, de uma perspectiva sociológica, torna-se particularmente interessante estudar a forma como as tecnologias sociais que a avaliação externa encerra é apropriada pelos diferentes segmentos do sistema educacional, em especial, pelos componentes da comunidade escolar; bem como compreender o modo como esse processo de apropriação transforma os próprios sujeitos da educação, inclusive o professor.

Neste artigo, menos do que aprofundar de modo empírico e detalhado os diferentes aspectos dessa agenda de pesquisa, o que se pretende é tão somente delimitá-la, assumindo, como hipótese, que estamos vivendo uma transição do lugar que a avaliação externa ocupa nos sistemas educacionais: de instrumento adotado para medir a qualidade do trabalho educacional, a avaliação externa converte-se em uma linguagem, que incide sobre as formas de comunicação, portanto, de ação, dos diferentes atores que participam do trabalho escolar.

Talvez ajude na compreensão do argumento deste artigo explicitar que sua abordagem segue de perto a proposta teórica de Habermas (1997), por sugerir que a avaliação externa estaria deixando, gradualmente, de ser operacionalizada por meio de uma linguagem sistêmica, e passando a fazer parte de um renovado senso comum escolar, aproximando-se, por essa via, do agir comunicativo. E a razão comunicativa, como ensinou Habermas (1997, p.20), “distingue-se da razão prática por não estar adscrita a nenhum ator singular nem a um macrossujeito sociopolítico”. O que a torna possível, prossegue Habermas, “é o *medium* linguístico, através do qual as interações se interligam e as forma de vida se estruturam” (Idem). Nossa sugestão é a de que a linguagem da avaliação externa estaria assumindo essa condição de *medium* linguístico, o que significa que ela passa a conformar um novo ambiente, tornando-se fonte de uma racionalidade orientada para o entendimento.

De um modo geral, onde quer que o uso da avaliação externa tenha se difundido, encontrou, em um primeiro momento, grande resistência por parte dos profissionais da educação. Não por acaso, foi relativamente comum a per-

cepção do uso da tecnologia da avaliação externa como instrumento de submissão do sistema educacional a uma racionalidade puramente gerencialista, não faltando, também, quem a acusasse de ser um instrumento a serviço da ideologia neoliberal. Nessa toada, a avaliação externa tem sido muitas vezes encarada como uma espécie de veneno que mina a relação entre a educação escolar e a cultura democrática.

No Brasil, onde se poderia alegar que a fragilidade da escola a deixaria mais vulnerável a esse tipo de efeito redutor do sentido mais geral do trabalho escolar, esse tipo de leitura crítica foi e ainda é muito influente. Mas ela também esteve no centro dos debates públicos acerca da educação nos Estados Unidos da América, onde a defesa da escola democrática e para a democracia tem como um de seus alvos o uso pouco cuidadoso e, por vezes punitivista, da avaliação externa (APPLE; BEANE, 2001; RAVITCH, 2011).

É verdade que a crítica ao uso da avaliação externa, por parte das estruturas de comando dos sistemas educacionais que inclui a acusação de um certo fetichismo dos dados, não é de modo algum injusta, colocando em foco práticas que, de fato, não raro desvirtuam o bom uso da avaliação, fazendo dela uma espécie de fim em si mesma, e não um meio de regulação e coordenação voltado para a melhoria escolar. Aqui a crítica, formulada em geral por professores que sentem sua autonomia comprimida e seu trabalho ameaçado por novas formas de controle, é em princípio salutar, obrigando a um permanente aprimoramento da relação entre o uso da avaliação externa e os interesses mais imediatos da comunidade escolar. O questionamento de que a avaliação externa serve basicamente como um instrumento de controle externo, que não traz insumos efetivamente úteis para a melhoria do ensino e aprendizagem; e o questionamento da forma como a estrutura de avaliação costuma estar organizada, mais reforçando a condição da escola como mera unidade de uma estrutura administrativa, do que favorecendo sua capacidade própria para se apropriar dos resultados de modo a fazer deles aliados da comunidade escolar, são apenas duas das pontas de lança da resistência oferecida pelos profissionais de educação ao uso intensivo da avaliação externa.

No fundo, é em decorrência desse embate entre, de um lado, agentes políticos – às vezes investidos do papel de gestores – empenhados em ampliar o alcance do uso da avaliação externa, chegando a fazer dela instrumento para a definição de estratégias de incentivo e de superação da resistência à mudança, e âncora de propostas de reformas educacionais em larga escala (BROOKE, 2012); e de outro, profissionais que se ressentem do efeito transformador provocado pelo uso dessa tecnologia. É desse embate, em suma, que,

gradualmente, a avaliação externa vai deixando de ser apenas instrumento de controle e de prestação de contas, para se converter em uma tecnologia cada vez mais percebida como aliada da valorização da cultura de responsabilidade profissional e, portanto, como ferramenta indispensável ao trabalho dos administradores escolares e dos docentes.

Muda o uso da tecnologia e mudam os sujeitos da educação escolar. Entra em cena um administrador gestor, que passa a formular um cálculo estratégico, parametrizado por evidências, sobretudo, por resultados; ganham novo protagonismo as instâncias regionais, como dimensões fundamentais para auxiliar a escola e seus gestores; e mudam os professores, que perdem seu *ethos* característico de profissional liberal, detentor de um saber especializado, em favor do de um funcionário de uma estrutura de prestação de serviços públicos. Embora trate especificamente do caso inglês, há algo de geral na caracterização que Lawn (2000) faz dessa virada, que ele situa nos anos de 1990, quando o professor já não é mais o protagonista da reconstrução nacional do contexto do pós-guerra, nem o profissional “tranquilamente limitado pelas paredes da sala de aula” (LAWN, 2000, p.81). Sua nova identidade segue sendo baseada na sala de aula, mas agora ele é “um trabalhador da escola, como deveres para além da sala de aula, sobre os quais serão inspecionados” (LAWN, 2000, p.81).

Este artigo pretende tratar dessa virada, caracterizando os contornos mais amplos desse processo no cenário internacional e apresentando, em linhas gerais, a forma peculiar como ela vem se dando no Brasil. A caracterização dessa virada no debate internacional será tratada na seção seguinte e, para tanto, mobilizamos a noção de pós-gerencialismo, para melhor caracterizar como a mudança de concepção da administração escolar e das políticas educacionais em geral vem dando lugar a uma ampla redefinição do uso das ferramentas de avaliação externa. Na terceira seção deste artigo, será realizada a análise das mudanças ora em curso na realidade brasileira. Nosso argumento é de que aqui se vive, desde o final dos anos de 1990, um processo de transição que, de certo modo, altera a antinomia entre a escola da “gestão da pobreza” (PEREGRINO, 2010) e a escola orientada para resultados (DUSI, 2017), que teria animado o impulso reformista dos anos de 2000, em favor de uma nova configuração, na qual estaria ganhando força nas redes educacionais e nas escolas, a importância do uso das evidências como referência para a tomada de decisões. E que essa crescente afirmação da cultura do uso de evidências, em grande parte oriunda de avaliações externas – mas também de fontes de dados geradas pela própria escola –, seria a base para um processo de amadurecimento da relação entre o compromisso com os resultados e as formas

mais participativas de reflexão e de planejamento da educação escolar. Na última seção, à guisa de conclusão, realçamos o aspecto central do argumento deste artigo, que é o da conformação de novos sujeitos na educação, portadores da linguagem da avaliação externa, com sua lógica própria de ação.

O PÓS-GERENCIALISMO E AS NOVAS FORMAS DE USO DA AVALIAÇÃO EXTERNA

Recuperamos aqui parte do argumento apresentando em artigo anterior (BURGOS; BELLATO, 2019), no qual sustentamos que estaria em curso uma leitura crítica do modelo adotado desde os anos de 1980, em países como Inglaterra, Estados Unidos, Canadá, entre outros, o qual tinha como tônica a introdução de elementos gerencialistas na administração pública e, em especial, na área da educação. De modo mais específico, essa marca gerencialista se faria presente pela valorização de um modelo de administração orientado pela valorização da definição de metas a serem alcançadas, da responsabilização pelos resultados e da organização de um sistema de monitoramento e avaliação externa do desempenho escolar.

A resenha da literatura internacional sobre o assunto permitiu identificar, no entanto, a conformação de uma nova tendência que, mesmo sem romper completamente com a plataforma gerencialista, introduziu mudanças importantes que nos autorizam a falar, no mínimo, da emergência de um pós-gerencialismo.

Especialmente importante para o presente artigo, é a delimitação de como essa tendência, que denominamos pós-gerencialista, lida com a avaliação externa e, em um sentido mais amplo, com o uso de evidências para a gestão e a administração escolar. Nesse sentido, um primeiro ponto diz respeito ao reconhecimento de que, apesar da inegável importância para mensurar o quanto as escolas estão cumprindo sua função educativa, a tendência pós-gerencialista defende que as avaliações externas são limitadas e, em alguns casos, podem esconder importantes questões contextuais. Disso se seguiria a necessidade de se combinar instrumentos de avaliação externa com formas de avaliação interna. Ou seja, a recomendação é a de que, além do monitoramento da performance do estudante por meio de testes periódicos, também se valorize uma autoavaliação como um processo contínuo de reflexão, que considere o reconhecimento de que os aprendizados do aluno e do professor estão integralmente conectados ao aprendizado organizacional (MACBEATH, 2010). Além disso, também se pontua a importância do contexto escolar como

parte da avaliação, já que a abstração das condições concretas da realidade escolar poderia desestimular processos de melhoria justamente nas escolas mais sensíveis (TEDDLIE, 2010).

Esse tipo de pretensão leva a que se enfrente o problema da sustentabilidade das reformas. Para autores como Hopkins (2010), costuma predominar o uso da avaliação como um espelho, obedecendo a uma visão de curto prazo. Com isso, ainda que resultados imediatos possam aparecer, não se alteraria nada de mais profundo no sistema. A saída estaria na conciliação entre uma concepção sistêmica, que assegure um alcance geral, mas sempre com o risco da homogeneização, e uma maior abertura à participação dos profissionais da educação, que trariam as especificidades de cada escola.

Mas é na problematização da sustentabilidade de reformas guiadas pela métrica dos resultados escolares que se verifica uma inflexão fundamental na maneira como o pós-gerencialismo lida com o uso da avaliação externa. Agora, ganha relevância o foco na construção de capacidades que possam impactar na conformação de uma nova cultura profissional, comprometida com o desempenho estudantil. Disso, se segue a valorização da discricionariedade e da criatividade (pontos centrais neste novo equilíbrio entre abordagens verticais e horizontais de reforma) daqueles que, como os professores, se fazem, às vezes, de “burocrata de nível de rua” para evocar a conhecida imagem formulada por Lipsky (1980). A valorização da base, neste caso, não deve ser tomada como uma simples inversão de uma cadeia de comando antes verticalizada, mas como uma aposta no profissionalismo e na cooperação docente. E para isso acontecer, é necessário que as estratégias de reforma abordem o contexto do ensino e da aprendizagem, bem como o da capacitação no nível escolar. Desloca-se, com isso, a tônica do controle exercido verticalmente, em favor de um novo tipo de controle, mais horizontal, e baseado na aposta no profissionalismo e na cooperação.

Quanto a isso, Joan Talbert (2010) vai ao ponto, quando acusa a ênfase nos ganhos de curto prazo e nos resultados dos testes de impor um ritmo de mudança que mina o desenvolvimento de comunidades profissionais de aprendizado (TALBERT, 2010, p.560). Para ela, os sistemas centralizados diluem a responsabilidade coletiva, já que a preocupação é com a conformidade e não com o efeito da ação do profissional. Por outro lado, o desenvolvimento da responsabilidade mútua dos professores se daria a partir da valorização da categoria profissional, bem como da sua articulação com a comunidade. A aposta no profissionalismo e na cooperação remete à valorização de uma perspectiva mais em rede do que hierárquica, bem como à construção de

capacidade e responsabilidade mais do que de controle. Na mesma linha, Hargreaves (2009) sustenta que a ênfase sobre a responsabilidade profissional, ao invés da vigilância constante, preservaria, segundo ele, a criatividade e a possibilidade de customização em face de diferentes realidades.

Essa é também a perspectiva defendida por Linda Darling-Hammond (2009) que, mesmo admitindo que a padronização traz alguma eficiência, sustenta que estas “*regulamentações sem rosto*” acabam se tornando o bode expiatório do fracasso escolar, já que ninguém no sistema assume a responsabilidade pelos seus efeitos sobre os alunos, ao passo que o incremento do profissionalismo, por sua vez, poderia promover uma melhora contínua da prática ao substituir a ideia de conformidade pela de responsabilidade. Afinal, assevera, a autora, que o trabalho docente é muito complexo para ser prescrito de longe, por isso, o controle não deve vir de cima, mas, sobretudo, dos mecanismos de incentivo e sanção decorrentes da responsabilidade em face dos pares e da reputação profissional.

Em síntese, a emergência do tema do profissionalismo, aí incluída a ênfase na cooperação profissional, revela a valorização das formas coletivas de organização em torno do mundo da escola, sobretudo daquelas comprometidas com uma melhoria sustentável. A prioridade passa a ser a da qualidade da comunicação entre os profissionais, bem como a da valorização do controle a partir de uma ética profissional, e não mais a conformidade aos desígnios burocráticos ou a vontade exclusiva de uma elite modernizadora.

É esse o espírito de uma literatura que, sobretudo no contexto anglo-saxônico, vem buscando reconciliar o projeto de reforma do sistema educacional com critérios afins à construção de sociedades democráticas. Aqui, importa enfatizar que, enquanto na modelagem tipicamente gerencialista, a defesa da natureza democrática da escola ancorava-se apenas no resultado escolar, agora, está em jogo a valorização do ambiente democrático da escola, sem descuidar da valorização dos resultados. É esse desafio, afinal, que a literatura pós-gerencialista está se propondo a enfrentar: reconectar a escola com valores e princípios democráticos sem prejuízo da retomada dos aspectos considerados preciosos dos diferentes períodos da reforma educacional, tais como a aposta em um ambiente propício à inovação, a valorização da definição de diretrizes amplamente divulgadas e dos avanços técnicos das avaliações externas, a busca de equidade e o uso de dados e de evidências empíricas para informar gestores e professores em suas intervenções (HARGREAVES, 2009).

Nesse novo enquadramento, portanto, tem-se um cenário no qual o uso dos resultados da avaliação externa desloca-se de sua condição de linguagem do estado e do poder público para a condição de linguagem do profissional da educação. A valorização da apropriação profissional dessa linguagem técnica demarca uma redefinição do próprio significado da relação da escola com a democracia ao redefinir a relação entre os profissionais eles mesmos e a relação com os estudantes. A responsabilidade pelos resultados passa a ser, agora, o marco de uma nova ética, de um profissional que não pode mais se dar ao luxo de prescindir do uso da avaliação externa como aliada desse novo padrão de controle.

Uma vez vislumbrada essa nova e promissora possibilidade, ganha renovado impulso o investimento público em novas formas de avaliação, com a crescente valorização da avaliação formativa, e das formas de monitoramento do aprendizado. Assim, a prevalecer essa tendência, poder-se-ia afirmar que a tecnologia da avaliação externa deixará de ser prerrogativa do vértice dos sistemas educacionais para ganhar, cada vez mais, lugar privilegiado na cultura profissional. No Brasil, como veremos na próxima seção, essa tendência já se faz presente, ainda que timidamente. Algo disso está visível, especialmente quando se valoriza a relação que vem sendo estabelecida entre formação continuada do professor e uso dos dados de avaliação externa como evidência fundamental à melhoria escolar.

TRANSIÇÃO DA AVALIAÇÃO EXTERNA NO BRASIL: DA ESCOLA ORIENTADA POR RESULTADOS À ESCOLA QUE VALORIZA AS EVIDÊNCIAS COMO BASE PARA A PARTICIPAÇÃO E A COOPERAÇÃO

A massificação do acesso à escola no Brasil foi acompanhada por um processo que tem sido caracterizado por autores como Mônica Peregrino (2010) como de “precarização institucional”. Ao longo da primeira grande onda de massificação, ocorrida entre os anos de 1970 e 1980, quanto mais o acesso à escola se expandia, mais ela se precarizava, e, para usar um termo proposto pela autora, mais se “desescolarizava”. Esse enquadramento se torna inteligível, quando se leva em conta que a chegada maciça de crianças pobres à escola faz com que ela assuma parte importante do trabalho de “gestão da pobreza”, anteriormente realizado pelos empregadores (vale lembrar que o trabalho infantil ainda era amplamente utilizado), por organizações filantrópicas, em geral assistencialistas, e pela Igreja Católica.

É com a crescente denúncia da baixa capacidade da escola para ensinar, fato que começa a ocorrer com a divulgação dos primeiros resultados de avaliação externa, que ganha força um processo de reformas educacionais as quais, pouco a pouco, vão alterando o sentido mais geral do trabalho escolar no país. No lugar da gestão da pobreza, entra em cena um modelo tipicamente gerencialista, que anima o governo federal, além dos estaduais e municipais, a realizarem esforços no sentido de reformar o sistema educacional, segundo o modelo de gestão para resultados, como bem a define Cristina Dusi (2017). E foi como parte dessa concepção sistêmica de reformas orientadas para a melhoria escolar, que se desenvolveu no Brasil um contínuo esforço de aprimoramento do uso de resultados de avaliação externa.

Mas essa onda reformista animada pelo receituário gerencialista tem dado sinais de esgotamento. E como argumentamos em já citado artigo (BURGOS; BELLATO, 2019), dois processos políticos parecem denotar bem isso: o movimento de ocupação das escolas, realizado entre 2015 e 2016, por estudantes secundaristas, em boa parte das unidades da federação; e a reforma do ensino médio, promovida pelo governo federal, em 2016. O principal alvo do movimento de ocupação das escolas, importa lembrar, foi exatamente a crítica a um modelo de gestão considerado tecnocrático, pouco transparente, e pouco aberto à participação estudantil. Já a reforma do ensino médio, proposta pelo Ministério da Educação, logo após a divulgação dos resultados do IDEB, vem como resposta do governo à estagnação do desempenho no ensino médio, logo tratado como uma “tragédia”. E foi sob o signo da tragédia, que se propôs, por meio de Medida Provisória nº 748/2016, uma reforma basicamente voltada para a alteração da estrutura curricular do ensino médio, criando, entre outras mudanças, a figura dos itinerários para, supostamente, oferecer maior liberdade aos estudantes². O fato de ter sido realizada em caráter de urgência (para justificar o uso da MP) e a adoção da categoria “tragédia” para caracterizar o ensino médio denotam bem o despreço pelo acúmulo anterior, realizado a partir de uma matriz gerencialista. O fato é que, com essa reforma do ensino médio, a dimensão da gestão perde centralidade.

Para além desses processos políticos ou subjacentes a eles, dois tipos de desafios administrativos explicitam bem esse esgotamento do modelo gerencialista. O primeiro tem a ver com as dificuldades em face da resistência à mudança de comportamento profissional, que costuma ser muito forte na área da educação. De fato, apesar dos avanços no esforço de qualificação e certi-

² “Não vamos fazer de conta que essa tragédia não existe”, declarou, à época, o recém-empossado ministro da Educação, Eduardo Mendonça Filho. Disponível em: <http://g1.globo.com/educacao/noticia/ideb-no-ensino-medio-fica-abaixo-da-meta-nas-escolas-do-brasil.ghtml>. Acesso em 12 set. 2017.

ficação profissional dos diretores, a onda reformista realizada pelos estados e municípios não transformou senão parcialmente a cultura profissional de tipo personalista que caracteriza a relação dos diretores com as escolas (BURGOS; CANEGAL, 2011), diretores esses que, vale lembrar, ainda são, em boa parte das redes do país, alçados ao cargo por indicação política. Entre os professores, o investimento de algumas redes na formação continuada, associado ao uso intensivo de mecanismos de monitoramento e avaliação, tem produzido resultados interessantes no sentido de alterar padrões de comportamento docente, aumentando sua responsabilidade com o resultado escolar. Mas mesmo nesses casos, de sucesso ainda relativo, a sua sustentabilidade depende de que o modelo gerencialista se transforme em formas mais horizontais de gestão do trabalho escolar³.

O segundo desafio tem a ver com as dificuldades de se produzir alterações mais consistentes no desempenho dos estudantes, em especial nos anos finais do ensino fundamental e no ensino médio. A questão da melhoria dos resultados escolares esbarra, entre outros fatores, na dificuldade de se conceber escolas mais responsivas em relação aos seus alunos, com mais foco no estudante e em sua realidade, e maior abertura à sua participação. Essas condições exigiriam, no mínimo, muitas mudanças na proposta de se nortear a busca de resultados apenas com base na lógica das metas, incentivos e bonificação por resultados. Ironicamente, a reforma do ensino médio vai em outra direção, não valorizando a agenda do aprimoramento da gestão escolar e tampouco contemplando a necessidade de se fortalecer a participação docente e discente no trabalho escolar.

Assim é que, para além das estéreis disputas ideológicas tão persistentes na cena contemporânea e que atrapalham e eludem o desafio real ora colocado para o avanço da educação no Brasil, o que está ou deveria estar em jogo nesse momento é o aprofundamento e não o seu menosprezo das significativas conquistas das últimas décadas. Mas se é verdade que aquelas conquistas geraram novos sujeitos, como argumentaremos melhor mais a frente, e que esses sujeitos defenderão a continuidade dos processos em curso antes de 2016, então, talvez faça sentido apostar que, em médio prazo, uma nova onda de inovações processuais na gestão escolar deverá ocorrer. E seu ponto de partida deverá ser uma agenda organizada em torno da ampliação e da diversificação dos usos das avaliações externas, com uma aplicação mais intensiva das avaliações formativas. Articuladas com essas mudanças no uso

³ Uma boa evidência desse tipo de política voltada para a transformação da cultura docente é a experiência do Programa de Alfabetização na Idade Certa, desenvolvida no Ceará. Para conhecer este programa e seu impacto no comportamento docente, ver Rossi (2010).

da avaliação externa e, em parte, em decorrência delas, outras mudanças na gestão deverão se aprofundar, entre elas: a produção e o uso de evidências oriundas da própria escola e do aprimoramento dos sistemas de gestão de dados administrativos da rotina escolar; o fortalecimento das instâncias regionais, como parte fundamental da inteligência de uma rede escolar concebida como uma engrenagem orientada para a cooperação e a formação de comunidades de aprendizagem; o investimento consequente e responsivo na formação continuada de professores e gestores; e, não menos importante, a valorização crescente da participação discente. Tudo isso tende a concorrer para a conformação de um cenário de inovações nas áreas da gestão e na produção e apropriação de tecnologias sociais. E também para a conformação de uma cultura de responsabilidade profissional em torno dos resultados escolares. Tais tendências, já insinuadas nas diferentes redes educacionais brasileiras, apontam, no final das contas, para um processo de ultrapassagem da onda gerencialista, mas sem prejuízo de algumas de suas importantes conquistas, que têm a ver com a valorização das metas e resultados, e dos mecanismos de avaliação externa e monitoramento.

Quanto a isso, talvez nada reflita melhor essa evolução do que a trajetória da avaliação educacional como parte cada vez mais indispensável à gestão escolar. De ferramenta considerada hostil à escola, por pretensamente mobilizar no ambiente escolar uma lógica que lhe parecia estranha, muitas vezes percebida como aliada da meritocracia e do economicismo, o uso da avaliação externa passa, gradualmente, a ser compreendido como um instrumento a serviço da tomada de decisão com base em evidências. Nesse sentido, ganha crescente importância como elemento que confere legitimidade a ações tanto no âmbito da escola quanto no da rede de ensino. O que se constata, portanto, é um processo de superação, que pode ser mais rápido ou mais moroso dependendo do caso, da desconfiança inicial acerca de seu uso, e uma afirmação de sua condição de aliada fundamental do trabalho escolar. E é especialmente interessante observar como, apesar da instabilidade de muitas políticas educacionais, a produção e o uso da avaliação externa no Brasil já contam com um notável grau de constância e de aprimoramento nas diferentes redes de ensino.

Disso se segue que, se a aceitação do uso da avaliação externa por parte dos atores escolares, principalmente no que diz respeito aos professores, já foi um problema de maior vulto, agora, um novo conjunto de desafios levanta-se diante da difusão da avaliação, desta vez, de natureza mais técnica. A superação da resistência política quanto ao instrumento vem dando lugar a dúvidas em relação a como se usar, de maneira produtiva e eficiente, os resultados

que ele gera, de modo a realmente produzir efeitos positivos no trabalho escolar, em especial na melhoria da aprendizagem.

Como a mera divulgação de resultados tornou-se insuficiente, apesar de necessária, para promover a almejada mudança na qualidade da educação com o apoio dos instrumentos de avaliação, a apropriação dos resultados passou a exigir uma contínua capacitação dos atores escolares para o trabalho com essas informações. Essa transição foi ressaltada por Alícia Bonamino (2013), quando indica as diferentes gerações de avaliação com base na forma como a divulgação dos resultados vem se dando, permitindo seu uso mais amplo ou mais restrito. Para ela, estamos “aprendendo a valorizar as evidências e os dados para orientar ações de melhoria da qualidade do ensino” (BONAMINO, 2013, p. 57), e estaria se afirmando entre nós uma “perspectiva diagnóstica e formativa, com vistas a identificar os fatores sobre os quais os profissionais da escola precisam concentrar seus esforços de intervenção” (BONAMINO, 2013, p. 57-58). E esse esforço tem feito com que os sistemas próprios de avaliação desenvolvidos por diversas secretarias estaduais e municipais de educação venham sendo combinados com programas de desenvolvimento profissional para fins de apropriação de resultados.

Com objetivos pragmáticos, essas capacitações costumam ter como suporte a tríade: resultados-metas-planejamento. Com efeito, de posse dos resultados da avaliação, o diagnóstico sobre o trabalho da escola pode ser melhor formulado, tomando como referência o mapeamento das principais vicissitudes que incidem na aprendizagem dos alunos. Esse primeiro passo abre espaço para a organização do trabalho colaborativo na escola e para o estabelecimento de metas de aprendizagem de maneira coletiva, bem como para a definição de estratégias mais finas de se assegurar equidade no acesso à aprendizagem.

É verdade que a definição de metas ainda não é ponto pacífico no âmbito de todas as redes e escolas, em especial, por conta do vínculo ainda percebido por muitos como inevitável entre a definição de metas e seu impacto negativo sobre a autonomia escolar. A despeito disso, as metas têm sido tratadas como elementos caros ao planejamento escolar, vinculadas às chances de se promover um aumento da qualidade educacional. Com isso, o planejamento escolar passa a ser percebido como dimensão fundamental da construção dos objetivos escolares, concebidos coletivamente e suportados pelas evidências oferecidas pela avaliação.

De todo esse processo, em grande medida deflagrado pelo uso crescente da avaliação externa, um dos efeitos mais importantes é o da redefinição do papel da administração escolar. De figura burocrática, pouco comprometida

com a dimensão propriamente pedagógica, a administração escolar ganha renovado protagonismo na organização do processo de ensino e aprendizagem na escola, e isso se deve, em grande medida, ao fato de passar a pautar sua atuação por resultados da avaliação externa, somativa e formativa. Mas o vigor desse processo de articulação entre avaliação e gestão escolar também tem levado ao fortalecimento do uso de instrumentos de monitoramento e, muito especialmente, a uma profunda transformação da formação continuada, que passa a ser, ela mesma, cada vez mais lastreada e animada por resultados de avaliação.

AVALIAÇÃO, MONITORAMENTO E FORMAÇÃO CONTINUADA

A articulação entre avaliação e monitoramento tem especial relevância para países como o Brasil, onde a principal inovação a ser conquistada não é tanto a do avanço em novas tecnologias de informação ou mesmo na mudança da forma escolar, mas, principalmente, a da conquista da estruturação de uma rotina escolar, onde a aula efetivamente “aconteça” e onde a escola realmente faça a diferença, sobretudo, para estudantes oriundos de famílias pouco escolarizadas. A maior inovação entre nós seria, portanto, estruturar uma rotina escolar capaz de cumprir aquilo a que se propõe, de acordo com um determinado currículo, deixando de funcionar como uma máquina de exclusão de parcela de estudantes que, privados da aprendizagem básica, vão sendo sistematicamente reprovados até se tornar estudantes retardatários e, finalmente, evadidos.

Possivelmente por sua potência estratégica para contextos educacionais como o brasileiro, é que se vem observando no país um movimento até certo ponto mais vigoroso que nos países mais avançados, de articulação entre sistemas de avaliação e monitoramento. De antemão, pode-se observar que a relação entre avaliação e monitoramento reclama o desenvolvimento de novas formas de avaliação formativa, capazes de dar um retorno mais rápido e tempestivo aos profissionais da escola; e também anima o desenvolvimento de instrumentais de avaliação interna que ofereçam dados de monitoramento do processo de ensino e aprendizagem. De fato, um sistema mais constante e tempestivo de avaliação formativa se confunde com a lógica do monitoramento; ou se poderia dizer que, nesse caso, a avaliação é monitoramento. Reside aí, sem dúvida, um aprimoramento da tecnologia da avaliação, fazendo com que ela passe a ser experimentada como dispositivo pedagógico fundamen-

tal. Mas é claro que, para essa articulação entre avaliação e monitoramento acontecer, é necessário que se crie nas redes escolares e entre seus profissionais uma cultura de valorização dos dados e informações propiciados por esses instrumentos. Por seu turno, para que essa cultura possa se difundir, é necessário imaginar espaços institucionais onde isso possa ocorrer.

Abre-se aqui um amplo leque de alternativas. Na literatura internacional, como se viu, uma boa parte da energia em torno dessa questão tem sido canalizada para o tema da valorização da cultura profissional. A ideia básica é a de que, para além da forma sindical de vocalização dos interesses dos professores, seria necessário incentivá-los a participarem de espaços de troca profissional, por meio de comunidades de aprendizagem, de associações ou de fóruns profissionais. O fortalecimento do sentimento de responsabilidade profissional seria, portanto, uma via de transformação das tecnologias de avaliação e monitoramento em um subsídio valorizado pelo profissional preocupado com a qualidade de seu trabalho.

Esse tipo de agenda é, sem dúvida, muito desejável, mas ainda é pouco aplicável em contextos como o brasileiro, nos quais os professores de educação básica ainda se ressentem de melhores condições de trabalho, bem como de uma formação qualificada para fazer frente aos desafios de escolas que, muitas vezes, atendem a públicos muito heterogêneos, boa parte deles oriundos de famílias pouco ou nada escolarizadas.

Diante desse quadro, uma via que vem se mostrando bastante promissora em algumas redes escolares do país é a da articulação entre avaliação, monitoramento e formação continuada do professor⁴. Interessante como essa dinâmica também permite que se defina um novo e potente significado para a avaliação somativa, que deixa de ser utilizada basicamente como fonte de prestação de contas e de responsabilização dos profissionais pelos resultados, e passa a ser encarada como dimensão fundamental de um sistema de vasos comunicantes de produção de inferências sobre o trabalho educacional. Vai ficando claro que, isoladamente, uma avaliação somativa pode muito pouco no que se refere à melhoria escolar, mas como inferência mais agregada de um sistema de monitoramento e avaliação, articulada ao trabalho de formação continuada, responde pela função primordial de servir de referência para todas as demais atividades de gestão voltadas para a melhoria do desempenho dos estudantes, aí incluída a maior preocupação com a equidade do sistema educacional.

⁴ Sobre a potencialidade dessa engrenagem envolvendo avaliação, monitoramento e formação continuada, ver relatório de pesquisa realizada pelo CAEd/UFJF (2019).

Ainda que amparada pelo marco comum da Lei federal do Piso Salarial (Lei nº 11.738/2008) (BRASIL, 2008), que reserva 1/3 da jornada de trabalho do professor para capacitação e planejamento das aulas, a formação continuada tem sido implementada com intensidades e formas muito diversas por municípios e estados. Mesmo assim, pode-se afirmar que ela é melhor sucedida justamente naqueles contextos em que se compreendeu a formação como um momento de articulação entre sistemas de avaliação e monitoramento, a começar pelo fato de que a definição do plano de formação deve, a cada ano, ser retroalimentada por resultados de avaliação somativa e formativa. Mas para isso ocorrer, é necessário que se crie uma estrutura de gestão composta de profissionais capazes de fazer essa apropriação dos resultados e a sua tradução em material estruturante da formação.

Na medida em que é continuada, a formação somente se realiza se for capaz de lidar empiricamente com os problemas concretos da sala de aula, e se o formador se sentir corresponsável pelo resultado das escolas cujos professores ele forma. Note-se que, para isso, o formador tem que se converter ele próprio em supervisor da rotina escolar, atuando junto e em parceria com o professor. E para se fazer isso, precisará lançar mão de dados e de informações de monitoramento que orientem seu trabalho de formação. Uma dessas fontes de informação deverá ser a dos resultados de avaliação formativa e de instrumentais de monitoramento do resultado escolar.

Cria-se, com isso, uma articulação entre avaliação, monitoramento da rotina escolar e formação continuada, que dá lugar a uma engrenagem que conecta avaliação somativa, avaliação formativa, outras inferências do processo de ensino e aprendizagem, formação e supervisão da rotina escolar. O que a engrenagem que articula avaliação, monitoramento e formação permite entender, afinal, é que, na condição de mediador de conhecimento, o professor precisa estar permanentemente municiado por um sistema de supervisão que, mais do que controle e responsabilização, propicie suporte e apoio ao seu trabalho. E, para isso, não pode mais se dar ao luxo de dispensar dados de avaliação externa, somativa e formativa, tampouco dados de monitoramento do processo de seu trabalho.

É claro que, para essa roda girar, muitas premissas precisam ser cumpridas, de ordem política e administrativa, e também pedagógica. Mas ter em mente a necessidade de se conceber essa articulação pode ser sem dúvida um importante trunfo no sentido de se enfrentar obstáculos que aparecem como externos à escola e que, em boa medida, o são mesmo. Contudo, para serem superados, não podem prescindir da participação ativa da escola e de um sistema educacional estruturado.

À GUIA DE CONCLUSÃO: A FORMAÇÃO DOS NOVOS SUJEITOS DA EDUCAÇÃO

Um dos efeitos mais interessantes da difusão da avaliação externa é a sua relação com a emergência de novos sujeitos no universo profissional escolar. É claro que, para estudar mais a fundo esse processo na cena contemporânea brasileira, seria necessário incluir na análise o que chamamos, em outra oportunidade, de construção social do aluno, que diz respeito ao processo de transformação de crianças e famílias pobres e pouco escolarizadas em seres afeitos à lógica escolar (BURGOS, 2014). Mas tal empreendimento iria além da proposta deste artigo. Por isso, optamos por recortar nosso enfoque, privilegiando como objeto central de nossa análise a relação entre o uso das novas tecnologias de avaliação externa e as transformações ora em curso no mundo profissional da educação.

Neste artigo, procuramos salientar que a linguagem da avaliação passou a fazer parte da cultura profissional e, junto com ela, um tipo novo de expertise começou a ser valorizado. Assim, ler a educação a partir dos resultados da avaliação externa pressupõe pensá-la desde uma perspectiva sistêmica e comparada, a uma perspectiva analítica que permite redefinir as relações de causalidade existentes entre desempenho estudantil e trabalho escolar.

A decantação da linguagem da avaliação encontra seu principal difusor na figura dos técnicos, inspetores, superintendentes etc., lotados nas secretarias de educação e em suas respectivas instâncias regionais. São eles os novíssimos sujeitos da educação, porta-vozes dessa nova forma de encarar o processo educacional. Muitos deles ingressaram nas repartições das secretarias de educação com a missão de difundir essa nova gramática, outros tiveram que se adequar a ela, para permanecerem em seus cargos. O fato é que se criou no Brasil, em um espaço de duas décadas, um novo tipo de profissional, fiador da leitura sistêmica do trabalho escolar, bem como da valorização dos resultados escolares.

Mas esse processo de decantação da linguagem da avaliação externa já alcançou o professor, produzindo mudanças que apontam para, pelo menos, duas dimensões. A primeira, talvez a menos visível, é a da repercussão sobre o trabalho docente propriamente dito, afinal, faz muita diferença planejar as aulas, organizar atividades transversais, investir mais em determinados aspectos do curso, ou mesmo conferir maior atenção a determinados estudantes, tendo como parâmetro os resultados de avaliação externa. Esse tipo de mu-

dança, que já impacta a formação docente em serviço, tende, cada vez mais, a incidir sobre as licenciaturas, ainda que as universidades nem sempre estejam prevenidas para enfrentá-la.

A segunda dimensão tem a ver com a cultura profissional docente. A produção de evidências sistêmicas sobre os resultados escolares contribui para o desenvolvimento de uma leitura que, fatalmente, desloca o professor de seu lugar tradicional, fortemente centrado na sala de aula. A tendência é a de que a concepção mais individualista e atomizada do trabalho docente dê lugar a novas formas de cooperação e de sentimento de cumplicidade coletiva pelo resultado escolar. Na medida em que impacta o trabalho de avaliação interna e a liberdade de que o professor gozava para realizá-la, o impacto da difusão da gramática da avaliação externa sobre o *ethos* docente tende a reduzir sua autonomia individual, ao mesmo tempo em que amplia, exponencialmente, seu poder coletivo de atuação. E se no início os professores, naturalmente, sentiram essa mutação como perda de poder e de prestígio, a afirmação de um *ethos* decididamente coletivo tende a se traduzir em novas formas de empoderamento. Não é este o lugar para aprofundar essa questão, mas não custa notar que tal mudança afeta, a um só tempo, o lugar dos sindicatos na vida docente, o trabalho das universidades na formação de professores e as políticas públicas que precisarão, cada vez mais, desenhar formas de incentivar cooperativas de desenvolvimento do trabalho escolar. Bem a propósito, cabe salientar que o modelo competitivo fartamente adotado pelo receituário gerencialista, como forma de incentivar a melhoria dos resultados escolares, mostra-se agora incompatível com as novas formas de organização de um trabalho escolar cada vez mais animado pela tópica da cooperação.

Neste artigo, procuramos argumentar que está em curso no país a afirmação de um novo *medium* linguístico no senso comum escolar e que a conformação desse novo ambiente está intrinsecamente associada a um processo de renovada valorização profissional tanto do professor como das demais funções das estruturas administrativas das redes educacionais. Mas talvez pudéssemos ir além, sustentando que dessa configuração segue uma nova lógica de reprodução institucional, até certo ponto imune às mudanças e descontinuidades políticas. Afinal, os novos sujeitos da educação, com sua gramática própria, vieram para ficar e se tornaram, eles próprios, defensores do novo modelo que encarnam: o de uma escola ligada aos resultados e ciosa de sua responsabilidade profissional, mas também ciente da importância do trabalho coletivo, cooperativo e sistêmico.

Em meio a um ambiente político como o que se vive nesse momento no Brasil e em outras partes do planeta, marcado por ataques puramente ideológicos à educação pública e pela tentativa de se substituir a preocupação com a efetividade das políticas educacionais pelos temas da moralidade, esses novos sujeitos da educação, forjados na linguagem da avaliação externa, são certamente os principais fiadores da continuidade do projeto estabelecido pela Constituição de 1988 para a educação escolar.

REFERÊNCIAS

APPLE, Michael; BEANE, James. **Escolas Democráticas**. Editora Cortez: São Paulo, 2001.

BONAMINO, Alícia. Avaliação Educacional no Brasil 25 anos depois: onde estamos?. In: BAUER, Adriana; GATTI, Bernadete (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: Editora Insular, 2013.

BROOKE, Nigel (org.). **Marcos Históricos na reforma da educação**. Editora Fino Traço: Belo Horizonte, 2012.

BURGOS, Marcelo Baumann; BELLATO, Caíque. Gerencialismo e pós-gerencialismo: em busca de uma nova imaginação para as políticas educacionais no Brasil. **Revista Sociologia & Antropologia**, Rio de Janeiro, v.9, p. 919–943, set./dez., 2019.

BURGOS, Marcelo (coord.). **A escola e o mundo do aluno**. Rio de Janeiro: Garamond, 2014.

BURGOS, Marcelo; CANEGAL, Carolina. Diretores Escolares em um contexto de reforma da educação. **Revista Pesquisa e Debate em Educação**, v.1, n.1. Disponível em: [file:///C:/Users/C.sociais/Downloads/2-6-1-PB%20\(1\).pdf](file:///C:/Users/C.sociais/Downloads/2-6-1-PB%20(1).pdf) . Acesso em: 21 ago. 2020.

CAEd/UFJF. **Para Alfabetizar na idade certa**: protocolos para a organização de uma política de formação continuada e de incentivo ao compromisso com os resultados para professores alfabetizadores. Juiz de Fora: CAEd/UFJF, 2019. Disponível em: <http://www.mestrado.caedufjf.net/wp-content/uploads/2019/11/PROTOCOLOS-ALFABETIZA%C3%87%C3%83O-NA-IDADE-CERTA-C05.pdf>. Acesso em: 21 ago. 2020.

DARLING-HAMMOND, Linda. Teaching and the change wars: The professionalism hypothesis In: A. Hargreaves; M. Fullan (Eds.), **Change wars**. Bloomington: Solution Tree, 2009, p. 45–68.

DUSI, Cristina Sayuri Cortes Ouchi. **Os efeitos da gestão para resultados na educação: uma análise das políticas públicas educacionais de sete estados brasileiros**. 2017. Tese (Doutorado em Ciências Sociais) – Faculdade de Ciências Sociais, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2017.

HABERMAS, Jürgen. **Direito e Democracia**: entre facticidade e validade. vol I. Rio de Janeiro, Tempo Brasileiro, 1997.

HARGREAVES, Andy. The fourth way of change: Towards an age of inspiration and sustainability. In: HARGREAVES, Andy; FULLAN, Michael (eds.) **Change wars**. Bloomington: Hawker Brownlow Education, 2009, p.11-44.

HOPKINS, David. Every School a Great School – Realising the Potential of System Leadership. In: HARGREAVES, Andy, et alli (eds.). **Second International Handbook of Educational Change**. vol. 23. Springer Science & Business Media, 2010, p. 741-764.

LAWN, Martin. Os professores e a fabricação de identidades. In: Antonio Novoa (org.) **A Difusão Mundial da Escola**. Editora Educa: Lisboa, 2000.

LIPSKY, Michael. **Street-level bureaucracy**: dilemmas of the individual in public services. NY: Russell Sage Foundation, 1980.

MACBEATH, John. Self-Evaluation for School Improvement. In: HARGREAVES, Andy et alli (eds.). **Second International Handbook of Educational Change**. vol. 23. Springer Science & Business Media, 2010, p. 901-912.

PEREGRINO, Mônica. **Trajetórias desiguais**: Um estudo sobre os processos de escolarização pública de jovens pobres. Rio de Janeiro: Ed. Garamond, 2010.

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

ROSSI, Jocelaine Regina Duarte.- **Entre o estável e o fortuito**: a formação continuada em serviço e as rotinas pedagógicas em alfabetização, 2010. Dissertação (Mestrado em Educação) –Faculdade de Educação, Programa de Pós-graduação em Educação, Universidade Federal do Ceará, Fortaleza, 2010.

TALBERT, Joan E. Professional Learning Communities at the Crossroads: How Systems Hinder or Engender Change. In: HARGREAVES, Andy et alli (eds.). **Second International Handbook of Educational Change**. vol. 23. Springer Science & Business Media, 2010, p. 555-572.

TEDDLIE, Charles. The Legacy of the School Effectiveness Research Tradition. In: HARGREAVES, Andy et alli (eds.). **Second International Handbook of Educational Change**. vol. 23. Springer Science & Business Media, 2010, p. 523-554.

Recebido em: 06 de dezembro de 2019

Aceito em: 19 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E *ACCOUNTABILITY*: UMA BREVE ANÁLISE DA EXPERIÊNCIA BRASILEIRA

Reynaldo Fernandes¹

Amaury Patrick Gremaud²

¹Professor Titular do Departamento de Economia – FEA/RP Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil.Contato: refernan@usp.br.

²Professor Doutor, Departamento de economia – FEA /RP Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil.Contato: agremaud@usp.br.

Resumo

O presente texto procura avaliar o movimento de avaliação educacional em larga escala e *accountability* que ocorreu no Brasil nas duas últimas décadas. No artigo, é feita uma breve revisão da literatura internacional sobre *accountability* educacional, destacando os argumentos em sua defesa, a posição dos críticos e a avaliação de seus resultados. Em seguida, é apresentado um histórico do movimento de avaliação em larga escala e *accountability* no Brasil e, ao final, é realizada uma avaliação desse movimento, abordando três questões: a) até onde a melhoria dos dados educacionais brasileiros no ensino fundamental refletem uma melhoria no aprendizado ou reflete uma “inflação de notas?; b) até que ponto os programas de avaliação e *accountability* contribuíram para esta melhoria? e c) Por que a melhoria no ensino fundamental não atingiu o ensino médio da mesma forma?

Palavras-chaves: Avaliação educacional. Accountability. Educação básica. SAEB. IDEB.

Abstract

This paper seeks to evaluate the large-scale educational assessment and accountability movement that has occurred in Brazil over the past two decades. In the article a brief review of the international literature on educational accountability is made, highlighting the arguments in its defense, the position of the critics and the evaluation of its results. This is followed by a review of the large-scale assessment and accountability movement in Brazil. At the end, an evaluation of this movement is conducted, addressing three issues: a) to what extent does the improvement of Brazilian educational data in fundamental education reflect an improvement in learning or reflect a “score inflation”? b) to what extent have evaluation and accountability programs contributed to this improvement? and c) why did the improvement in primary education not reach high school in the same way?

Keywords: Educational evaluation. Accountability. Basic education. SAEB. IDEB.

INTRODUÇÃO

Enquanto as primeiras iniciativas voltadas à implantação e desenvolvimento do Sistema de Avaliação da Educação Básica (SAEB) datam do final dos anos 80 e que ocorreram duas aplicações anteriores (1990 e 1993), podemos considerar o ano de 1995 como o ano de institucionalização definitiva do SAEB. A partir dessa data, o SAEB vem divulgando resultados de leitura e matemática para uma amostra de estudantes do final de cada uma das etapas do ensino básico. Em conjunto com as informações de movimentação e fluxo escolar, extraídas do censo da educação básica, o SAEB permitiu a realização de um detalhado diagnóstico da qualidade da educação ofertada no Brasil e em cada uma de suas unidades federativas.

Nos anos 2000, a avaliação educacional em larga escala no Brasil passou a incorporar a noção de *accountability*. A criação da Prova Brasil em 2005, a divulgação do Exame Nacional do Ensino Médio (ENEM) por escolas em 2006 e o lançamento do Índice de Desenvolvimento da Educação Básica (IDEB) em 2007, são marcos dessa nova orientação. Os resultados passaram a ser divulgados não apenas para o país e unidades da federação, mas também por redes de ensino e escolas individuais. Esse movimento de avaliação e *accountability* não ficou restrito ao governo federal. Hoje, diversos estados e alguns municípios possuem sistemas próprios de avaliação e programas de *accountability*, incluindo bônus para professores com base no desempenho dos alunos nos exames.

Os programas de *accountability* educacional costumam gerar polêmica, seja no Brasil ou no exterior. Seus defensores destacam as dificuldades existentes para monitorar o trabalho da escola (por pais, autoridades e a sociedade em geral). Esse quadro facilitaria o surgimento de um problema típico de agente-principal, onde os interesses dos agentes (professores, diretores e gestores educacionais) podem não estar totalmente alinhados com os interesses dos principais (governantes, eleitores, estudantes e seus familiares). Assim, os exames externos providenciariam informações independentes às autoridades e ao público sobre o desempenho das escolas em disciplinas chaves. Isso, associado a outros incentivos para que os alunos desempenhem bem nos exames, encorajaria os educadores a se concentrarem em tarefas que elevem o desempenho dos estudantes nos exames, potencialmente alterando o método de ensino. Por sua vez, seus opositores sustentam que tais programas, além de não terem impactos comprovados na melhora do aprendizado, distorcem os incentivos das escolas. Duas das principais preocupações dizem

respeito à exclusão dos estudantes com mais dificuldades de aprendizado e ao estreitamento do currículo (concentrar o ensino naquilo que é pedido nos exames externos).

Ao completarmos 25 anos da institucionalização do SAEB, parece ser uma boa oportunidade para avaliar esse movimento de avaliação educacional em larga escala e *accountability*. O presente texto pretende caminhar nessa direção. O restante do artigo está organizado em três seções, além da conclusão. Na próxima seção, realizamos uma breve revisão da literatura internacional sobre *accountability* educacional, destacando os argumentos em sua defesa, a posição dos críticos e, principalmente, a avaliação de seus resultados. A seção III traz um histórico do movimento de avaliação em larga escala e *accountability* no Brasil. Por fim, na seção IV, faz-se uma avaliação desse movimento no Brasil.

ACCOUNTABILITY EDUCACIONAL: SUA RACIONALIDADE, A POSIÇÃO DOS CRÍTICOS E AS EVIDÊNCIAS DISPONÍVEIS DE SEUS RESULTADOS.

O uso de avaliações educacionais em larga escala com objetivo de proporcionar estudos sobre aprendizagem e para monitorar sistemas educacionais não é algo novo. Um marco no uso de avaliações em larga escala no estudo da aprendizagem é dado pelo relatório Coleman (COLEMAN et al., 1966), que tinha como objetivo estudar a segregação racial no sistema educacional dos EUA. Esse trabalho é importante por uma série de razões, entre as quais a mudança de foco na forma de identificar a qualidade das escolas: de insumos e processos para resultados. Para efeitos de monitoração do sistema educacional, o National Assessment of Educational Progress (NAEP) dos Estados Unidos surge em 1969. Para essas duas finalidades, as avaliações tinham base amostral (cujos resultados não eram apresentados por escolas, professores ou estudantes) e não costumavam dar origem a grandes polêmicas, ainda que alguns questionamentos já estivessem presentes e incidiam, principalmente, sobre a limitação dos instrumentos (questionários, testes e matrizes) em captar o trabalho que se realizava nas escolas e sobre os processos infra escolares.

É só a partir do final dos anos 80 que as avaliações com fins de *accountability* ganham corpo, tendo como marco a reforma educacional inglesa de 1988. Ao incorporar a função de *accountability*, as avaliações em larga escala passam a ser universais e dão origem a uma grande polêmica sobre sua validade. Apesar da polêmica, as avaliações educacionais universais para fins de *accountability* se espalharam ao redor do mundo. Hoje, a maioria dos principais

países desenvolvidos e muitos países em desenvolvimento possuem sistemas universais de avaliação que impõem algum grau de *accountability* sobre seus sistemas educacionais, escolas e mesmo professores individuais.

Por *accountability* educacional, entende-se o processo de avaliar o desempenho de sistemas educacionais, escolas e professores individuais com base em medidas de desempenho dos alunos. Seguindo Hanushek e Raymond (2005), vamos dividir os programas de *accountability* em dois tipos: i) aqueles que se limitam a divulgar os resultados dos estudantes por escolas e sistema educacional, a “*accountability* fraca” e ii) aqueles que, além disso, atrelam prêmios, sanções e assistência a tais resultados, a “*accountability* forte”. Enquanto nos programas do segundo tipo, as recompensas e sanções são explícitas – como bônus para professores em escolas consideradas excelentes ou ameaças de reestruturação ou fechamento de escolas de baixo desempenho – , nos programas do primeiro tipo, elas estão implícitas – operando menos por ação direta dos gestores do programa e mais por pressão da comunidade.

Os programas de *accountability* educacional têm por objetivo mudar a estrutura de incentivos para que professores, escolas ou sistemas educacionais proporcionem um melhor aprendizado aos seus estudantes. A racionalidade desses programas tem como base o problema do agente-principal, em que os interesses dos agentes (professores, diretores e gestores educacionais) não estariam totalmente alinhados com os interesses dos principais (pais, autoridades e a sociedade como um todo).³ Nesse quadro e na presença de informação imperfeita (assimétrica e/ou incompleta), os educadores poderiam se comportar de uma maneira que não esteja totalmente de acordo com os interesses dos estudantes e/ou de seus responsáveis.⁴ Admite-se que os responsáveis pela oferta de educação (professores, diretores e gestores) podem alterar suas condutas e, assim, proporcionar aos estudantes um melhor ensino. Entretanto, tais mudanças são consideradas custosas e, por esse motivo, os educadores não as efetivam. Então, ao providenciar informações independentes às autoridades e ao público em geral sobre o desempenho das escolas em disciplinas chaves, os programas de *accountability* atuariam no sentido de promover incentivos para que os educadores se concentrem em tarefas

³ Por exemplo, os educadores podem preferir gastar recursos em ações não diretamente associadas ao ensino (tornar o ambiente mais aprazível para eles ou usar da estrutura da escola para execução de serviços pessoais), podem conceder “benefícios” a eles mesmos, que acabam por comprometer a aprendizagem dos estudantes (faltar sem ter que repor as aulas ou focar o ensino na parte do currículo que mais gostam, ao invés daquelas mais importantes para os estudantes) etc.

⁴ Para uma discussão mais detalhada sobre *accountability* e incentivos baseados em avaliações, ver, por exemplo, Fernandes e Gremaud (2009), Figlio e Loeb (2011) e Hout e Elliott (2011).

que elevam o desempenho dos estudantes nos exames. Se a simples divulgação dos resultados de desempenho não for considerada suficiente, pode-se atrelar recompensas e sanções às escolas com base no desempenho de seus estudantes (*accountability* forte).

De modo geral, os programas de *accountability* assumem que é do interesse dos estudantes e/ou de seus responsáveis que as escolas concentrem seus esforços no ensino de algumas disciplinas chaves, cujo aprendizado pode ser aferido por avaliações em larga escala. Evidentemente, quanto mais alinhado com esse objetivo estiver o interesse dos educadores, menos efetivo tenderá a ser o programa de *accountability*. Assim, alguém que considere que os professores já fazem o máximo para proporcionar aos estudantes o melhor aprendizado, tenderia a considerar que qualquer política de incentivos para eles seria, no mínimo, inócua. Por outro lado, aqueles que acreditam que a qualidade da educação pode melhorar implicitamente admitem que alguns dos responsáveis pela educação (professores, diretores, gestores de rede ou governantes) podem fazer algo diferente do que vêm fazendo. Nessa perspectiva, em algum nível (professores, escolas ou sistemas de ensino), a *accountability* pode ser necessária.

É importante destacar que diferentes desenhos de programas fornecem diferentes estruturas de incentivos. Por exemplo, se a medida *accountability* utilizada for a proporção de estudantes considerados proficientes (aqueles com pontuação acima de determinado nível), o incentivo fornecido para as escolas é que elas se concentrem naqueles estudantes logo abaixo do nível de proficiência, retirando atenção daqueles que já ultrapassam esse nível e daqueles com desempenho muito baixo (considerados com baixa probabilidade de ultrapassar o nível fixado). Por sua vez, se a medida de *accountability* for a pontuação média da escola, o incentivo é para as escolas se preocuparem com todos os estudantes.

Os críticos dos programas de *accountability* levantam uma série de pontos, entre os quais destacam-se: 1) os programas são incompletos, pois não consideram todos os resultados importantes das escolas; 2) suas medidas de aprendizagem são imprecisas; 3) são injustos, ao responsabilizar os educadores por aspectos sobre os quais eles não possuem total controle; e 4) podem gerar distorções como o estreitamento curricular e a exclusão de alunos com maiores dificuldades de aprendizado.

Em relação ao primeiro ponto, é difícil discordar da alegação de que os objetivos dos programas existentes são limitados, frente aos múltiplos objetivos que podemos atribuir às escolas. A questão fundamental, no entanto, é sa-

ber se podemos considerar correto sinalizar para as escolas que priorizem suas ações no aprendizado de determinadas disciplinas consideradas chaves. Evidentemente, não há uma resposta óbvia para essa questão, a qual pode variar entre diferentes programas. Programas com objetivos mais estreitos podem incentivar as escolas a tirar o foco de aspectos importantes do ensino, enquanto programas com objetivos muito amplos podem ter dificuldade de obter medidas confiáveis de todos eles, além de proporcionar uma sinalização confusa sobre quais deveriam ser os principais objetivos das escolas. A escolha das medidas de desempenho é uma das decisões mais importantes do desenho do programa.

A segunda crítica é a de que os resultados das avaliações são uma medida imprecisa do aprendizado dos estudantes na disciplina considerada. De fato, o resultado dos estudantes nas provas não depende apenas da aprendizagem. Depende também da motivação e preparação específica para realizar o exame; das condições da aplicação; da sorte etc. Kane e Staiger (2002) mostram que os resultados de exames padronizados são medidas sujeitas a muito ruído, particularmente entre as pequenas escolas. A volatilidade das escolas no ranking de desempenho pode desacreditar o indicador utilizado. Mas, é possível adotar procedimentos para minorar esse problema como, por exemplo, usar a média das últimas edições do exame ao invés dos resultados de uma única edição, ou adotar um índice composto que agrega mais de uma medida de resultado. Uma preocupação na avaliação dos programas de *accountability* é que sua implantação pode levar as escolas a adotar medidas que elevam a pontuação nos exames sem que a aprendizagem seja afetada como, por exemplo, motivar e treinar os estudantes para o teste. Um fenômeno conhecido como inflação de notas (*score inflation*).

Quanto ao ponto três, é verdade que os resultados dos exames padronizados incorporam, além do esforço da escola e de seus professores, influências advindas da família, dos amigos e das habilidades inatas dos estudantes, bem como do erro aleatório de medida. Entretanto, isso não é necessariamente um problema em um programa de “*accountability* fraca”, limitado à ampla divulgação dos resultados. O público interessado pode “extrair o sinal de qualidade” de uma escola, por exemplo, por comparar os resultados de mais de uma edição dos exames com o de escolas próximas e/ou que possuem público similar. Esse não é o caso, no entanto, para os programas de “*accountability* forte”. Para esses, a questão da medida de desempenho é um elemento sensível, já que as premiações e/ou punições são automaticamente atreladas a ela. Nesse caso, seria necessário adotar alguma medida de valor adicionado.⁵

⁵ Para uma discussão sobre Modelos de Valor Adicionado ver, entre outros, Reardon e Raudenbush (2009).

Por fim, os programas podem promover distorção de incentivos, mas existem formas de, se não as eliminar, reduzi-las. Se os objetivos dos programas são adequados (concentram-se no que é prioritário), o estreitamento do currículo não vem a ser um problema. Por outro lado, pode haver incentivo para a exclusão de alunos de baixa proficiência. Assim, os programas devem procurar incluir formas de penalizar a exclusão de alunos com baixa proficiência.

Existe hoje uma importante literatura avaliando o impacto de diversos programas de *accountability* implementados nos Estados Unidos e em outros países. Duas revisões dessa extensa literatura são Figlio e Loeb (2011) e Hout e Elliott (2011). As duas revisões apontam que a resposta das escolas aos programas de *accountability* podem ir tanto na direção intencionada pelo programa, quanto em adotar ações que visam a elevar o desempenho nos exames sem a correspondente melhora no aprendizado. Por exemplo, Hout e Elliott (2011, p. 62-63) apontam:

Além das mudanças no ensino de determinado assunto, há evidências de tentativas de aumentar as pontuações de maneiras completamente não relacionadas à melhoria do aprendizado. As tentativas incluíram o ensino de habilidades para a realização dos testes, a exclusão de alunos de baixo desempenho dos testes, a alimentação dos alunos com refeições de alto teor calórico nos dias de teste, o fornecimento de ajuda aos alunos durante um teste e até a alteração das respostas dos alunos após a conclusão do teste.

Os dois estudos também encontram evidências de inflação de notas (*score inflation*), no sentido que o impacto positivo dos programas nos exames utilizados pelo programa (*high-stakes test*) tendem a ser maior do que nos exames não utilizados pelo programa (*low-stakes test*). Quanto a avaliação dos programas em elevar a aprendizagem nas disciplinas consideradas, as revisões chegam a conclusões diversas. Figlio e Loeb (2011) chegam a uma conclusão mais positiva dos programas, enquanto Hout e Elliott (2011) têm uma posição mais negativa. Isso é interessante, uma vez uma parcela importante dos artigos revisados é comum às duas revisões.⁶ Figlio e Loeb (2011, p. 383) concluem que “A preponderância de evidências sugere efeitos positivos, do movimento de *accountability* nos Estados Unidos durante os anos 90 e início dos anos 2000, sobre o desempenho dos alunos, especialmente em matemática”.

⁶ Enquanto Hout e Elliott (2011) consideram apenas programas com consequências explícitas, Figlio e Loeb (2011) consideram todos os tipos de programas (*accountability* “fraca” e “forte”). Hout e Elliott (2011) excluem também programas que se utilizam de regressões descontínuas (por avaliar o impacto de apenas uma parcela dos alunos da escola) e programas para os quais não foi possível obter uma medida de um teste não considerado para efeitos de *accountability* (*low-stakes test*).

Por sua vez, Hout e Elliott (2011, p. 92) concluem que “Apesar de usá-los por várias décadas, os formuladores de políticas e os educadores ainda não sabem como usar os incentivos baseados em testes para gerar consistentemente efeitos positivos no desempenho e melhorar a educação”.

Apesar da diferença nas conclusões, ao analisar objetivamente as duas revisões, vemos que elas apresentam um quadro muito parecido. Por exemplo, na tabela 4-1 da revisão de Hout e Elliott (2011), que resume os resultados dos artigos considerados, 17 impactos de programas com base em *low-stakes tests* são apresentados e os resultados são: 8 positivos, 7 não estatisticamente significativos e 2 negativos. Portanto, em sintonia com Figlio e Loeb (2011), os resultados são majoritariamente positivos. Entretanto, Hout e Elliott (2011) preferem destacar o pequeno valor das estimativas. A valor médio das estimativas por eles considerada é de 0,08 desvios padrão da distribuição de notas dos alunos do estado norte-americano (ou país) considerado. Eles também destacam que os impactos positivos se concentram em matemática e nas séries iniciais. Ainda que o uso do valor de 0,08 não esteja isento de críticas, ele dá uma dimensão dos impactos obtidos.⁷

Ao rever esses trabalhos, uma conclusão mais apropriada seria: as evidências disponíveis sugerem que, em média, os programas de *accountability* educacional apresentam efeitos positivos, mas modestos. Além disso, se concentram em matemática e nas séries iniciais. Para finalizar essa seção, seria importante destacar dois aspectos relacionados à conclusão acima. Primeiro, encontrar impactos modestos e concentrados nas séries iniciais em matemática não é uma exclusividade dos programas de *accountability*. Análises de impacto de reformas educacionais mais dispendiosas que *accountability* (aumento de salários dos professores, redução do tamanho da turma etc.) têm, na melhor das hipóteses, encontrado resultados modestos e, da mesma forma, concentrado em matemática e nas séries iniciais.⁸ Por fim, a utilização de um valor médio das estimativas esconde variações não apenas entre programas, mas também entre escolas de um mesmo programa. Parece que, quando o programa é posto em funcionamento, algumas escolas respondem de modo

⁷Hanushek (2012), em texto bastante crítico a Hout e Elliott (2011), contesta o uso do valor de 0,08. Ele incluiria estudos que não encontram os critérios definidos para inclusão na revisão, além de tirar o foco de programas que apresentam um excelente desempenho.

⁸Por exemplo, Aos e Pennucci (2003) revisam 53 estudos que avaliam o impacto da redução no tamanho da turma. Em todos os casos, a elasticidade do desempenho em relação ao tamanho da classe, em termos absolutos, é menor que 0,15. A maioria dos estudos na revisão mediu os resultados dos alunos (mudança de um ano) com pontuações padronizadas em testes; alguns examinaram as taxas de conclusão do ensino médio. Isso significa que reduzir o tamanho de sala de 30 para 27 alunos eleva o aprendizado anual médio dos estudantes em, no máximo, 1,5%. O efeito positivo de diminuir o tamanho das turmas é mais forte nas séries iniciais.

significativo, enquanto outras não dão qualquer resposta. Se for esse o caso, seria interessante investigar o que essas escolas têm de diferente. O que faz com que umas respondam aos incentivos e outras não.

UM BREVE HISTÓRICO DA AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E ACCOUNTABILITY NO BRASIL⁹

Como visto, é praxe considerar como marco inicial para os processos de avaliação educacional externa, o Relatório Coleman, com testes de desempenho aplicados a mais de 650 mil estudantes e de “surveys” aplicados aos próprios estudantes, pais, professores e diretores de escola, buscando levantar características do contexto e do processo educativo. Junto com este, o relatório Plo-wden na Inglaterra e a criação do NAEP fazem do final dos anos 60 um marco internacional das avaliações diagnósticas externas das redes educativas.

Nas décadas de 60 e 70, não se observam, no Brasil, levantamentos como estes. Boa parte das pesquisas e dos debates aqui travados diziam respeito a questões acerca dos fluxos educacionais: entrada, abandono e reprovações dos estudantes nas escolas e sistemas educativos. Nesse período, as questões de acesso à escola eram graves e preocupações sobre quantos e que tipo de alunos avançavam dentro do sistema estavam na base das pesquisas. Na década de 70, as pesquisas também foram muito influenciadas pela hipótese atrelada à teoria do capital humano de que a concentração de renda brasileira era consequência, pelo menos em parte, do baixo nível educacional da população brasileira. Essas pesquisas se baseavam nas informações acerca do (não) acesso ao sistema educacional e na (baixa) quantidade de séries que os diferentes grupos da população brasileira alcançavam.¹⁰

Nessa mesma década, os problemas da evasão e/ou abandono e da retenção e/ou reprovação também eram pontos importantes do debate dando sustentação a iniciativas, ainda na década de 70, de políticas de progressão continuada. Essas discussões avançaram pelos anos 80 junto com o próprio crescimento do número de jovens que adentravam no sistema educacional brasileiro. Se, por um lado, se observa, ao longo da década de 80, fortes

⁹ Para uma discussão das avaliações externas e do processo de institucionalização do SAEB, ver Bonamino e Franco (1999), Castro (2016), Horta Neto (2007) e Pestana (2016).

¹⁰ Segundo Horta Neto (2007), as primeiras medições da educação brasileira fizeram parte do *Anuário Estatístico Brasileiro*, produzido a partir de 1906, concentrando informações principalmente do Distrito Federal sobre número de escolas, de pessoal docente, matrículas e repetências. Estes dados foram interrompidos em 1918 e retomados, agora nacionalmente, em 1936.

melhorias nos dados de acessibilidade para as crianças nos anos iniciais do ensino fundamental, o abandono e a repetência continuavam a estar no centro das controvérsias e acabaram por trazer diversos questionamentos aos levantamentos empíricos realizados sobre estes fluxos educacionais e suas interpretações. Por exemplo, os alunos que deixavam de frequentar a escola e voltavam a se matricular na mesma série no ano subsequente eram considerados evadidos pelas estatísticas oficiais em vez de repetentes. Isso inflava as taxas de evasão e subestimava as taxas de repetência, o que ensejou a introdução de novos modelos para a interpretação dos dados como, por exemplo, o modelo Pro Fluxo.¹¹

Nos anos 80 surgem, junto com algumas pesquisas de cunho etnográficos, as primeiras pesquisas de rendimento escolares (testes cognitivos) atrelados a fatores associados (questionários e observações contextuais). Tais avaliações foram realizadas dentro do Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro (EDURURAL), que tinha por objetivos expandir o acesso à escola primária, diminuir as taxas de repetência e evasão e melhorar o rendimento escolar dos alunos. Fazia parte do programa – que contou com apoio do Banco Mundial, da Fundação Carlos Chagas, da Fundação Cearense de Pesquisa e da Universidade Federal do Ceará, além de pesquisadores estrangeiros¹² – avaliar os impactos dos investimentos realizados sobre os rendimentos dos alunos e sobre as taxas de reprovação. Esta investigação ocorreu em 1981, 1983 e 1985 e envolveu a aplicação de testes de Português e Matemática em aproximadamente 6.000 alunos da 2ª e 4ª séries do ensino fundamental de 600 escolas em 60 municípios dos estados do Ceará, Piauí e Pernambuco (HORTA NETO 2007; BONAMINO; FRANCO, 1999).

Interessante notar que, nessas pesquisas, juntam-se de modo explícito a questão dos fluxos, notadamente do abandono e da repetência, e dos problemas relativos à aprendizagem. As conclusões desses estudos, segundo Bonamino e Franco (1999, p. 105) foram:

Com relação ao aspecto da melhoria da performance dos alunos a partir de investimentos realizados, a pesquisa detectou resultado positivo apenas no Estado do Piauí. Mas quanto à melhoria da taxa de promoção, os resultados foram negativos nos três estados, o que sugere que a melhoria nas taxas de promoção é um objetivo ainda mais difícil de ser alcançado do que a melhoria do desempenho dos alunos. O resultado sugere que, no contexto das novas condições viabilizadas pelos investimentos,

¹¹ Ver Fletcher e Ribeiro (1989) e Klein e Ribeiro (1991).

¹² Ver Harbison e Hanushek (1992) e Gatti (1994).

os professores optam, ao menos em um primeiro momento, por aumentar seus padrões de exigência, o que redundará na tendência de manutenção, ou mesmo de elevação, das taxas de repetência. (...) o baixo rendimento curricular das crianças e o grande número de repetência decorrente de um conjunto de fatores, como baixos salários, influência política na designação de professores e infraestrutura curricular insuficiente, além de precariedades associadas às condições de vida dos alunos e suas famílias, em especial no que se refere às condições de saúde.

As bases dessas primeiras investigações com testes e *surveys* foram aproveitadas nos passos iniciais de institucionalização de um programa de avaliação educacional nacional. De modo a ampliar as informações disponíveis sobre a educação, incorporando, então, mais elementos em torno do processo de aprendizagem do alunado, a Fundação Carlos Chagas foi contratada, em 1987, pelo INEP para a realização do estudo “Avaliação do Rendimento dos Alunos de Escola do 1º Grau da Rede Pública: Um Estudo em 15 Capitais e 24 Cidades”. A própria fundação estendeu este estudo no ano seguinte, com o Governo do Estado do Paraná, e acrescentou 29 cidades ao estudo.

Pestana (2016, p. 74) destaca que o próprio Ministério da Educação nesse período sente:

necessidade de informações sobre as diversas realidades educacionais. Somente com base em um sistema de dados abrangente e robusto seria possível conhecer amplamente essas realidades e melhorar a capacidade de proposição e execução de políticas educacionais e de auditoria social. Os processos de avaliação educacional destacam-se então, como meio privilegiado de geração de informações do tipo requerido por essa forma de fazer política educacional.

Também em 1988, o MEC anunciou a criação do Sistema de Avaliação do Ensino Público de 1º Grau – SAEP. A pretensão era desenvolver mecanismos de avaliação de abrangência nacional, estendendo para todo o país os entendimentos que então se debatiam com o Banco Mundial, no interior do Projeto Nordeste, de promover avaliações semelhantes às realizadas no EDURURAL.¹³ Uma questão importante, na implementação desses testes, dizia respeito à falta de um currículo nacional a ser utilizado como base para a construção das matrizes de referência. Essas foram construídas com base em consultas aos professores e na verificação sobre o que efetivamente era ministrado nas sa-

¹³ A literatura, por exemplo Bonamino e Franco (1999), destaca nesse momento inicial o apoio de organismos internacionais como o Banco Mundial para a promoção destes mecanismos de avaliação. No caso do SAEP, Horta Neto (2007) aponta o financiamento do Instituto Interamericano de Cooperação para a Agricultura – IICA.

las de aula e contou com as duas experiências anteriores da Fundação Carlos Chagas. Ainda em 1988, pré-testes foram realizados, mas a primeira aplicação só ocorreu em 1990 envolvendo alunos das então 1ª, 3ª, 5ª e 7ª séries. Os dados foram publicados em 1992 já com uma mudança no nome do estudo que passou a adotar o nome de SAEB, sendo esta aplicação de 1990 considerado 1º ciclo do SAEB. Pretendia-se institucionalizar ciclos a cada dois anos, o que foi efetivamente realizado a partir do segundo ciclo que, no entanto, ocorreu apenas em 1993. Atualmente completamos o 15º ciclo avaliativo.

Esses ciclos avaliativos do SAEB podem ser divididos em três fases, os dois primeiros (1990 e 1993) compondo a fase de implementação do SAEB. Depois, a partir do terceiro ciclo de 1995, temos a segunda fase, de consolidação do SAEB como instrumento de diagnóstico e monitoramento da educação básica brasileira e, a partir de 2005, uma terceira fase marcada por acoplar ao SAEB elementos de *accountability* que, como visto, já tinham começado a se destacar nas avaliações internacionais no final dos anos 80.

Quanto aos dois ciclos iniciais destaca Pestana (2016, p. 75):

O arranjo institucional inicial do sistema de avaliação, em período de grande luta por espaços de atuação em todos os níveis, contava com a participação e a contribuição efetiva das administrações estaduais de educação em termos técnicos, operacionais e financeiros. Em menor grau, algumas administrações municipais, especialmente capitais que possuíam grandes redes de ensino, também participaram nessa fase inicial. Nessa interação, que previa divisões de responsabilidades, foi possível aprofundar o debate sobre o significado, os meios utilizados e as consequências do uso da avaliação como uma política de melhoria da qualidade da educação, além de aspectos técnicos do processo de avaliação. Como resultado, essa dinâmica auxiliou na formação de uma primeira leva de técnicos e especialistas aptos a operar o sistema que se desenhava. Em relação aos custos, os estados assumiram diversas atribuições, entre elas o levantamento de dados, aspecto decisivo e o que mais contribuiu para a institucionalização do sistema e, principalmente, para introdução e disseminação de uma cultura de avaliação no setor educacional brasileiro.

Ainda segundo Pestana (2016, p. 78):

Embora sejam inegáveis os ganhos técnicos obtidos com as modificações realizadas em 1995 (no método de análise dos testes e na metodologia de amostragem adotada), o mesmo não se pode afirmar em relação ao arranjo institucional do Saeb, que centralizou atribuições no MEC e diminuiu as atividades realizadas em parceria com estados e municípios.

Depois dos dois primeiros ciclos, em 1994, institucionalizou-se o SAEB como um sistema nacional de avaliação para diagnóstico e monitoramento. Algumas características foram introduzidas e se mantiveram pelos ciclos seguintes com poucas mudanças até 2005. Além da centralização, as principais modificações implementadas a partir do 3º ciclo de 1995 foram: a) a definição da aplicação dos testes nos alunos nas chamadas “series conclusivas”, ou seja, na 4ª e 8ª série do ensino fundamental (atualmente 5º e 9º anos do ensino fundamental), além da inclusão da 3ª série do ensino médio; b) a inclusão de escolas particulares nas amostras; c) a manutenção dos testes versando sobre habilidades e competências nas áreas de Português (leitura) e Matemática (resolução de problemas) – ainda que, em alguns ciclos, terem sido aventados e, por vezes, acrescidos testes em outras disciplinas; d) as avaliações com base em amostras complexas, representativas em diferentes estratos, especialmente com a possibilidade de resultados em termos estaduais e por dependência administrativa; e) a consolidação de instrumentos contextuais, fornecendo informações das características socioeconômicas, culturais e dos hábitos e práticas dos alunos – além dos aspectos infra escolares e das práticas dos docentes e gestores educacionais, que já eram colhidos nos primeiros ciclos; f) a consolidação do uso da teoria da resposta ao item (TRI) como metodologia de apuração dos resultados e definição das escalas (uma única escala por disciplina, envolvendo todas as séries).¹⁴

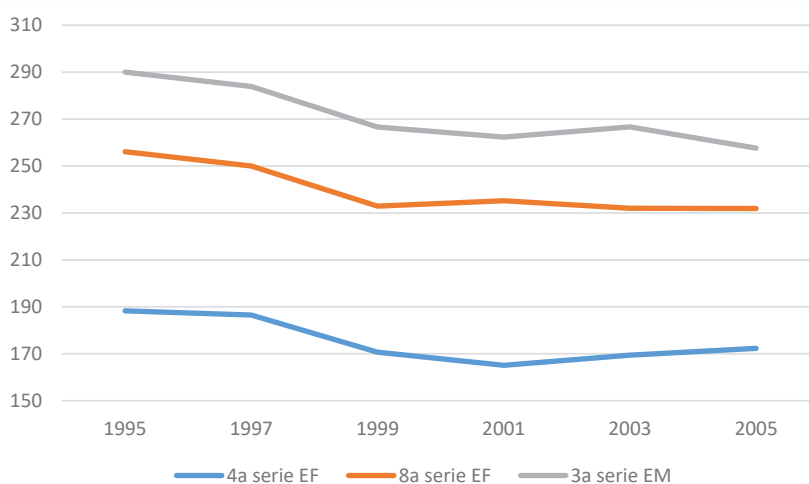
As duas últimas alterações acima permitiram importantes mudanças nas análises até então realizadas. Por um lado, a interpretação dos resultados da avaliação brasileira passou a dialogar com investigações internacionais, em que a preocupação sobre a influência das condições socioeconômicas e culturais sobre os resultados das proficiências dos alunos sempre foi objeto de atenção. No caso brasileiro, se confirma a importância dos aspectos socioeconômicos na explicação das diferenças de resultados entre os alunos, mesmo que o chamado “efeito escola” ainda tenha um peso bastante evidente e, nesse sentido, políticas educacionais tenham amplo espaço para melhorar a qualidade da educação nacional.

Por sua vez, o uso da TRI permite a comparação dos resultados ao longo do tempo, e a escala do SAEB passou a ser o termômetro da avaliação da qualidade brasileira. Ou seja, os ciclos avaliativos puderam ser comparados a partir de 1995 nos estratos amostrais semelhantes, mesmo que os testes aplicados ao longo dos anos sejam realizados com provas diferentes e sobre diferentes alunos. Os resultados dessas avaliações acabaram por colocar a questão da

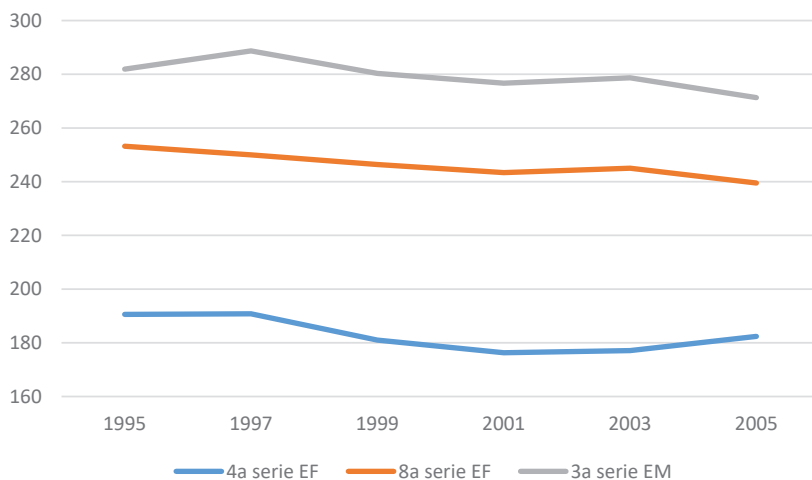
¹⁴ Ver Bonamino e Franco (1999).

aprendizagem nos debates educacionais ao lado das históricas discussões sobre fluxos (repetência, evasão etc.). Mesmo que existam indefinições sobre quais os resultados que deveriam efetivamente ser esperados para considerar, na escala do SAEB, a educação brasileira proficiente, adequada ou em um nível razoável, era claro que os indicadores mostravam que estes resultados não estavam sendo plenamente alcançados, o processo de aprendizagem deixava a desejar e, no período, estavam em queda. Pelos Gráficos 1 e 2 abaixo, esta comparação das médias de proficiências dos alunos brasileiros nas respectivas séries e disciplinas ao longo desta fase do SAEB (1995 -2005) mostra, de uma forma geral uma queda no desempenho. A própria queda das proficiências é em parte atribuída à ampliação dos fluxos, ao crescimento do acesso ao sistema, e a diminuição da qualidade é atribuída, em boa medida, à expansão do sistema e à dificuldade de se manter um padrão estável de qualidade.

Gráfico 1 – Brasil: a evolução das proficiências médias do SAEB de Língua Portuguesa



Fonte: dados básicos INEP (2006).

Gráfico 2 – Brasil: a evolução das proficiências médias do SAEB de Matemática

Fonte: dados básicos INEP (2006).

Ainda em relação à avaliação nesse período, três outros elementos podem ser destacados: a) o desenvolvimento de sistemas avaliativos em outros entes da federação brasileira, especialmente alguns Estados; b) a criação de outras estratégias nacionais avaliativas e; c) a participação do país em avaliações internacionais.

Nos primórdios do SAEB, como já observado, houve participação ativa de diversos Estados e alguns municípios no processo. Parte dessas iniciativas deu origem a sistemas estaduais de avaliações, alguns dos quais se mantiveram até hoje, enquanto outros sofreram descontinuidades. Estados como São Paulo, com o SARESP, Minas Gerais, com o SIMAVE e o Ceará, com o SPAECE, são exemplos de sistemas estaduais de avaliação que se iniciaram no final do século passado e se mantiveram ativos até hoje. Um aspecto importante é que tais sistemas estaduais possuem a mesma escala de proficiência do SAEB.

Nacionalmente, dois outros sistemas de avaliação foram criados: o Exame Nacional para a Certificação de Competência de Jovens e Adultos (ENCCEJA) e o Exame Nacional do Ensino Médio (ENEM). O primeiro se tornou um exame que possibilitava a certificação para os níveis fundamental e médio de ensino para jovens e adultos que não tiveram a oportunidade de concluir seus estudos no período ideal.

Quando criado, em 1998, o ENEM tinha como objetivo principal fornecer um autodiagnóstico aos estudantes que terminavam o ensino médio. O exame, de caráter voluntário, permitia a comparação dos resultados obtidos por um

participante com a média dos demais participantes na edição daquele ano. Um dos problemas metodológicos do ENEM era que ele, diferentemente do SAEB, não permitia a comparabilidade ao longo do tempo dos seus resultados, impossibilitando qualquer afirmação sobre a evolução de desempenho dos estudantes ao final do ensino médio. Ao longo dos anos, várias instituições de ensino superior passaram a se valer dos resultados do ENEM como critério, parcial ou exclusivo, de seleção para ingresso nos seus cursos universitários. Em 2005, o governo federal também passou a utilizar o ENEM como critério para a concessão de bolsas no âmbito do Programa Universidade Para Todos (ProUni). Nesse sentido, o ENEM acabou também cumprindo as funções de seleção e credenciamento. Tais elementos, junto com a gratuidade do exame para aqueles oriundos de escolas públicas, fizeram com que o ENEM, apesar de continuar sendo voluntário, tenha atingido mais de 3 milhões de inscritos em 2005 com forte crescimento da participação dos concluintes do ensino médio no exame, que passou a ter uma boa cobertura nesse segmento.

Neste período, o Brasil passou a participar também de avaliações internacionais, o MEC decide pela participação do Brasil nos estudos internacionais do Laboratório Latino-americano de Avaliação da Qualidade da Educação (LLECE). O LLECE foi criado em 1994 com outros 14 países latino americanos, sob a coordenação do Escritório Regional sobre Educação da UNESCO para a América Latina e o Caribe (OREALC/UNESCO). Em 1997, o Laboratório aplicou o Primeiro Estudo Regional Comparativo e Explicativo (PERCE) sobre o desempenho da aprendizagem entre os alunos da 3ª e 4ª séries do ensino fundamental, em leitura e matemática. Anos depois, em 2006, foi aplicado o segundo estudo e, em 2013, o terceiro estudo.

Em 2000, o Brasil foi convidado a participar do Programa Internacional de Avaliação de Alunos (PISA). O PISA é uma avaliação internacional (como o TIMSS, PIRLS, CIVICS) que se desenvolve sob a coordenação da OCDE e que, ao longo do tempo, ganhou destaque internacional. A pretensão do PISA é medir a capacidade dos jovens de 15 anos de usar seus conhecimentos e habilidades de leitura, matemática e ciência para enfrentar os desafios da vida real. No primeiro ciclo avaliativo, em 2000, o Brasil foi o único país não membro da OCDE a participar daquele ciclo e também é o único país não OCDE a participar de todos os ciclos que passaram a ocorrer de três em três anos.

Com os dados do PISA, fica claro que a aquisição e o desenvolvimento de habilidades e competências por parte dos jovens brasileiros, independente das diferenças de interpretação em torno da escala do SAEB, estava muito abaixo daquelas desenvolvidas na média dos países da OCDE. Outro elemento que

ganha destaque com a participação no Brasil no PISA é a volta da questão dos fluxos. O Brasil chama a atenção por ter ainda vários de seus jovens de 15 anos cursando a 6ª série do ensino fundamental, ao invés da 1ª série do ensino médio (ou, ao menos, a 8ª série do ensino fundamental) como seria o esperado.

A terceira fase do SAEB se inicia no ano de 2005. A normatização é alterada, e o SAEB passa a englobar um teste com aplicação amostral – inicialmente chamada de Avaliação Nacional da Educação Básica (ANEB) – e um teste com aplicação censitária – a Prova Brasil (oficialmente, Avaliação Nacional do Rendimento Escolar – ANRESC). A Prova Brasil foi aplicada pela primeira vez em 2005 e sua principal diferença em relação ao SAEB é que, dentro de seu universo de referência, que são os alunos das escolas públicas do 5º ano e do 9º ano do ensino fundamental, ela é censitária.¹⁵

No ano de 2005, o SAEB e a Prova Brasil foram realizados separadamente, com instrumentos diferentes (testes e questionários). Já em 2007, estas duas avaliações se fundiram e voltamos a ter uma única avaliação externa federal, em que os segmentos que compõem o universo da Prova Brasil foram avaliados censitariamente e seus resultados foram divulgados por escola, municípios, estados e por redes de ensino; enquanto que os outros segmentos que compõem o SAEB, mas não a Prova Brasil, continuaram a ser avaliados amostralmente. Esses outros segmentos são as escolas privadas e o 3º ano do ensino médio, para os quais não existem resultados divulgados por município ou por escola. Os ciclos avaliativos se mantiveram a cada dois anos. Assim, desde 2005, o Brasil passou a contar com o tradicional sistema de avaliação para diagnóstico e, também, com um programa de “*accountability* fraca”, por escolas e redes de ensino.

Ainda dentro desta perspectiva, uma mudança importante no ENEM foi introduzida em 2006. Como estudantes de praticamente todas as escolas do ensino médio participavam do exame que passava de 3 milhões de examinados, viu-se a possibilidade de divulgar os resultados do ENEM agregados por escola.¹⁶ Assim, além de autodiagnóstico e de credenciamento, passou a ser possível a utilização do ENEM como um instrumento de diagnóstico e de *accountability* para o ensino médio. Em 2009, o ENEM foi novamente modificado e nele se introduziu a TRI, permitindo sua comparabilidade ao longo do tempo,

¹⁵ A preocupação com a avaliação de outros níveis de ensino marcou também o período e, em termos federais, duas iniciativas foram introduzidas em relação ao processo de alfabetização: a Provinha Brasil e a Avaliação Nacional de Alfabetização (ANA).

¹⁶ Das mais de 24.250 escolas de ensino médio que constam do Censo Escolar, cerca de 23.000 tinham alunos inscritos no Enem em 2007.

e expandido os currículos avaliados, buscando, assim, diminuir os problemas de estreitamento de currículos (que para o ensino médio eram evidentes). Um dos principais objetivos dessa mudança foi o de atrair as Universidades Federais para utilizar o ENEM em seus processos seletivos. Em 2019, o ENEM teve mais de 5 milhões de inscritos, sendo que 77% desses realizam, de fato, o exame.

O ENEM possui algumas vantagens em relação ao SAEB para efeitos de *accountability*: vai além de matemática e leitura (avalia escrita, ciências naturais e ciências humanas), inclui escolas particulares e, por ser instrumento de seleção das universidades, os alunos tendem a realizá-lo com mais comprometimento. O fato de o ENEM ser de adesão voluntária não deveria ser impedimento para sua ampla divulgação, pois: (a) isso traz pouco impacto na ordenação das escolas e (b) a literatura especializada dispõe de diversos corretores de participação que poderiam ser usados.¹⁷

Apesar disso, o INEP anunciou que não mais divulgaria os resultados do ENEM por escolas e, a partir de 2017, mudou a Prova Brasil, que atualmente voltou a ser chamada de SAEB, a qual passou a ser universal para as escolas públicas também no 3º ano do ensino médio. Assim, a base do sistema de *accountability* no Brasil partir de 2017 é apenas o SAEB, que por enquanto, para o ensino médio, voltou a enfrentar um problema de estreitamento de currículo.

Uma outra questão dentro do processo de *accountability* é a ligação entre as questões de aprendizagem dos alunos e as tradicionais e sempre presentes questões de fluxos na educação brasileira. Como a proficiência em exames padronizados e o fluxo escolar não são independentes, restringir a cobrança aos resultados da Prova Brasil poderia incentivar os professores, diretores e gestores a adotarem medidas dentro das escolas que aumentassem tanto o desempenho médio dos estudantes nos testes padronizados quanto às reprovações, por exemplo, endurecendo dos critérios para aprovação. Assim, foi introduzida uma nova estatística educacional, o IDEB. Este surge com o objetivo de ancorar a *accountability* em um sistema de metas educacionais, sem que este estivesse baseado apenas nos resultados da Prova Brasil e, portanto, com o risco de contribuir para agravar o já dramático quadro de repetência e evasão escolar. Acerca do IDEB, podemos acompanhar a descrição feita por Fernandes (2016, p. 103) ¹⁸:

¹⁷ Para as escolas da amostra do SAEB 2011, a correlação entre as notas do SAEB e do ENEM foi de 0,87 e 0,91 para leitura e matemática, respectivamente. Para aplicação de corretores de participação no SAT dos Estados Unidos ver Dynarski (1987); Dynarski e Gleason (1993); Behrendt, Eisenach e Johnson (1986); e Clark, Rothstein e Schanzenbach (2009).

¹⁸ Para uma discussão sobre as propriedades do IDEB, ver Fernandes (2007).

O Ideb é obtido pela multiplicação da proficiência média dos alunos da escola (N) pela taxa média de aprovação da escola (P): $Ideb = NP$. Sob certas hipóteses ele pode ser interpretado como a razão entre a proficiência média dos alunos da escola (N) e o tempo médio que os alunos levam para concluir uma série (T): $Ideb = N/T$. Por exemplo, se, em média, os alunos precisam de dois anos para concluir uma série, o Ideb será igual à metade da proficiência média dos alunos da escola. Ele foi construído como forma de eliminar as reprovações improdutivas: reprovações que não contribuem para elevar o desempenho dos estudantes. Se as reprovações contribuírem para melhorar o desempenho dos estudantes da escola – seja porque incentivam os alunos a estudarem mais, seja porque tornam as turmas mais homogêneas –, a taxa ótima de reprovação seria diferente de zero, mas, provavelmente, muito baixa. Isso se confirmada a crença, de grande parte dos pesquisadores em educação, que reprovações são pouco produtivas.

Em 2007, com o objetivo de obter um maior comprometimento das redes e escolas com a melhoria da educação brasileira, foi pactuado, entre o Ministério da Educação e secretarias de educação de estados e municípios, um sistema de metas, e foi estabelecido o Plano de Metas Compromisso Todos pela Educação com base no IDEB. As metas do IDEB foram estipuladas para 2021, com metas intermediárias estabelecidas de dois em dois anos, a partir de 2007.

Para a meta de 2021, adotou-se um padrão externo: o desempenho educacional que, em média, era observado nos países da OCDE. Para cálculo do IDEB dos países da OCDE, supôs-se uma taxa de aprovação de 96%. A principal dificuldade foi fixar as notas, uma vez que os países da OCDE não fazem a Prova Brasil. Para tanto, admitiu-se que o PISA ordena os alunos da mesma forma que a Prova Brasil e, então, verificou-se que o percentil da distribuição de notas do Brasil no PISA era correspondente à média de desempenho dos países da OCDE. Encontrado esse percentil, obteve-se a nota correspondente a ele na Prova Brasil de 2005, a qual passou a ser referência para a meta do IDEB. Para definição de metas para redes de ensino e escolas individuais, considerou-se que todas as redes e escolas deveriam contribuir para que o país atingisse a meta estipulada, mas quem partisse de uma situação melhor no início também teria que obter melhores resultado ao final. Ainda que as diferenças de desempenho não fossem eliminadas, as metas consideram uma redução da desigualdade entre redes de ensino e escolas quando comparado com o ano base de 2005. A metodologia adotada considerou que a trajetória do IDEB ao longo do tempo, tanto para o Brasil como para os demais níveis de abrangência, segue o comportamento de uma função logística. Des-

sa forma, foi possível calcular o “esforço” que o Brasil e cada uma das redes e escolas teria que fazer para atingir a meta em 2021, partindo do desempenho observado em 2005.

O plano contemplava diferentes incentivos para que as diferentes escolas e redes de ensino acolhessem tais metas e se comprometessem com elas. As escolas que atingissem as metas eram beneficiadas com o aumento de seus recursos no Programa Dinheiro Direto na Escola (PDDE), mas a principal atenção do governo foi às redes que tinham piores índices. O MEC estabeleceu convênios com estados e municípios, por meio da elaboração local de um Plano de Ações Articuladas (PAR). Pelo PAR, os gestores municipais e estaduais se comprometiam a promover um conjunto de ações, responsabilizando-se pelo alcance das metas estabelecidas no âmbito federal. Em contrapartida, passavam a contar com transferências voluntárias e assessoria técnica da União.

Quando se observa a ação dos Estados e Municípios brasileiros, deve-se destacar que apesar da adesão das redes estaduais e municipais à Prova Brasil ser voluntária, apenas alguns poucos municípios não participaram de algumas das avaliações.¹⁹ Esta elevada participação é confirmada por pesquisa recentemente realizada (BAUER et al, 2017)²⁰. Elemento importante revelado pela pesquisa que pelo menos 20 Estados desenvolveram indicadores e sistemas próprios de avaliação e aproximadamente 30% dos municípios brasileiros também desenvolveram indicadores e processos avaliativos próprios.

Quanto ao movimento de *accountability*, este também não ficou restrito ao governo federal. Diversos estados e alguns municípios, além de manterem ou desenvolverem sistemas próprios de avaliação, introduziram diferentes programas de *accountability*. Conforme destaca Brooke (2006), mesmo antes do governo federal, estados como o Ceará, já em 2001, estabeleceram uma conexão entre o seu sistema de avaliação – o SPAECE – e o Projeto Melhoria da Escola, com prêmios e recompensas em dinheiro para escolas e suas equipes que obtivessem os melhores resultados. Esta experiência foi, ao longo dos anos, modificada, mas a ideia de *accountability* segue até os dias de hoje no Ceará. Indicadores de qualidade educacional estão atualmente presentes, inclusive na própria transferência de parte do ICMS do Estado para os municípios cearenses.²¹

¹⁹ Na primeira aplicação da Prova Brasil, o Estado de São Paulo produziu apenas resultados por regional administrativa, mesmo que todas as escolas da rede estadual tenham participado.

²⁰ Segundo Bauer et al (2017) os municípios não apenas participam da Prova Brasil (96% dos respondentes da pesquisa), mas 97% destes municípios também utilizaram a Provinha Brasil e 90% participaram da ANA.

²¹ Um sistema semelhante está sendo implementado em Pernambuco atualmente.

Além do Ceará, Rio de Janeiro e Paraná, segundo Brooke (2006), também desenvolviam sistemas de responsabilização no mesmo momento em que o governo federal desenvolvia o seu. Depois dos passos dados pelo MEC, outros indicadores semelhantes ao IDEB foram desenvolvidos, por exemplo, em São Paulo, e avaliações em outras séries e disciplinas foram produzidas em diferentes municípios e/ou Estados. Políticas de difusão e análise de resultados foram implementadas em várias localidades, assim como políticas de bônus para professores com base no desempenho dos alunos nos exames e/ou em indicadores assemelhados também foram implementados. Segundo Bauer et al. (2017), existe uma parte substancial dos municípios (85%) que utilizam os resultados das avaliações (próprios ou não) em processos de difusão de resultados para diferentes públicos²² - *accountability* fraca. Quanto à *accountability* forte, com a concessão de prêmio, bônus ou outras consequências, pouco menos de 30% dos municípios respondentes disseram fazer uso deste tipo de iniciativa.

AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E ACCOUNTABILITY: UMA ANÁLISE DA EXPERIÊNCIA BRASILEIRA.

Ainda que não unânime, a implantação e a ampla divulgação das avaliações censitárias, a exemplo da Prova Brasil e do IDEB, tiveram uma boa aceitação por parte da opinião pública. As divulgações têm tido grande repercussão na mídia e têm despertado o interesse de professores e gestores públicos. Botelho et al. (2014) mostram que 80,4% dos professores da rede pública de ensino conheciam o IDEB de sua escola. Há também evidência de que os resultados das avaliações são levados em consideração pela população e impactam a eleição de prefeitos (FIRPO; PIERI; SOUZA, 2017).

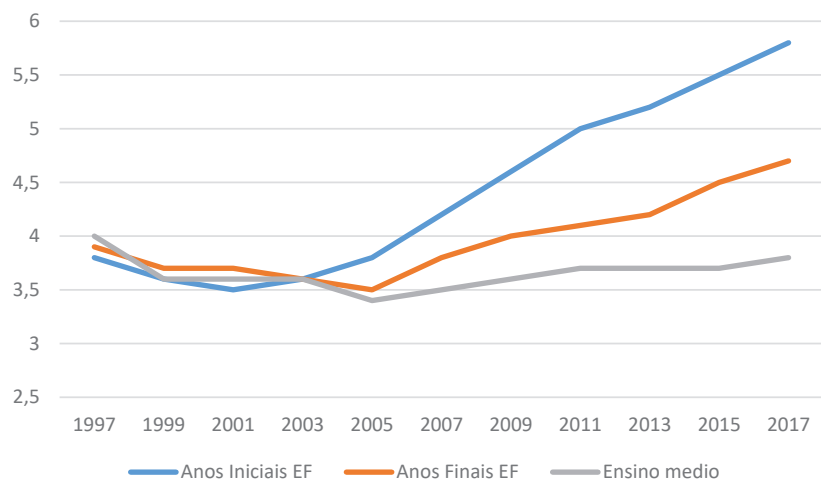
Entretanto, saber qual o impacto de todo esse movimento de avaliação e *accountability* educacional sobre a qualidade da educação do Brasil não é algo que possua uma resposta simples. O gráfico 3 sugere uma inflexão nos IDEBs, justamente em 2005, ano de implantação da Prova Brasil e a partir do qual o IDEB passou a ser divulgado.²³ Evidentemente, não podemos atribuir, automaticamente, a melhora observada no IDEB à política de *accountability*

²² Uma destas formas de divulgação de resultados que foi objeto de polêmicas é a colocação de placas nas escolas com os seus resultados.

²³ Na realidade, a inflexão observada para o ensino fundamental 1 ocorre um pouco antes, em 2001. O IDEB oficialmente é divulgado a partir de 2005, mas é possível calculá-lo desde 1995, com base nos dados do SAEB e do Censo da Educação Básica.

em questão. A melhora poderia ser consequência de outros fatores que se deram independente e simultaneamente à Prova Brasil e ao IDEB. Por outro lado, o crescimento do IDEB é bastante distinto entre as diferentes etapas de ensino. Ele é bastante expressivo na primeira etapa do ensino fundamental, menos pronunciado na segunda etapa do fundamental e pequeno para o ensino médio.

Gráfico 3 – Brasil: evolução do IDEB 1997-2017

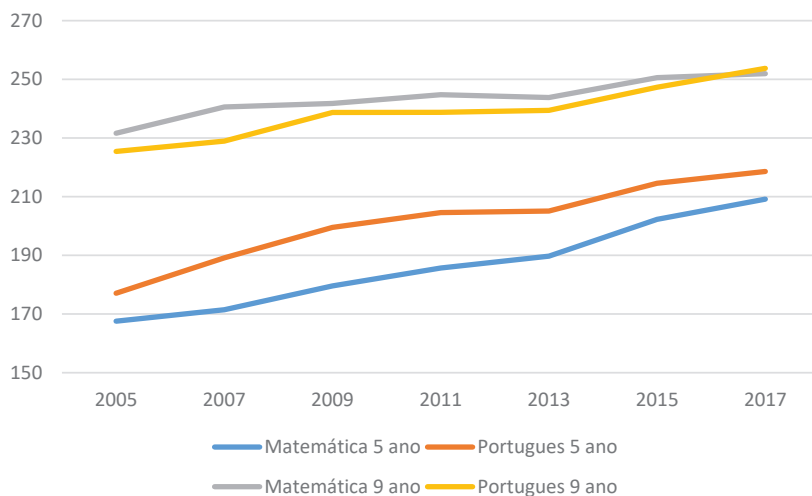


Fonte: elaboração dos autores dados originais do INEP

O gráfico 3 levanta três questões importantes a respeito da evolução do IDEB. Primeira, o aumento significativo do IDEB do ensino fundamental a partir de 2005 é genuíno ou está contaminado por inflação de notas? Segunda, se genuíno, pode ser explicado, ao menos em parte, pela política de *accountability*? Por fim, por que a elevação do IDEB no ensino fundamental não se verifica para o ensino médio? Sem ter a pretensão de dar uma resposta definitiva a essas questões, esta seção tem por objetivo tecer algumas reflexões sobre elas.

Crescimento do IDEB do Ensino Fundamental a partir de 2005: Melhoria no Aprendizado ou Inflação de Notas?

Pelo gráfico 3, percebemos a elevação do IDEB no ensino fundamental, especialmente a partir do segundo lustro da primeira década do século XXI. Se decomposmos os dados do IDEB, podemos perceber que a melhoria se deve muito aos avanços nas proficiências obtidas na Prova Brasil, ou seja, o avanço do IDEB se deve mais às melhoras na proficiência da Prova Brasil/SAEB do que aos avanços nos dados de aprovação. Os avanços nas proficiências podem ser observados no gráfico 4.

Gráfico 4 – Brasil: proficiências Prova Brasil ensino fundamental 2005-2017

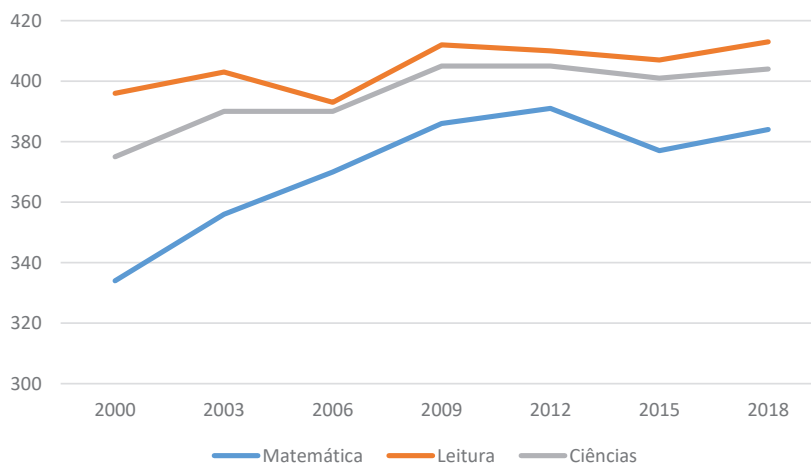
Fonte: elaborado pelos autores com dados do INEP.

Como visto anteriormente, uma possibilidade é que esta melhoria nas proficiências se deva a fenômenos como os de inflação de notas (*score inflation*), quando diretores e professores das escolas adotam medidas que elevam a pontuação nos exames sem que a aprendizagem seja afetada como, por exemplo, motivando e treinando os estudantes para o teste. Esta é efetivamente uma possibilidade, mas não temos conhecimento de análises que tenham conseguido estabelecer quanto da ampliação das notas demonstradas no gráfico 4 se devem a este fenômeno ou quanto é um crescimento genuíno da aprendizagem.

Por outro lado, é importante destacar duas coisas. A primeira é que a elevação das notas, especialmente as do 5º ano, mesmo que pequenas se olhadas de um ciclo avaliativo para outro; quando se olha todo o período, é de mais de 40 pontos (ou por volta de 25 pontos no 9º ano), isto se aproxima a um desvio padrão na escala do SAEB e pode significar algo como o equivalente a quase dois anos adicionais de aprendizagem (um ano no caso do 9º ano), o que não é algo pequeno e dificilmente se obteria apenas com sucessivos processos de treinamento para a Prova Brasil. O segundo elemento a configurar a hipótese que este crescimento não se deve apenas ao fenômeno da “inflação de notas” é o desempenho do Brasil em um outro teste, diferente e aplicado amostralmente no Brasil, no caso os dados do PISA. Estes dados poderiam ser comparados especialmente com os resultados do 9º ano (apesar do resultado do PISA ser negativamente impactados pelos alunos de 15 anos que frequentam o 7º e 8º anos). No gráfico 5, o resultado dessa avaliação (PISA) é

mostrado e possui um comportamento relativamente parecido com os dados da Prova Brasil, reforçando, assim, a hipótese de que, se não podemos afastar totalmente fenômeno do “*score inflation*” no caso brasileiro, ele claramente também não pode explicar todo o ganho de proficiência e efetivamente existe aprendizagem genuína nos dados sobre a evolução do SAEB/Prova Brasil.

Gráfico 5 – Brasil: resultados do PISA 2000-2018



Fonte: elaborado pelos autores com base em dados da OCDE.

Os Programas de Avaliação e Accountability Contribuíram para O Aumento do IDEB a partir de 2005?

Como vimos na subseção anterior, a elevação das pontuações do SAEB e do IDEB do ensino fundamental (especialmente na 1ª fase) foi expressiva e não pode ser justificada apenas por inflação de notas. Assim, somos levados a concluir que houve melhora das condições de ensino-aprendizagem após 2005. Evidentemente, não temos como afirmar que tal fato não teria ocorrido se a política de *accountability* relacionada à Prova Brasil e ao IDEB não tivesse sido implantada. Uma melhor interpretação seria, talvez, que tal crescimento decorre de uma maior mobilização de diversas esferas de governo (federal, estadual e municipal) e da sociedade em geral para a melhoria da educação, onde a política relacionada à Prova Brasil e o IDEB seja apenas um elemento desse movimento mais geral.

Vários fatores poderiam ser relacionados para ajudar a explicar o crescimento das pontuações obtidas nas avaliações. Em primeiro lugar, o grande aumento na taxa de frequência à pré-escola ocorrido entre 1985 e 2005 pode ter contribuído para que as crianças que ingressaram no ensino fundamental no pe-

ríodo subsequente tivessem maior capacidade de aprendizado. Além disso, o aumento educacional das mães, que ocorreu entre 1995 e 2005, também pode ter contribuído para aumentar os investimentos familiares nas crianças.

Uma medida que merece destaque diz respeito à ampliação do ensino fundamental de oito para nove anos.²⁴ Os gráficos 6 e 7 mostram a evolução do IDEB da primeira fase do ensino fundamental para duas redes de ensino: rede estadual de Minas Gerais e rede municipal de Ribeirão Preto (SP).

Gráfico 6 – Minas Gerais IDEB – anos iniciais



Fonte: elaborado pelos autores com base em dados do IBGE.

Gráfico 7 – Ribeirão Preto IDEB – anos finais



Fonte: elaborado pelos autores com base em dados do IBGE.

²⁴ Em 06/02/2006, foi sancionada lei Federal regulamentando o ensino fundamental com 9 séries, dando prazo até 2010 para que todas as redes de ensino do país se enquadrassem na nova regra. Entretanto, algumas redes já haviam adotado o ensino fundamental de nove anos antes de 2006.

Em Minas Gerais, o ano 2009 é o primeiro ano em que os alunos sem repetência chegam ao final da primeira fase do ensino fundamental com cinco anos de escolaridade, ao invés de quatro. Em Ribeirão Preto, isso ocorre no ano de 2011. Assim, o impacto da implantação do ensino fundamental de nove anos parece notório. Eleva o IDEB em cerca de um ponto. Como diferentes redes implantaram o ensino fundamental de nove anos em anos distintos, o impacto da medida na média do IDEB tende a ser mais suave no tempo. Peña (2014) avalia o impacto do ensino fundamental de nove anos na primeira fase do fundamental entre 2007 e 2011, concluindo que a medida elevou o SAEB de matemática em 4,43 pontos (0,09 DP) e de leitura em 5,11 pontos (0,1 DP). Isso explicaria apenas 11% e 14% da variação observada no período, respectivamente, na pontuação de matemática e leitura.

Por fim, iniciativas para melhorar a gestão em alguns municípios, tais como Sobral no Ceará, também trouxeram bons resultados em termos de aprendizado.²⁵

Note que as políticas de *accountability* não concorrem com várias medidas adotadas pelas escolas e sistemas educacionais na explicação da melhora do desempenho observado. A *accountability* não afeta diretamente as práticas de ensino. Ela altera a estrutura de incentivos e, se funcionar, leva os responsáveis a ofertar educação a adotarem medidas que afetam diretamente o processo de ensino-aprendizagem: controle de faltas docentes, redução do tamanho das turmas, alteração do currículo, tutoria etc.

De todo modo, dado a magnitude da evolução dos indicadores e o período da reversão de tendência, é difícil argumentar que as políticas de avaliação e *accountability* não contribuíram em nada com a melhoria da aprendizagem no ensino fundamental no Brasil.

Por Que o Crescimento do IDEB no Ensino Fundamental não Atingiu o Ensino Médio?

O crescimento do IDEB observado para o ensino fundamental não se repete para o ensino médio. O crescimento do IDEB entre 2005 e 2017 foi de apenas 0,4 pontos nesse segmento. E mais, esse crescimento foi comandado pela redução da repetência. No período, a proficiência cresceu 10,9 pontos em leitura e diminuiu 0,7 em matemática. Em matemática, a proficiência atinge em 2015 (267,5 pontos) o menor valor observado para o SAEB/Prova Brasil, desde sua implantação em 1995. Esse desempenho pouco favorável do ensino médio tem gerado preocupação e controvérsia.

²⁵ Ver Rocha, Komatsu; Menezes-Filho (2018).

Uma parcela da diferença de crescimento entre os IDEBs do ensino fundamental e do ensino médio poderia ser explicado pela inflação de notas. Como o ensino médio não tem prova Brasil e IDEB por escolas e rede de ensino, os resultados SAEB e IDEB do ensino médio são *low-stakes* e, em princípio, não contaminados pela inflação de notas. Como vimos, entretanto, a inflação de notas não explica toda a elevação do IDEB e Prova Brasil do ensino fundamental.

A interpretação mais pessimista da evolução dos indicadores educacionais a partir de 2005 é de que, em virtude da falta de mudanças no ensino após os anos iniciais do ensino fundamental, os ganhos iniciais vão sendo perdidos ao longo dos anos finais do ensino fundamental e do ensino médio. As gerações beneficiadas pelo melhor ensino nos anos iniciais do ensino fundamental apresentariam, ao final do ensino médio, o mesmo desempenho que teriam atingido na ausência da melhoria do ensino nos anos iniciais. Todo esforço ocorrido nos anos iniciais seria perdido!

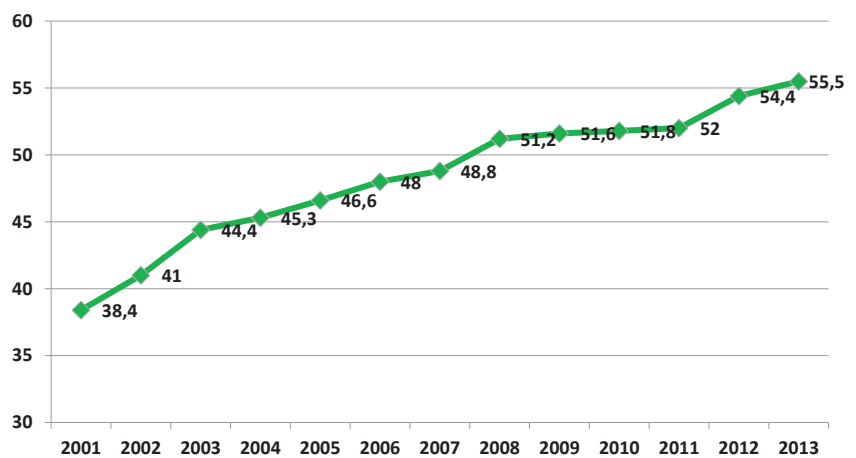
Entretanto, essa interpretação não é necessariamente correta e, do nosso ponto de vista, não parece ser a mais plausível. Em primeiro lugar, a composição daqueles que fazem a prova em determinada série pode estar mudando ao longo do tempo. Tais efeitos de composição podem ser decorrentes de alterações nos padrões de repetência e evasão. É possível também que seja mais difícil avançar na escala do SAEB quando a pontuação obtida é mais elevada. Assim, em termos de aprendizado, um crescimento de cinco pontos no final da primeira fase do ensino fundamental (por exemplo, de 210 para 215) pode ser menos significativo do que um crescimento de 2 pontos ao final do ensino médio (por exemplo, de 300 para 302).

O primeiro ponto (efeitos de composição) chama a atenção para o fato de que, na presença de repetência e evasão, não é tão claro como deveríamos avaliar se o desempenho dos estudantes está melhorando ao longo do tempo. Podemos pensar em três alternativas: (i) comparar o desempenho de determinada série ao longo do tempo, independentemente da idade dos alunos; (ii) comparar o desempenho de gerações sucessivas em determinada idade, independentemente da série cursada; (iii) comparar o desempenho de gerações sucessivas em determinada série, independentemente do ano em que o aluno é testado. Se todos os alunos ingressassem na escola na idade correta e não houvesse repetência nem evasão, essas três comparações seriam idênticas. No entanto, na presença de repetência e evasão, elas podem apresentar resultados muito distintos em caso dessas taxas sofrer alteração ao longo do tempo.

Por exemplo, uma redução das taxas de evasão pode levar a um aumento no desempenho entre gerações sucessivas, avaliadas em determinada idade, mas a uma redução no desempenho dos estudantes em determinada série ao longo do tempo. Isso porque uma parcela de estudantes de determinada geração (que antes não seriam esperados a atingir a série em questão) agora é observada na série e esses estudantes possuem, presumivelmente, desempenho inferior daqueles que, nas condições anteriores, seriam esperados a atingir a série em consideração.

Parece claro que tanto a repetência quanto a evasão vêm caindo ao longo do tempo. O Gráfico 8 mostra a evolução da taxa líquida de matrícula do ensino médio. Ela sugere que uma parcela de jovens que, nas condições que vigoravam 10 anos atrás, não chegariam ao final do ensino médio, hoje chegam. Como esses jovens, presumivelmente, possuem pior desempenho, esse fato poderia explicar, ao menos em parte, o porquê o crescimento observado na pontuação do SAEB e IDEB sete anos atrás na primeira fase do ensino fundamental não é observado hoje ao final do ensino médio.

Gráfico 8 – Taxa líquida de matrícula no ensino médio



Fonte: elaborado pelos autores dados básicos INEP.

Em relação ao segundo ponto (a dificuldade de avançar na escala do SAEB na medida em que pontuações mais elevadas vão sendo obtidas), ele tem sido observado em diversos exames que utilizam uma escala comum para diversas séries. Por exemplo, a tabela 1 apresenta as pontuações da amostra de normatização do *Comprehensive Test of Basic Skills* (CTBS), modelo U, desenvolvido pela CTB/McGraw-Hill, verticalmente escalado com base na TRI e utilizando um modelo logístico de três parâmetros. Em ambas disciplinas,

o ganho entre séries cai drasticamente entre as séries mais baixas e as mais elevadas. Além disso, o desvio padrão das pontuações diminui na medida em que a pontuação média sobe.

A redução do crescimento da pontuação média na medida em que a pontuação média aumenta pode ser uma consequência da forma como a escala é construída. Como destaca Ballou (2009), as escalas com base na TRI admitem que o aumento de proficiência necessário para elevar, em determinado montante, a probabilidade de responder um item corretamente (dado uma probabilidade inicial) é independente do grau de dificuldade do item. Isso significa que, considerando itens que se diferenciam apenas pelo grau de dificuldade, o aumento de proficiência requerido para um examinado com baixa proficiência elevar a probabilidade de responder corretamente um item fácil de 0,1 para 0,9 é o mesmo que o requerido para outro examinado de alta proficiência elevar a probabilidade de acerto de 0,1 para 0,9 de um item muito difícil. Ou seja, os conhecimentos e habilidades que o primeiro examinado precisaria adquirir têm a mesma medida que os conhecimentos e habilidades que o segundo examinado teria que adquirir.

Tabela 1 – Pontuação no *Comprehensive Test of Basic Skills* (CTBS/U) – 1981, Amostra de Normatização

Série	Leitura/Vocabulário			Matemática		
	Pontuação Média	Desvio Padrão	Variação	Pontuação Média	Desvio Padrão	Variação
1	488	85	-	390	158	-
2	579	78	91	576	77	186
3	622	65	43	643	44	67
4	652	60	30	676	35	33
5	678	59	26	699	24	23
6	697	59	19	713	20	14
7	711	57	14	721	23	6
8	724	54	13	728	23	7
9	741	52	17	736	17	8
10	758	52	17	739	16	3
11	768	53	10	741	18	2
12	773	55	5	741	20	0

Fonte: Yen (1986).

Alguém pode considerar que a noção de proficiência implícita nos modelos de TRI não casa com a sua noção intuitiva de aquisição de conhecimentos e habilidades. Desse modo, o fato de o crescimento da proficiência se reduzir na medida em que a proficiência aumenta, não significa, necessariamente, que o aprendizado diminui com a elevação da proficiência, uma vez que isso depende da noção de aprendizado que cada um possui.

Se for verdade que é mais difícil (envolve mais esforço) atingir uma probabilidade de acerto de 0,9 para um item muito difícil (quando os conhecimentos e habilidades adquiridos fixam uma probabilidade de acerto em 0,1), do que atingir essa mesma probabilidade para um item fácil (quando a probabilidade de acerto é também 0,1), então, avançar na escala de proficiência vai se tornando mais difícil na medida em que se atingem níveis mais elevados de proficiência. Sendo assim, se ocorresse um aumento de esforço igual para todos os estudantes de todas as séries (ou uma melhora equivalente na qualidade de ensino), seria de esperar que o aumento de proficiência fosse maior nas séries iniciais (onde a proficiência é menor) do que nas séries mais elevadas. Do mesmo modo, o ganho de proficiência nas séries iniciais, em virtude de um melhor ensino, tende a se reduzir na medida em que vai se avançando no sistema, supondo que as condições de ensino tenham se mantido constante nas séries mais elevadas.

Note que em toda discussão acima foi admitido que proficiência é unidimensional, como é o caso do modelo de TRI adotado no SAEB. Enquanto a hipótese de unidimensionalidade é questionável para um exame restrito a uma única série, sua aplicação para uma escala vertical, cobrindo diversas séries, é, sem dúvida, bem mais problemática. Por exemplo, vamos admitir que o ensino de matemática nos anos iniciais do ensino fundamental se restrinja apenas à aritmética e que nos anos finais do ensino fundamental, além da aritmética, entrassem outros conteúdos, como álgebra e geometria. Assim, os itens do caderno de teste para os alunos do final da primeira fase do ensino fundamental cobririam apenas aritmética, enquanto os itens do caderno de teste para os alunos do final da segunda fase do ensino fundamental cobririam aritmética, álgebra e geometria. Vamos admitir agora que um aluno, ao final da primeira fase do ensino fundamental, apresente uma proficiência de matemática maior do que a de um aluno ao final da segunda fase do ensino fundamental. De acordo com a metodologia de ligação adotada pelo SAEB (grupos não equivalentes e itens comuns), isso significa que o aluno da primeira fase do ensino fundamental possui mais conhecimentos e habilidades em aritmética, mas não, necessariamente, em matemática. Isso porque o aluno da segunda fase do ensino fundamental pode ser capaz de acertar itens de álgebra e geometria que o aluno da primeira fase do ensino fundamental não é capaz, pelo simples fato de não ter sido exposto a esse material. Se o aluno dos anos iniciais do ensino fundamental fizesse o teste dos anos finais (ao invés do teste dos anos iniciais) apresentaria, provavelmente, um desempenho inferior ao verificado ao realizar o teste referente ao seu nível. Assim, o avanço da proficiência média entre os alunos da primeira e segunda fase do ensino fundamental “subestimaria” o “real” avanço médio de conhecimentos e habilidades matemáticas.

Em suma, mudanças na composição dos alunos que realizam o SAEB de ensino médio é a maior dificuldade de avançar na escala, na medida que a pontuação aumenta poderiam explicar, ao menos em parte, o porquê o desempenho no ensino médio não apresenta a mesma evolução do desempenho observado no ensino fundamental alguns anos antes.

CONCLUSÃO

Este artigo buscou avaliar o movimento de avaliação educacional em larga escala e *accountability* no Brasil. Realizou-se uma breve revisão da literatura internacional sobre *accountability* educacional, destacando os argumentos em sua defesa, a posição dos críticos e, principalmente, a avaliação de seus resultados. Traçou-se um histórico do movimento de avaliação em larga escala e *accountability* no país. Por fim, fez-se uma avaliação desse movimento.

O Brasil possui hoje um sistema de avaliação da educação básica que, sem exagero, poderia ser classificado entre os melhores do mundo. Mas, evidentemente, pode ser aprimorado. Além de retratar o passado, as avaliações têm o papel de sinalizar para o sistema o que se espera das escolas. Então, ao menos para a segunda fase do ensino fundamental e para o ensino médio, as avaliações deveriam ir além de leitura e matemática e incluir ciências da natureza e humanidades. No ensino médio, isso era feito com o ENEM, mas foi revertido com a decisão do INEP de não mais divulgar os resultados do ENEM por escolas, e universalizar o SAEB para as escolas públicas de ensino médio.

O artigo enfatiza as políticas de avaliação e *accountability* implantadas pelo Ministério da Educação. Em um outro trabalho, seria interessante apresentar e avaliar mais detalhadamente as experiências de avaliação e *accountability* realizadas por estados e municípios.

REFERÊNCIAS

- AOS, S. PENNUCCI, A. K-12 **Class Size Reductions and Student Outcomes: a Review of the Evidence and Benefit-Cost Analysis**. Documento n. 13-01-2201, Washington State Institute for Public Policy, 2013.
- BALLOU, D. Test Scaling and Value-Added Measurement. **Education Finance and Policy**, MIT Press, v. 4, n. 4, p. 351-383, 2009.
- BAUER, A. et al. Iniciativas de avaliação do ensino fundamental em municípios brasileiros. **Revista Brasileira de Educação**, v. 22, n. 71, p. 1-19, out. 2017.
- BEHRENDT, A.; EISENACH, J. JOHNSON, W. R. Selectivity Bias and the Determinants of SAT Scores. *Economics of Education Review*, 5 (4), p. 363-371, 1986.
- BONAMINO, A.; FRANCO, C. Avaliação e Política Educacional: o Processo de Institucionalização do SAEB. **Cadernos de Pesquisa**, Rio de Janeiro, n. 108, p. 101-132, nov. 1999.
- BOTELHO, F. B. et al. Sistemas de Accountability nas Escolas Públicas Brasileiras: Identificando a Eficácia das Diferentes Experiências. In: Fernandes, R., Souza, A. P. F., BOTELHO, F.; SCORZAFAVE, L. G. (orgs). **Políticas públicas educacionais e desempenho escolar dos alunos da rede pública de ensino**. Ribeirão Preto: FUNPEC-Editora, 2014, p. 59-80.
- BROOKE, N. O futuro das políticas de responsabilização educacional no Brasil. **Cadernos de Pesquisa**. v. 36, n. 128, p. 377-401, maio./ago. 2006.
- CASTRO, M. H. G. O Saeb e a Agenda de Reformas Educacionais: 1995- 2002. **Em Aberto**, Brasília, v. 29, n. 96, p 85-98, maio/ago. 2016.
- CLARK, M.; ROTHSTEIN, J.; SCHANZENBACH, D. W. Selection Bias in College Admissions Test Scores. **Economics of Education Review**, Cambridge, n. 28, p. 295-307, aug. 2008.
- COLEMAN, J. et al.; **Equality of Educational Opportunity**. Washington DC, 1966.
- DYNARSKI, M. The Scholastic Aptitude Test: Participation and performance. **Economics of Education Review**, Elsevier, v. 6, n.3, p.263-273, jun. 1987.
- DYNARSKI, M. GLEASON, P. Using Scholastic Aptitude Test Scores as indicators of state educational performance. **Economics of Education Review**, Elsevier, v. 12, n. 3, p. 203-211, jun. 1993.

FERNANDES, R. A universalização da avaliação e a criação do IDEB: pressupostos e perspectivas. **Em Aberto**, Brasília, v. 29, n. 96, mai./ago. 2016.

_____. Índice de Desenvolvimento da Educação Básica (Ideb). Brasília: Inep, **Textos para Discussão**, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 26p., 2007.

FERNANDES, R. GREMAUD, A.. Qualidade da Educação: Avaliação, Indicadores e Metas. In: Veloso, F., Pessôa, S., Henriques, R. e Giambiagi, F. (orgs). **Educação Básica no Brasil: construindo o país do futuro**. Editora Campus/Elsevier, 2009, p. 213-238.

FIGLIO, D. LOEB, S. School Accountability. In Hanushek, E., Machin, S. e Woessmann, L. (Eds) **Handbook of the Economics of Education**, v. 3, North-Holland, 2011, p. 383-421.

FIRPO, S.; PIERI, R.; SOUZA, A. P. Electoral Impacts of Uncovering Public School Quality: Evidence from Brazilian Municipalities. **Revista Economia**, v. 18, p. 1-17, 2017.

FLETCHER, P. R.; RIBEIRO, S. C. **Modeling education system performance with demographic data: an introduction to the PROFLUXO model**. Paris: Unesco, 1989.

GATTI, B. Avaliação educacional no Brasil: experiências, problemas, recomendações. **estudos em avaliação educacional**, Fundação Carlos Chagas, n.10, p. 67-80, 1994.

HANUSHEK, E. A. Grinding the antitesting ax: more bias than evidence behind NRC panel's conclusions. **Educational Next**, v. 12, n. 2, p. 49-55, 2012.

HANUSHEK, E. A.; RAYMOND, M. Does school accountability lead to improved student performance? **Journal of Policy Analysis & Management**, v. 24, n. 2, p. 297-327, 2005.

HARBISON, R.W.; HANUSHEK, E. A. **Educational performance of the poor**. New York: Oxford University Press, 1992.

HORTA NETO, J. L. Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o SAEB de 2005. **Revista Ibero Americana de Educacion**, Brasília, v. 42. n. 5, p 1 -13, abr. 2007.

HOUT, M.; ELLIOTT, S. W. **Incentives and test-based accountability in education**. National Research Council of National Academies. Washington, DC: The National Academies Press, 2011.

KANE, T.; STAIGER, D. The promise and pitfalls of using imprecise school accountability measures. **Journal of Economic Perspectives**, v.16, n. 4, p. 91–114, fall 2002.

KLEIN, R.; RIBEIRO, S. C O Censo Educacional e o Modelo de Fluxo: o problema da repetência. **Revista Brasileira de Estatística**, v. 52, n. 197-198, p. 5-45, nov. 1991.

PEÑA, P. A. **Impact of extension of elementary education in Brazil on test scores**. Estudo realizado para o programa Todos pela Educação, 2014. Disponível em: <http://www.todospelaeducacao.org.br/biblioteca/1524/impact-of-elementary-school-extension-on-test-scores/>. Acesso em: 1 de novembro de 2019

PESTANA, M. I. Trajetória do SAEB: criação, amadurecimento e desafios. *Em Aberto*, Brasília, v. 29, n. 96, p. 71-84, mai./ago/, 2016.

REARDON, S. F. RAUDENBUSH, S. W. Assumptions of value-added models for estimating school effects. **Education Finance and Policy**, v. 4, n. 4, p. 492-519, 2009.

ROCHA, R.; KOMATSU, B.; MENEZES-FILHO, N. . **Avaliando o impacto das políticas educacionais em Sobral, Texto para discussão**, Insper, n.35, out. 2018.

YEN, W. The choice of scale for educational measurement: an IRT perspective. **Journal of Educational Measurement**, v. 23, n. 4, p.299-325, 1986.

Recebido em: 03 de dezembro de 2019

Aceito em: 01 de junho de 2020

Publicado em: 30 de junho de 2020

AVALIAÇÃO EDUCACIONAL: FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR EM CENÁRIOS DE ALTA VULNERABILIDADE SOCIAL¹

Edivaldo Cesar Camarotti Martins²

Adolfo-Ignacio Calderón³

¹O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

²Doutorando e Mestre em Educação pela Pontifícia Universidade Católica de Campinas (PUC-Campinas), Campinas, São Paulo, Brasil. E-mail: edivaldoccmartins@gmail.com.

³ Professor titular do Programa de Pós-Graduação em Educação (PPGE) da Pontifícia Universidade Católica de Campinas (PUC-Campinas), Campinas, São Paulo, Brasil. E-mail: professoradolfoalderon@gmail.com.

Resumo

Este artigo objetiva analisar os fatores contextuais de eficácia escolar que contribuíram para que uma escola, localizada em região de alta vulnerabilidade social no município de Ferraz de Vasconcelos, Estado de São Paulo, superasse as metas projetadas no Índice de Desenvolvimento da Educação Básica nos anos de 2007, 2009 e 2011. Realizamos pesquisa empírica, de natureza qualitativa, com 21 entrevistados (professores, diretor, alunos, pais e funcionários). Os dados foram sistematizados e discutidos quali-quantitativamente à luz da literatura acadêmico-científica. Os resultados apontam que o compromisso com a aprendizagem e as práticas pedagógicas foram os fatores contextuais de eficácia escolar mais relevantes para o alcance dos objetivos de aprendizagem e a superação das metas, resultado de um conjunto articulado de ações, recursos e práticas desenvolvidas pelos diversos atores escolares.

Palavras-chave: Avaliação educacional. Educação básica. Eficácia escolar. Fatores contextuais. Alta vulnerabilidade social.

Abstract

This article aims to analyze the contextual factors of school effectiveness that contributed to a school located in a region of high social vulnerability in the municipality of Ferraz de Vasconcelos, State of São Paulo, exceeding the goals projected in the Basic Education Development Index (Índice de Desenvolvimento da Educação Básica in portuguese) in the years 2007, 2009 and 2011. Qualitative empirical research was conducted with interviews with 21 respondents (teachers, director, students, parents and employees). The data were systematized and discussed quantitatively in light of academic-scientific literature. The results indicate that the commitment to learning and pedagogical practices were the most relevant contextual factors of school effectiveness to overcome goals. The result of an articulated set of actions, resources, and practices developed by the various school actors.

Keywords: Educational evaluation. Basic education. Schools effectiveness. Contextual factors. High social vulnerability.

INTRODUÇÃO

Na década de 1990, com o surgimento dos sistemas de avaliação em larga escala no Brasil, o poder público passou a avaliar a qualidade do ensino sob o crivo dos resultados das avaliações externas. A qualidade do ensino, que até então era relacionada com informações e estatísticas censitárias, como, por exemplo: número e taxas de matrículas, de aprovação ou de repetência; passou a ser vinculada à efetiva aprendizagem dos alunos, traduzida no desempenho escolar observado nas avaliações em larga escala (CALDERÓN, 2017).

A Carta Magna de 1988 (BRASIL, 1988), em especial o Artigo 206, inciso VII, que estabelece a “garantia de padrão de qualidade”, bem como o Artigo 9º, inciso VI, da Lei de Diretrizes e Bases da Educação Nacional (LDB) (BRASIL, 1996), que assegura um “processo nacional de avaliação do rendimento escolar”, determinaram a existência de toda uma estrutura de avaliação da qualidade do ensino fundamental, médio e superior, nos diversos níveis de governo.

Na avaliação da qualidade da educação básica construiu-se, desde então, um sólido sistema de avaliação que passou por um contínuo processo de aprimoramento e que permanece até os dias atuais, mesmo com alternâncias significativas de governos. O Sistema de Avaliação da Educação Básica (SAEB), criado no ano de 1990, durante o curto mandato do Presidente Fernando Collor de Mello, do Partido da Reconstrução Nacional (PRN), foi formalmente institucionalizado no ano de 1994, no final do mandato do Presidente Itamar Franco, do Partido do Movimento Democrático Brasileiro (PMDB) (PESTANA, 1996), e aprimorado nos dois mandatos do Presidente Fernando Henrique Cardoso (1995-2003), do Partido da Social Democracia Brasileira (PSDB), foi ampliado nos dois mandatos do Presidente Luiz Inácio Lula da Silva (2003-2011), do Partido dos Trabalhadores (PT), e manteve-se até os governos subsequentes, como o principal sistema de monitoramento da aprendizagem de crianças e jovens em idade escolar.

Essa cultura avaliativa que se instaurou nos sistemas de ensino contribuiu para consolidar no Brasil uma espécie de algoritmo governamental: desempenho escolar \equiv qualidade do ensino. Essa analogia ao símbolo matemático de equivalência significa afirmar que, na perspectiva dos governos, quanto maior o desempenho escolar, melhor a qualidade do ensino. Nesse contexto, o aperfeiçoamento técnico dos instrumentos de avaliação em larga escala ocorrido no Brasil, a partir da década de 1990, foi fundamental para a criação de indicadores, com destaque para o Índice de Desenvolvimento da Educação Básica (IDEB), que mensura a evolução da qualidade do ensino na educação básica em cada escola e nas redes de ensino no Brasil.

Apesar das profundas críticas existentes às limitações desse indicador de qualidade (HORTA NETO, 2018; PONTES; SOARES, 2016; FIGUEIREDO, 2018), uma característica inerente ao IDEB é a projeção de metas para as escolas e redes de ensino (BRASIL, 2013), de tal forma que, quando os resultados do IDEB são divulgados, juntamente com o desempenho obtido pelas escolas e redes de ensino, são divulgadas as metas projetadas, evidenciando escolas com alto e baixo desempenho dentro de uma mesma rede de ensino, além de mostrar a evolução das escolas e redes de ensino em termos do cumprimento ou não cumprimento das metas.

A periódica divulgação dos resultados do IDEB acaba tornando públicos alguns casos de escolas eficazes que, mesmo imersas em regiões de alta vulnerabilidade social, conseguem atingir e até superar as metas de desempenho escolar projetadas pelo governo. Diante disso, emergiu nosso problema de pesquisa: Quais fatores escolares podem contribuir para que escolas localizadas em regiões de alta vulnerabilidade social consigam alcançar os objetivos de aprendizagem de seus alunos, atingindo ou até mesmo superando as metas de desempenho escolar projetadas no IDEB?

Para encontrarmos possíveis respostas a essa questão, realizamos um estudo de caso que analisou Fatores Contextuais de Eficácia Escolar (FCEE) em uma escola que obteve alto desempenho em seguidas edições do IDEB. Selecionamos uma escola localizada em uma região de alta vulnerabilidade social⁴, pertencente à rede municipal de ensino de Ferraz de Vasconcelos – São Paulo, instalada em um bairro periférico da cidade, com características típicas de regiões carentes e de pobreza. Ao contrário de outras escolas localizadas em regiões de alta vulnerabilidade social da mesma rede de ensino, a escola selecionada cumpriu todas as metas projetadas pelo IDEB nos anos de 2007, 2009 e 2011.

A escola possuía, no ano de 2014, aproximadamente 700 alunos, atendia aos anos iniciais do ensino fundamental (do 1º ao 5º ano) e funcionava nos períodos matutino e vespertino. Não possuía um prédio suntuoso, mas, além de 12 salas de aula, contava com: biblioteca; sala de informática com 28 computadores com acesso à internet; sala de atendimento especializado (educação especial); refeitório; sanitários adequados; secretaria; sala de professores; salas para os gestores; quadra de esportes (descoberta) e área de convivência (pátio coberto). O espaço escolar, de um modo geral, aparentava ser bem

⁴ Como parâmetro na definição do que seja alta vulnerabilidade social, adotamos o Índice Paulista de Vulnerabilidade Social (IPVS) da Fundação Sistema Estadual de Análise de Dados (SEADE), que considera as regiões de alta vulnerabilidade social (IPVS = 5) como regiões de pobreza, baixa renda ou socioeconomicamente desfavorecidas.

cuidado, tanto na questão da limpeza, quanto da conservação dos ambientes e dos equipamentos escolares.

Neste estudo, foi realizada uma pesquisa empírica, de natureza qualitativa, por meio de 21 entrevistas semiestruturadas, em que foram ouvidos em separado: 5 professores, 1 diretor, 5 funcionários, 5 alunos e 5 pais⁵. Todos os professores eram efetivos na escola, sendo que dois deles possuíam mais de 10 anos de tempo no magistério, sendo 5 anos naquela escola. Os demais professores possuíam entre 5 e 7 anos de tempo no magistério e pouco mais de 2 anos naquela escola. O diretor era professor efetivo da rede municipal, designado para o cargo já havia dois anos. Em relação aos funcionários, dois deles trabalhavam na secretaria, dois eram inspetores de alunos e um trabalhava na limpeza. Todos os alunos entrevistados estavam matriculados no 5º ano e todos os pais eram membros do Conselho de Escola.

Durante as entrevistas, a questão norteadora que ensejava respostas livres foi a seguinte: sua escola obteve alto desempenho em três edições seguidas do IDEB, atingindo as metas projetadas pelo governo. A que fatores você atribui esse resultado? Cite-os buscando exemplificar com evidências.

A sistematização dos dados das entrevistas consistiu em relacionar a frequência de citação de temas, palavras ou ideias para posteriormente dimensionar o peso atribuído pelos próprios respondentes aos fatos mencionados (CHIZZOTTI, 2006) e, dessa forma, identificarmos semelhanças e convergências discursivas nas informações coletadas sobre os fatores contextuais de eficácia escolar.

A partir da percepção dos atores escolares, os fatores contextuais identificados foram categorizados em quatro dimensões: recursos escolares, gestão escolar, práticas pedagógicas e clima escolar (MARTINS, 2015). Nessa ótica, as opiniões dos respondentes (dados qualitativos) foram transformados em quadros (dados quantitativos), categorizados, hierarquizados, analisados e comparados, de tal modo se configurando em dados quali-quantitativos.

Alicerçados no conceito de escola eficaz proposto por Murillo Torrecilla (2005), o qual considera como escolas eficazes, aquelas que conseguem promover o desenvolvimento integral de todos e cada um de seus alunos, indo além do esperado, quando consideradas as condições prévias dos alunos, estamos considerando como fatores contextuais de eficácia escolar, que contribuem

⁵ A pesquisa com seres humanos foi aprovada pelo Comitê de Ética em Pesquisa. Dados do Processo: CAEE: 19013613.2.0000.5481; Número do Parecer: 443.520; Data da Relatoria: 01/11/2013. Obtivemos autorização da Secretaria Municipal de Educação de Ferraz de Vasconcelos para realizar a pesquisa empírica na escola por meio do Ofício nº 303/2013, datado de 20/05/2013.

diretamente para que determinada instituição de ensino seja considerada uma escola eficaz: acontecimentos; ações; contextos; práticas; atividades ou recursos, intra ou extraescolares.

A abordagem de desenvolvimento integral supera visões limitadas em eficácia escolar, as quais consideram apenas o rendimento em matemática, leitura e escrita como parâmetro de qualidade de ensino. O objetivo inalienável da escola e dos sistemas educacionais deve ser justamente o desenvolvimento integral dos alunos, ou seja, matemática, leitura e escrita são importantes, mas também é fundamental o desenvolvimento de atitudes críticas e criativas nos alunos (MURILLO TORRECILLA, 2011).

Concordamos que as investigações em eficácia escolar precisam “ir além dos limites da metodologia estatística, buscando na abordagem qualitativa um conhecimento mais complexo do que os números indicam” (SOARES, 2002, p. 6), para que efetivamente “contribuam para formar um corpo de conhecimento que possa ajudar a otimizar os níveis de qualidade e equidade da educação” (MURILLO TORRECILLA, 2008, p. 467), sem correr o risco de permanecer esquecidos nas bibliotecas universitárias ou nos centros de pesquisa especializados no assunto. No contexto da nossa concepção de eficácia escolar, assumimos uma visão humanista e socialmente comprometida de qualidade do ensino, para muito além de antigas concepções tecnocráticas.

FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR EM CENÁRIO DE ALTA VULNERABILIDADE SOCIAL

Este artigo soma-se a toda uma literatura “que procura identificar as características organizacionais que são determinantes para a eficácia das escolas” (NÓVOA, 1999, p. 26). Trata-se de um campo de pesquisa que vem recebendo diversas críticas no cenário acadêmico-científico. Murillo Torrecilla e Hernández Castilla (2011, p. 4), por exemplo, destacam que, no âmbito internacional, essa linha de estudos é considerada como uma “visão mecanicista do processo educativo, e, com isso, neoliberal”. No Brasil, para Luiz Carlos Freitas, destacado crítico das pesquisas sobre eficácia escolar, esses estudos “são sonhos positivistas” (FREITAS, 2002) que integram uma “utopia liberal”. Dentro dessa concepção, analisar os fatores que afetam a qualidade da aprendizagem nada mais seria que uma atitude ingênua na perspectiva da equidade educacional (FREITAS, 2003).

Entretanto, mesmo entre os críticos aos estudos sobre eficácia escolar, é possível encontrar diversas variações interpretativas, alguns dos quais revelam entendimentos menos pessimistas sobre o assunto. Silva (2015), por exemplo, apesar de relacionar eficácia escolar com pressupostos economicistas, utilitaristas e intervencionistas aplicados à educação, afirma “que as escolas realmente têm grande influência no desenvolvimento dos alunos” e “podem fazer a diferença na qualidade da aprendizagem”, a fim de “contribuir para a garantia do direito à educação”, se “considerarem as características contextuais” (SILVA, 2015, p. 732) de cada realidade analisada.

Fato é que, mesmo em meio a críticas muitas vezes rígidas por parte de pesquisadores, intelectuais e acadêmicos, a literatura sobre eficácia escolar, vagarosa e paulatinamente, vem conquistando espaço na literatura acadêmico-científica e no contexto das políticas públicas no Brasil, seja por meio de teses, dissertações, artigos científicos ou publicações institucionais de estudos realizados pelo governo federal em parceria com agências multilaterais (MARTINS; CALDERÓN, 2016). Entretanto, apesar de crescente, a produção científica sobre eficácia escolar ainda pode ser considerada reduzida no Brasil, muito em função da predominância do discurso crítico a esse tipo de estudo no campo educacional brasileiro (MARTINS, CALDERÓN, 2015, 2019; JANUÁRIO, 2019).

No que se refere à literatura internacional, destacamos a coletânea de Nigel Brooke e José Francisco Soares, com textos estrangeiros traduzidos para a língua portuguesa, que têm como objetivo “descrever as origens, as trajetórias, os resultados e as polêmicas que caracterizam a pesquisa em eficácia escolar” (BROOKE; SOARES, 2008, p. 10). No contexto ibero-americano, a produção de Francisco Javier Murillo Torrecilla, pesquisador da Universidade Autónoma de Madri, toma como referência os estudos pioneiros desenvolvidos na Inglaterra e nos Estados Unidos e contextualiza a pesquisa sobre eficácia escolar para a realidade dos países ibero-americanos (MURILLO TORRECILLA, 2003, 2005, 2006, 2007, 2008, 2010, 2011 – e colaboradores).

De acordo com Silva, Bonamino e Ribeiro (2012), os estudos brasileiros vêm apresentando resultados convergentes com estudos internacionais sobre a influência positiva nos resultados escolares, proporcionada pela escola e pelos recursos escolares, pelo trabalho docente e interações professor-aluno, pela liderança pedagógica do gestor, comprometimento coletivo com o aprendizado dos alunos, pelo clima escolar e estilo pedagógico dos docentes.

A maioria das pesquisas em eficácia escolar se utiliza de metodologias quantitativas, na medida em que, por meio de modelos estatísticos, buscam estimar o chamado efeito escola, que é um índice que representa o impacto da escola

no desempenho acadêmico dos alunos, depois de controladas as características sociais dos alunos e contextuais das escolas (ALVES; SOARES, 2007). São estudos de alta complexidade técnica, que se valem de análises pautadas em modelos matemáticos de regressão multinível para analisar os fatores escolares não em apenas uma, mas em grupos de escolas ou até mesmo em redes de ensino. Em geral, são desenvolvidos com base nos dados das avaliações em larga escala realizadas pelo governo federal ou pelos governos estaduais (SILVA; BONAMINO; RIBEIRO, 2012).

Entretanto, há pelo menos duas grandes categorias de pesquisas em eficácia, definidas a partir dos seus objetivos de estudo. Uma delas é a que se vale de metodologias quantitativas, composta por estudos que objetivam estimar quantitativamente os efeitos escolares, analisando suas propriedades científicas, à qual se vinculam, por exemplo, as pesquisas de estimação do efeito escola. A outra categoria de estudos, à qual nosso estudo se vincula, apesar de considerar dados quantitativos, possui como objetivo identificar qualitativamente os fatores do contexto das escolas que as tornam uma escola eficaz (MURILLO TORRECILLA, 2003). Isso posto, o Quadro 1 fornece os fatores contextuais de eficácia escolar identificados na escola, a partir da percepção dos respondentes da pesquisa.

Quadro 1 – Fatores contextuais de eficácia escolar mencionados em pesquisa realizada em escola de alta vulnerabilidade social no município de Ferraz de Vasconcelos – São Paulo no ano de 2014, por tipo de categoria de respondentes, frequência discursiva e número de categorias que citaram o fator

Fatores contextuais de eficácia escolar	Categorias de Respondentes					FqD	Ncr
	Professores	Diretor	Alunos	Pais	Funcionários		
Compromisso com a aprendizagem	4	1	4	4	3	16	5/5
Boa Infraestrutura	4	1	3	3	4	15	5/5
Reforço escolar adequado	3	1	2	3	3	12	5/5
Clima escolar harmonioso e disciplinado	3	1	2	2	3	11	5/5
Participação das famílias nas atividades escolares	3	1	-	3	3	10	4/5
Liderança do diretor	2	1	-	2	3	8	4/5
Merenda escolar de boa qualidade	-	1	3	-	2	6	3/5
Materiais didáticos diversificados e de boa qualidade	2	-	2	1	-	5	3/5
Aulas de Educação Física	-	-	4	1	-	5	2/5
Formação Continuada	2	1	-	-	1	4	3/5
Professores experientes	1	-	1	-	1	3	3/5
Professor Auxiliar	1	-	2	-	-	3	2/5
Não haver aulas vagas	-	-	-	3	-	3	1/5
Distribuição de material escolar gratuito	-	-	1	1	-	2	2/5
Bom trabalho desenvolvido pelos funcionários	-	-	2	-	-	2	1/5

Fonte: Elaborado pelos autores.

Legenda: - FqD: Frequência discursiva; Ncr: Número de categorias de respondentes.

As colunas centrais do Quadro 1 registram os números que representam as quantidades de respondentes que mencionaram os fatores. Outra coluna registra a frequência discursiva (FqD), na qual constam os números que representam a quantidade total de respondentes que mencionaram o fator. A última coluna indica o número de categorias de respondentes (Ncr) que mencionaram cada fator, considerando-se as cinco categorias de respondentes: professores, diretor, alunos, pais e funcionários. Dessa forma, por exemplo, para o fator “Compromisso com a aprendizagem”, o Quadro 1 informa que 4 professores, 1 diretor, 4 alunos, 4 pais e 3 funcionários o mencionaram, totalizando uma frequência discursiva de 16 respondentes, dentre as cinco categorias de respondentes (5/5).

No total, foram mencionados 105 fatores que foram sistematizados em 15 fatores contextuais de eficácia escolar, que estão registrados na primeira coluna do Quadro 1. É importante destacarmos que os fatores contextuais foram reunidos a partir de pontos em comum, isto é, quando a análise descritiva de

fatores mencionados com nomenclaturas diferentes revelava que se tratava do mesmo acontecimento. Exemplos: ação, contexto, prática, atividade ou recurso, que foram enquadrados dentro do respectivo fator contextual, e o fator clima escolar harmonioso também foi designado como ambiente disciplinado e tranquilo.

Para melhor analisarmos os Fatores Contextuais de Eficácia Escolar (FCEE), a partir das frequências discursivas, optamos por criar duas categorias: principais FCEE e FCEE de menor incidência discursiva. Os principais FCEE são os que foram mencionados por pelo menos quatro das cinco categorias de respondentes (Ncr = 5/5 ou Ncr = 4/5). Os FCEE de menor incidência discursiva são os que foram mencionados por até três das cinco categorias (Ncr = 1/5, Ncr = 2/5 ou Ncr = 3/5).

ANÁLISE DOS PRINCIPAIS FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR

Tomando-se como referência os dados quantitativos do Quadro 1, construímos o Quadro 2, que apresenta os principais FCEE, hierarquizados da maior para a menor frequência discursiva.

Quadro 2 – Classificação dos principais fatores contextuais de eficácia escolar (FCEE) mencionados em pesquisa realizada em escola de alta vulnerabilidade social no município de Ferraz Vasconcelos – São Paulo no ano de 2014, por categoria de respondentes e frequência discursiva

Principais FCEE	Categorias de Respondentes					FqD	Clas
	Professores	Diretor	Alunos	Pais	Funcionários		
Compromisso com a aprendizagem	4	1	4	4	3	16	1º
Boa Infraestrutura	3	1	3	3	4	15	2º
Reforço escolar adequado	3	1	2	3	3	12	3º
Clima escolar harmonioso e disciplinado	3	1	2	2	3	11	4º
Participação das famílias nas atividades escolares	3	1	-	3	3	10	5º
Liderança do diretor	2	1	-	2	3	8	6º

Fonte: Elaborado pelos autores. Os principais fatores contextuais de eficácia escolar (FCEE) são os que foram mencionados por pelo menos quatro das cinco categorias de respondentes.

Legenda: - FqD: Frequência discursiva; Clas: classificação por ordem dos principais fatores

A análise das opiniões e das percepções dos respondentes evidenciou que o fator “Compromisso com a aprendizagem”, cuja frequência discursiva o caracterizou como o mais mencionado (1º na classificação), se sustentava no fato de que todos os envolvidos no processo se sentiam responsáveis pela aprendizagem dos alunos. Trata-se, pois, de uma responsabilidade compartilhada pelos próprios atores escolares, no sentido da existência de um sentimento comum de responsabilidade com o sucesso da escola.

O fator “Boa Infraestrutura” se materializava no fato de que, para os respondentes, a escola possuía uma boa distribuição dos ambientes escolares, atendendo aos requisitos de acessibilidade e, ainda, boas condições de higiene e de organização. A Unidade contava, segundo eles, com salas de aula arejadas, mobiliário adequado, material de apoio pedagógico, biblioteca e recursos humanos em número suficiente para as necessidades da escola. A sala de informática foi considerada como um ponto positivo na escola, na medida em que era identificada como um importante espaço de aprendizagem. A escola possuía uma quadra de esportes, a qual, apesar de ainda não possuir cobertura (ponto negativo apontado pelos respondentes), foi considerada como um importante espaço de aprendizagem e convívio escolar.

Ainda no quesito infraestrutura, uma característica mencionada como sendo positiva e que certamente é própria de determinadas regiões metropolitanas da Grande São Paulo, foi o fato de que o prédio escolar estava localizado em uma região elevada, onde não havia riscos de alagamentos e enchentes. Isto é, a topografia da região onde o prédio escolar estava localizado era entendida como um ponto positivo para o cotidiano escolar, pois evitava a perda de aulas e garantia o cumprimento dos dias letivos.

O fator “Reforço escolar adequado” foi considerado como sendo de eficácia escolar por apresentar bons resultados no que se refere à recuperação da aprendizagem. As atividades de reforço eram desenvolvidas com um número reduzido de alunos por turma (grupos de até 5 alunos), com ênfase no trabalho de leitura e escrita, por meio de atividades lúdicas, músicas, dinâmicas e atividades diversificadas. Isto é, se aplicavam estratégias de ensino diferentes das utilizadas durante as aulas regulares e as adaptava às necessidades dos alunos com dificuldade de aprendizagem.

Regras de convivência claras e de conhecimento de todos contribuía para a menção ao fator “Clima escolar harmonioso e disciplinado”. Os dados apontam que esse fator, apesar de não ser o de maior frequência discursiva, mostrou-se relevante para o sucesso da escola. Essa constatação decorre do fato de que o clima escolar, em geral, era vinculado ao desenvolvimento eficaz de

inúmeros projetos, ações e atividades da escola. Um aspecto quantitativo que reforça essa nossa análise é o fato de que, em cada categoria de respondente, houve a incidência de pelo menos um indivíduo que mencionou o clima escolar harmonioso e disciplinado como um fator contextual de eficácia escolar.

Quanto ao fator “Participação das famílias nas atividades escolares”, de acordo com os respondentes, era muito significativa a participação dos pais nas atividades escolares, como por exemplo, nas reuniões de pais e mestres e nos eventos promovidos pela escola e no acompanhamento das tarefas de casa. Esse aspecto também evidenciou a existência de uma boa comunicação entre os profissionais da escola e os pais, e o envolvimento das famílias nos assuntos escolares.

Foi mencionado que o diretor possuía liderança tanto para aceitar propostas e apoiar a realização de determinadas ações, quanto para se contrapor a outras, que ele deixava claro que não considerava adequadas. As informações indicaram que ele sempre orientava os professores sobre a importância de se alcançar os objetivos de aprendizagem, coletivamente construídos, bem como de se cumprir as metas projetadas para a escola. Outro aspecto mencionado para justificar o fator “Liderança do diretor” foi que ele possuía habilidade para envolver toda a equipe escolar e membros da comunidade nas atividades da escola nos momentos de tomada de decisão, tanto sobre questões administrativas, quanto pedagógicas.

Conforme Martins (2015), os fatores que influenciam positivamente no desempenho dos estudantes de escolas localizadas em regiões de alta vulnerabilidade social podem ser categorizados nas seguintes dimensões: práticas pedagógicas, gestão escolar, recursos escolares e clima escolar. Nessa ótica, uma outra leitura que podemos fazer dos dados referentes aos principais fatores contextuais de eficácia escolar decorre da análise qualitativa do somatório das suas frequências discursivas, quando categorizados nas dimensões do cenário escolar. O Quadro 3, com as quatro dimensões, apresenta uma síntese das características dos principais fatores contextuais de eficácia escolar.

Quadro 3 – Principais características dos fatores contextuais de eficácia escolar (FCEE) mencionados em pesquisa realizada em escola de alta vulnerabilidade social no município de Ferraz Vasconcelos – São Paulo no ano de 2014, por dimensão no cenário escolar, frequência discursiva e percentual

Dimensão do cenário escolar	Fatores contextuais de eficácia escolar	Principais características dos fatores contextuais	Σ FqD	Perc
Práticas Pedagógicas	Compromisso com a aprendizagem	<ul style="list-style-type: none"> • Sentimento comum de responsabilidade com o sucesso da escola. 	28	38,9%
	Reforço escolar adequado.	<ul style="list-style-type: none"> • Pequenos grupos de até cinco alunos; • Estratégias didáticas altamente diversificadas. 		
Gestão Escolar	Participação das famílias nas atividades escolares	<ul style="list-style-type: none"> • Acompanhamento das tarefas de casa; • Boa comunicação entre os profissionais da escola e os pais; • Envolvimento das famílias em algumas ações de natureza pedagógica. 	18	25%
	Liderança do diretor	<ul style="list-style-type: none"> • Capacidade de envolver a comunidade escolar nas atividades da escola e nas tomadas de decisões, tanto no âmbito administrativo, quanto no âmbito pedagógico; • Agir sempre orientando aos professores sobre a importância dos objetivos e metas da escola. 		
Recursos Escolares	Boa Infraestrutura	<ul style="list-style-type: none"> • Estrutura modesta, mas adequada; • Laboratório de informática; • Biblioteca; • Quadra de esportes; • Espaços de aprendizagem. 	15	20,8%
Clima Escolar	Clima escolar harmonioso e disciplinado	<ul style="list-style-type: none"> • Regras de convivência claras e iguais para todos; • Ambiente disciplinado. 	11	15,3%
Total			72	100%

Fonte: Elaborado pelos autores, a partir de resultados de entrevistas.

Legenda: - FqD: Frequência discursiva; Σ FqD: Somatório de frequência discursiva; Perc: Percentual.

O somatório das frequências discursivas dos fatores contextuais na dimensão práticas pedagógicas representou 38,9% do total, enquanto a frequência discursiva dos fatores relacionados com a dimensão clima escolar representou 15,3% do total da frequência discursiva dos principais FCEE. É uma leitura que confere, de modo mais concreto, posição de destaque ao conjunto de fatores relacionados com as práticas pedagógicas no que tange à relevância desses fatores para a melhoria do IDEB. Diante dessa constatação, podemos inferir que, na percepção dos respondentes, os fatores relacionados com as práticas pedagógicas foram mais relevantes para o sucesso da escola.

ANÁLISE DOS FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR DE MENOR INCIDÊNCIA DISCURSIVA

Cumpre-nos destacar que, apesar de menor incidência discursiva, os fatores listados no Quadro 4 também foram considerados como importantes e fundamentais para o sucesso da escola.

Quadro 4 – Classificação dos fatores contextuais de eficácia escolar (FCEE) de menor incidência discursiva mencionados em pesquisa realizada em escola de alta vulnerabilidade social no município de Ferraz Vasconcelos – São Paulo no ano de 2014, por categoria de respondentes, frequência discursiva e número de categoria de respondentes

Fatores contextuais de eficácia escolar	Categorias de Respondentes					FqD	Ncr	Clas
	Professores	Diretor	Alunos	Pais	Funcionários			
Merenda escolar de boa qualidade	-	1	3	-	2	6	3/5	1º
Materiais didáticos diversificados e de boa qualidade	2	-	2	1	-	5	3/5	2º
Aulas de Educação Física	-	-	4	1	-	5	2/5	3º
Formação Continuada	2	1	-	-	1	4	3/5	4º
Professores experientes	1	-	1	-	1	3	3/5	5º
Professor Auxiliar	1	-	2	-	-	3	2/5	6º
Não haver aulas vagas	-	-	-	3	-	3	1/5	7º
Distribuição de material escolar gratuito	-	-	1	1	-	2	2/5	8º
Bom trabalho desenvolvido pelos funcionários	-	-	2	-	-	2	1/5	9º

Fonte: Elaborado pelos autores. Os principais fatores contextuais de eficácia escolar (FCEE) são os que foram mencionados por pelo menos quatro das cinco categorias de respondentes.

Legenda: - FqD: Frequência discursiva; Ncr: Número de categoria de respondentes; Clas: Classificação por ordem dos fatores com menor incidência discursiva.

Assim como nos quadros anteriores, o Quadro 4 indica, por exemplo, que a merenda escolar de boa qualidade foi mencionada por 1 professor, 3 alunos e 2 funcionários, obtendo uma frequência discursiva igual a 6 (FqD = 6), tendo sido mencionada por três das cinco categorias de respondentes (Ncr = 3/5). A hierarquização desses fatores também foi realizada a partir da frequência discursiva, da maior para a menor.

Para os casos de mesma frequência discursiva, utilizamos como segundo critério de hierarquização (desempate) o número de categorias de respondentes que mencionaram o fator. Assim, por exemplo, o fator de menor incidência discursiva do fator “Materiais didáticos diversificados e de boa qualidade” (FqD = 5) foi considerado o segundo mais relevante, pois foi mencionado por três categorias de respondentes (Ncr = 3/5), enquanto o fator “Aulas de educação física” (FqD = 5) foi o terceiro, pois foi mencionado por apenas duas categorias de respondentes (Ncr = 2/5).

Na classificação dos Fatores contextuais de eficácia escolar (FCEE) de menor incidência discursiva mencionados (Quadro 4), a “merenda escolar de boa qualidade” foi o mais destacado fator de eficácia escolar, na medida em que havia um cardápio bem variado na escola, composto por frutas, verduras, legumes, arroz, feijão, carne, sucos naturais, leite e iogurtes. Lembrando que a escola estava localizada em uma região de alta vulnerabilidade social, atendendo a alunos de famílias de baixa renda e de extrema pobreza, a única refeição substancial do aluno, em muitos casos, era a merenda escolar. Por mais tosca ou ingênua que possa parecer a informação de que a merenda escolar é importante para que muitos alunos consigam se concentrar nas atividades escolares, no caso da escola pesquisada, onde um percentual significativo dos alunos não possuía acesso a uma alimentação digna em suas residências, a merenda escolar de boa qualidade se revelou como o fator contextual de eficácia escolar de menor incidência discursiva mais relevante para aquela realidade.

“Materiais didáticos diversificados e de boa qualidade”, planejados na oferta de livros didáticos e paradidáticos, materiais lúdicos, jogos pedagógicos, utilização de *softwares* educacionais e outras tecnologias com foco na aprendizagem, também foram considerados como fatores contextuais. As “aulas de educação física” foram mencionadas, na medida em que, na percepção dos respondentes, por meio de atividades lúdicas e desportivas, se preparava os alunos para os cuidados com a saúde, melhorando suas condições físicas e mentais, impactando positivamente nas atividades de leitura, escrita e cálculo.

No âmbito da “formação continuada”, os professores mencionaram que participavam regularmente de cursos de capacitação profissional, com a finalidade de melhorar o domínio dos conteúdos e a prática docente. Além dos cursos, indicaram as reuniões pedagógicas como momentos privilegiados para as atividades de formação. Para além da realização de cursos de capacitação ou de formação em serviço, a formação continuada foi considerada como um fator de eficácia escolar, porque estava diretamente relacionada com a melhoria

do ensino, na medida em que, nos momentos de formação, os objetivos de aprendizagem e as metas de desempenho da escola eram questões presentes em diversas atividades formativas.

O fato de a escola contar com “professores experientes” também foi considerado como sendo um fator de eficácia escolar. Todos os professores eram efetivos, e muitos contavam com mais de dez anos de experiência no magistério, sendo que alguns eram professores da escola por mais de cinco anos. A experiência de magistério, na opinião dos respondentes, contribuía para que os professores ministrassem as aulas com mais capacidade técnica, auxiliando e acompanhando de perto a aprendizagem e o desenvolvimento dos alunos. Outra característica positiva com a experiência dos professores foi o fato de que conseguiam estabelecer um bom relacionamento com as famílias dos alunos e com a comunidade em geral.

Um importante fator contextual de eficácia escolar de menor frequência discursiva foi o “professor auxiliar”, que era um professor contratado para ajudar o professor titular da classe durante as aulas. Configurava-se como um recurso humano que a escola dispunha para auxiliar e apoiar os professores, especialmente no atendimento pontual a alunos com dificuldades de aprendizagem, inclusive nas atividades de recuperação e reforço.

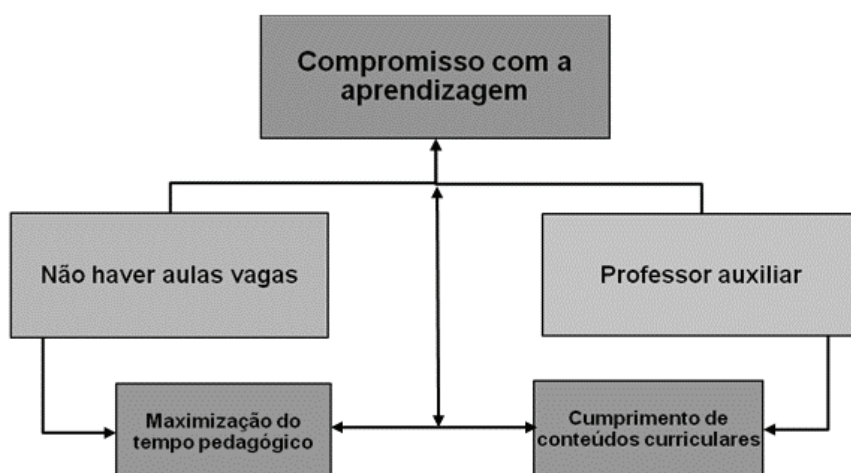
Em determinadas situações de imprevistos em que o professor titular faltasse e enquanto não houvesse substituto, era o professor auxiliar quem conduzia as atividades na classe. Essa dinâmica impactava positivamente na rotina escolar, pois era pouco provável que os alunos ficassem com aulas vagas, o que maximizava o tempo pedagógico e fortalecia as condições para a não interrupção do desenvolvimento de conteúdos curriculares. Nesse contexto, “não haver aulas vagas” foi outro fator contextual mencionado, justamente porque maximizava o tempo pedagógico e o cumprimento de conteúdos curriculares, além de estimular a frequência regular dos alunos.

Tais procedimentos proporcionavam maior continuidade das atividades pedagógicas, na medida em que os alunos não realizavam atividades sem vínculo direto com os objetivos da aprendizagem como, por exemplo, ficar brincando na quadra de esporte da escola isentos da mediação de um adulto ou mesmo retornar para suas casas devido à falta de professor. A importância do professor auxiliar e de não haver aulas vagas ganharam concretude na pesquisa e pode ser corroborada pela literatura científica, pois há autores que são enfáticos sobre essa questão, na medida em que:

Os alunos devem dispor do máximo de tempo possível para aprender, no sentido mais amplo da palavra. As interrupções, as desarticulações curriculares e pedagógicas, a deficiente organização dos tempos ou dos espaços, são fatores que perturbam o funcionamento das escolas. A capacidade de otimizar o tempo é uma característica das escolas de qualidade (NÓVOA, 1999, p. 27).

A articulação entre esses dois fatores foi um importante elemento no sucesso da escola, pois impactava diretamente no fator que foi considerado como o mais relevante na escola: o compromisso com a aprendizagem. Sem dúvida, a pesquisa constatou que, naquela realidade contextual, o trabalho desenvolvido pelo professor auxiliar, articulado com outras ações em prol de se garantir que não houvesse aulas vagas, fortalecia o compromisso com a aprendizagem na escola. A Figura 1 ilustra essa articulação de fatores observada na escola.

Figura 1 – Articulação entre os Fatores contextuais de eficácia escolar (FCEE) de menor incidência discursiva “não haver aulas vagas” e “professor auxiliar”, favorecendo o fator de maior incidência “compromisso com a aprendizagem.”



Fonte: Elaborada pelos autores, a partir de pesquisa realizada em escola de alta vulnerabilidade social no município de Ferraz de Vasconcelos – São Paulo no ano de 2014.

A “distribuição de material escolar gratuito” também foi um fator mencionado. No início do ano letivo, a rede municipal disponibilizava, gratuitamente, mochilas que eram entregues na escola, contendo produtos de material escolar: lápis, borracha, caneta, lápis de cor, cadernos, régua, entre outros itens básicos para os estudos. Além de suprir parte da carência de materiais de muitos alunos, cujas famílias não tinham dinheiro suficiente para comprá-los, de certo modo, essa ação reforçava os laços de confiança e respeito mútuo entre as famílias e a equipe escolar.

O trabalho desenvolvido pelos funcionários foi considerado, por dois alunos, como um fator de eficácia escolar, pois, na percepção deles, os funcionários eram pessoas dedicadas ao serviço e tratavam os alunos com atenção e respeito, inclusive nos momentos em que era necessário repreendê-los pelo descumprimento de alguma regra de convivência estabelecida pela escola.

CONSIDERAÇÕES FINAIS

O conceito de fatores contextuais de eficácia escolar que utilizamos se alinha ao apontado por Muijs (2003) no sentido de que os estudos em eficácia escolar têm mostrado que, embora muitos elementos de melhoria desenvolvidos nas escolas eficazes sejam semelhantes aos que produzem melhorias em qualquer outro contexto, há fatores escolares e boas práticas que são específicos de algumas escolas, em determinados contextos, inclusive para aquelas localizadas em regiões de alta vulnerabilidade social.

Entretanto, como respostas à nossa questão norteadora da pesquisa, podemos registrar que o agir com compromisso com a aprendizagem foi o fator de eficácia mais relevante para o sucesso da escola. Quando se toma como referência a frequência discursiva dos principais fatores contextuais, os relacionados com as práticas pedagógicas foram os que mais impactaram no desempenho escolar, evidenciando que, para aquela comunidade escolar, tais aspectos eram os mais relevantes para o alcance dos objetivos de aprendizagem, a obtenção de alto desempenho e cumprimento das metas projetadas para o IDEB nos anos de 2007, 2009 e 2011.

Como dois importantes fatores de eficácia escolar, próprios da escola pesquisada, destacamos “não haver aulas vagas” e existência do “professor auxiliar”, os quais, bem articulados pela equipe escolar, configuravam-se como um dos pilares de sustentação para o mais relevante fator apontado pelos respondentes: o “compromisso com a aprendizagem”. O fato de não haver aulas vagas foi considerado como promotor de maximização do tempo pedagógico, que por sua vez, fortalecia o cumprimento de conteúdos curriculares. A atuação do professor auxiliar, que no Brasil é mais comum em escolas privadas do que em escolas públicas, mas que, gradativamente, vem ganhando concretude nas políticas educacionais e sendo incorporado nos quadros de recursos humanos das escolas (MARTINS, 2015), impactava positivamente nas atividades escolares, principalmente porque evitava a existência de aula vaga.

No que se refere aos fatores relacionados com a gestão escolar, destacaram-se a liderança do diretor e o envolvimento das famílias nas atividades esco-

lares, como elementos articuladores de importantes ações. A formação continuada, apesar de ter obtido menor frequência discursiva, foi mencionada por pelo menos um membro de cada categoria de respondente da equipe escolar. A importância desse fator reside no fato de que as ações formativas estavam diretamente relacionadas com os objetivos e com as metas de aprendizagem.

Os fatores associados à dimensão dos recursos escolares representaram uma parcela significativa, os quais impactavam no desempenho escolar (40% dos respondentes os identificaram), convergindo para o apontado na literatura científica que destaca que, no Brasil, uma considerável parcela da explicação para o desempenho escolar satisfatório deve-se à existência de recursos escolares suficientes, pois, ao contrário do que se observa em muitos países, no nosso país ainda há uma “variabilidade nos recursos escolares com que contam as escolas” (FRANCO, BONAMINO, 2005, p. 44).

O clima escolar, apesar de não ter apresentado elevada frequência discursiva, foi mencionado como de grande importância para o sucesso da escola, pois foi considerado como um dos principais fatores contextuais de eficácia escolar e um dos pilares de sustentação para o desenvolvimento eficaz das atividades escolares.

Em suma, a análise dos fatores contextuais de eficácia escolar mostrou que o sucesso da escola nos resultados do IDEB nos anos de 2007, 2009 e 2011 foi o resultado não de um ou outro fator isoladamente, mas sim de um conjunto articulado de ações, recursos e práticas desenvolvidas por todos os atores escolares. Nessa ótica, não somente os principais fatores contextuais de eficácia escolar, mas também os de menor incidência discursiva foram fundamentais nas tarefas empreendidas naquela unidade escolar e que a levaram ao sucesso, evidenciando peculiaridades da escola e reafirmando que determinados fatores escolares podem ser decisivos para a melhoria da aprendizagem e, conseqüentemente, do desempenho escolar.

Os dados apresentados, sem dúvida, permitem outras análises, pois não se esgotam no que evidenciamos e nas análises aqui realizadas. Escudero Muñoz (2009), de modo muito assertivo, aponta que há fatores e práticas escolares que podem ser considerados como fatores de eficácia ou como boas práticas escolares para determinadas escolas, em determinados contextos, e talvez não o ser para outras. Nesse sentido, ao apresentarmos uma listagem de fatores escolares que, no contexto daquela instituição de ensino, foram considerados como de eficácia escolar, não significa afirmar que tais fatores são os que, necessariamente, poderão ou deverão se configurar como fatores de eficácia escolar em outras realidades contextuais.

REFERÊNCIAS

- ALVES, M. T. G.; SOARES, J. F. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. **Educação e Pesquisa**, São Paulo, v. 39, n. 1, p. 177-194, jan./mar. 2013.
- ALVES, M. T. G.; SOARES, J. F. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a Sociologia da Educação. **Sociedade e Estado**, v. 22, n. 2, p. 435-473, maio/ago, 2007.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Resultados e Metas do IDEB**. INEP. Disponível em: <http://ideb.inep.gov.br/resultado/resultado/resultado.seam?cid=248028>. Acesso em: 14 abr. 2013.
- BRASIL. Lei no. 9.394, de 20 de dezembro de 1996. **Estabelece as diretrizes e bases da educação nacional (LDB)**. Brasília, DF: 1996.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Senado Federal: Centro Gráfico, 1988, 292p.
- BROOKE, N.; SOARES, J. F. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: EdUFMG, 2008.
- CALDERÓN, A. I. Usos e apropriações das avaliações em larga escala: tensões e desafios. In: Amanda Sangy Quiossa, Diovana Paula de Jesus Bertolotti, Luiz Flávio Neubert, Priscila Campos Cunha. (Org.). **Planos de ação para a rede estadual de ensino de Minas Gerais**, 1ed. Juiz de Fora: CAEd-FADEPE/JF, v. 2, p. 31-45, 2017.
- CHIZZOTTI, A. **Pesquisa qualitativa em ciências humanas e sociais**. Editora Vozes: Petrópolis, 2006.
- ESCUADERO MUÑOZ, J. M. Buenas prácticas y programas extraordinarios de atención al alumnado en riesgo de exclusión educativa. **Revista de Currículum y Formación de Profesorado**, Madri, vol. 13, n. 3, p. 107-141, 2009.
- FIGUEIREDO, D. et al. Os cavalos também caem: Tratado das inconsistências do IDEB. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 26, n. 100, jul. 2018.
- FRANCO, C.; BONAMINO, A. A pesquisa sobre características de escolas eficazes no Brasil: breve revisão dos principais achados e alguns problemas em aberto. **Revista Educação**. Rio de Janeiro: PUC-Rio, n. 1, 2005.
- FREITAS, L. C. de. Ciclos de Progressão Continuada: vermelho para as políticas públicas. **Eccos – Revista Científica**, São Paulo, v. 4, n.1, p. 79-93, jun. 2002.

_____. **Ciclos, seriação e avaliação:** confrontos de lógicas. São Paulo: Moderna, 2003.

HORTA NETO, J. L. Avaliação educacional no Brasil para além dos testes cognitivos. **Revista de Educação PUC Campinas**, Campinas, v.23, n.1, p. 37-53, jan./abr. 2018.

JANUARIO, A. A. **Pesquisa em eficácia escolar no Brasil:** um estudo sobre a produção científica brasileira com foco nos fatores relacionados à eficácia escolar (2005-2017). 2019. 141f. Dissertação (Mestrado em Educação) – Pontifícia Universidade Católica de Campinas, Campinas, 2019.

MARTINS, E. C. C.; CALDERÓN, A. I. Eficácia escolar: boas práticas à luz de estudos do governo brasileiro e das agências multilaterais. **Revista Diálogo Educacional**, Curitiba, v. 19, n. 62, p. 1297-1327, jul./set. 2019.

_____. Boas práticas e elevado desempenho escolar em contexto de vulnerabilidade social com referência aos resultados do IDEB. **Revista Educação em Debate**, Fortaleza, anos 35-38, v. 38, n. 66-71 p. 130-144, jan./jul. 2016.

_____. Boas práticas escolares e avaliação em larga escala: a literatura ibero-americana em questão. **Estudos em Avaliação Educacional**, São Paulo, v. 26, n. 62, p. 264-293, maio/ago. 2015.

MARTINS, E. C. C. **Construindo uma escola eficaz:** boas práticas escolares e fatores de alto desempenho em escolas de alta vulnerabilidade social. 2015. 224f. Dissertação (Mestrado em Educação) – Centro de Ciências Humanas, Sociais e Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, 2015.

MUIJS, D. La mejora y la eficacia de las escuelas en zonas desfavorecidas: resumen de resultados de investigación. REICE - **Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, Madrid, v. 1, n. 2, 2003.

MURILLO TORRECILLA, F. J. (Coord.) et al. **La investigación sobre Eficacia Escolar en Iberoamérica.** Revisión internacional del estado del arte. Bogotá: Convenio Andrés Bello, 2003.

_____. (Coord.). **Estudios sobre Eficacia Escolar en Iberoamérica:** 15 buenas investigaciones. Bogotá: Convenio Andrés Bello, 2006.

_____. (Coord.) et al. **Investigación Iberoamericana sobre Eficacia Escolar.** Bogotá: Convenio Andrés Bello, 2007.

MURILLO TORRECILLA, F. J. Mejora de la Eficacia Escolar en Iberoamérica. **Revista Iberoamericana de Educación**, Madrid, v. 55, p. 49-83, 2011.

_____. Lecciones aprendidas de la investigación sobre Eficacia escolar en América Latina. **Educación y Ciudad**. Bogotá, v.19, p. 7-17, 2010.

_____. Enfoque, situación y desafíos de la investigación sobre eficacia escolar en América Latina y el Caribe. In: UNESCO. **Eficacia escolar y factores asociados en América Latina y el Caribe**. Santiago de Chile: UNESCO, 2008, p. 17-47.

_____. La Investigación en Eficacia Escolar y Mejora de la Escuela como motor para el incremento de la Calidad Educativa en Iberoamérica. **Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, Madrid, v. 3, n. 2, 2005.

MURILLO TORRECILLA, F. J. HERNÁNDEZ CASTILLA, R. La equidad en la investigación Sobre eficacia escolar. **Profesorado, Revista de Currículum y Formación de Profesorado**, Granada, v. 15, n. 3, p. 1-8, dec. 2011.

NÓVOA, A. (Coord.) **As organizações escolares em análise**. Lisboa: Dom Quixote, 1999.

PESTANA, M. I. Trajetória do Saeb: criação, amadurecimento e desafios. **Em Aberto**, Brasília, v. 29, n. 96, p. 71-84, maio/ago. 2016

PONTES, L. A. F.; SOARES, T. M. As metas escolares do Ideb: uma proposta alternativa de cálculo. **Estudos em Avaliação Educacional**, São Paulo, v. 27, n. 66, dez. 2016.

SILVA, I. F. Origem e evolução do paradigma da escola eficaz e seus desdobramentos no contexto atual. **Perspectiva**, Florianópolis, v. 33, n. 2, p. 707-738, maio/ago. 2015

SILVA, J. L.; BONAMINO, A. M. C.; RIBEIRO, V. M. Escolas eficazes na educação de jovens e adultos: estudo de casos na rede municipal do Rio de Janeiro. **Educação em Revista**, Belo Horizonte, v.28, n.2, p.367-392, jun. 2012.

SOARES, J. F. et al. **Escola Eficaz**: um estudo de caso em três escolas da rede pública de ensino do Estado de Minas Gerais. Belo Horizonte: UFMG, FAE, GAME: Fundação Ford, 2002.

Recebido em: 09 de dezembro de 2019

Aceito em: 20 de junho de 2020

Publicado em: 30 de junho de 2020

ESCALAS ESPECÍFICAS PARA MEDIR A INFRAESTRUTURA ESCOLAR DA EDUCAÇÃO BÁSICA

Joaquim José Soares Neto¹

Elianice Silva Castro²

¹Professor Titular do Programa de Pós-Graduação em Desenvolvimento, Sociedade e Cooperação Internacional do Centro de Estudos Avançados Multidisciplinares (Ceam) e no Mestrado Nacional Profissional em Ensino de Física da Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil. E-mail: jjsoaresneto@gmail.com.

²Doutora em Desenvolvimento, Sociedade e Cooperação Internacional pela Universidade de Brasília, Brasília, Distrito Federal, Brasil. E-mail: elianice.castro@gmail.com.

Resumo

A fim de orientar políticas públicas de investimento para educação básica brasileira, este estudo propõe escalas de infraestrutura específicas para cada um dos segmentos de ensino, identificando uma infraestrutura mínima a cada um desses ambientes escolares. Foram definidas constitutivamente e operacionalmente as características da infraestrutura escolar de segmento de ensino; estimados os parâmetros dos itens; definidos e descritos os níveis e das escalas e realizadas análises descritivas.

Palavras-chave: Infraestrutura escolar. Educação básica. Indicadores educacionais.

Abstract

In order to guide public investment policies for Brazilian basic education, this study aimed to propose three specific school infrastructure scales for each of the education segments, identifying a minimum infrastructure for each of these school environments. Thus, the infrastructure of each segment was constitutively and operationally defined; estimation of item parameters; definition of levels and description of scales and descriptive analysis.

Keywords: School infrastructure. Basic education. Educational indexes.

INTRODUÇÃO

A igualdade e equidade constituem um dos eixos de pesquisa na área de eficácia escolar e, conforme Reynolds et al. (2014), existe uma tendência internacional em associar a eficácia de sistemas educacionais à qualidade em educação. Nesse sentido, a eficácia escolar, ao englobar qualidade e equidade, compreende o quanto o desempenho de uma escola é melhor do que o de outras e o quanto a escola é capaz de compensar as diferenças de entrada dos estudantes (SAMMONS, 1995; REYNOLDS et al., 2014).

No Brasil, as pesquisas sobre eficácia escolar também se orientam pela perspectiva de análise da igualdade e da equidade, com o objetivo de identificar variáveis do nível da escola que diminuam a interferência de variáveis do nível do estudante nos resultados escolares. Por essência, essas pesquisas buscam compreender como as diferenças de entrada dos estudantes, relacionadas a gênero, etnia e nível socioeconômico, e as diferenças de recursos e dos processos escolares – infraestrutura escolar, formação de professores e gestão escolar – podem influenciar no desempenho e na diminuição das desigualdades educacionais (ALBANEZ et al., 2002; ANDRADE et al., 2003; SOARES; ALVES, 2003; SOARES, 2004; SOARES; ANDRADE, 2003; CASTRO, 2014; KARINO, 2016). Via de regra, os resultados desses estudos indicam que os fatores que geram eficácia escolar têm contribuído para o aumento das desigualdades no que se refere ao desempenho dos diferentes grupos de estudantes.

A fim de contribuir com a identificação desses níveis de desigualdade, especificamente no que se refere aos recursos escolares do país, Neto et al. (2013) propuseram um indicador capaz de aferir a infraestrutura escolar da educação básica brasileira. Nessa escala da infraestrutura, que foi subdividida em quatro níveis (elementar, básica, adequada e avançada), todas as escolas receberam um “score” interpretável e que expressa um significado prático e objetivo do que constitui a sua infraestrutura escolar. Esse estudo, além de inovar ao interpretar os níveis de infraestrutura escolar do país, apresentou como resultado indícios de grande desigualdade de recursos entre as escolas, especialmente entre aquelas localizadas nas zonas rurais e nas regiões Norte e Nordeste (NETO et al., 2013).

Esse indicador subsidiou outros estudos que também buscavam avaliar a influência da infraestrutura sobre o desempenho e no aumento das desigualdades. Castro (2014) utilizou a escala de infraestrutura escolar como uma das variáveis de processo associada ao desempenho e identificou que escolas que possuem nível socioeconômico baixo e naquelas em que, mais de 80%

dos estudantes atendidos são beneficiários do Programa Bolsa Família (PBF), a infraestrutura e o nível de escolaridade do docente exercem o mesmo nível de influência no desempenho dos estudantes. Em contrapartida, em escolas cujo percentual de estudantes beneficiários do programa é de apenas 20%, a infraestrutura tem um poder de influência no desempenho escolar muito maior em relação à formação docente, mas similar em relação ao nível socioeconômico. Nessa mesma linha de investigação, Karino (2016) verificou que a chance de sucesso acadêmico de um estudante pobre é de aproximadamente um terço da chance de um estudante de outro nível socioeconômico.

A fim de subsidiar uma investigação nesse campo, Castro, Karino e Neto (2018, no prelo) analisaram a série histórica da escala de infraestrutura escolar, no espaço de tempo compreendido entre 2007 e 2017. Um dos resultados apresentados é o inexpressivo aumento de 0,8% de escolas da educação básica com infraestrutura escolar adequada nesses onze anos, nível representado por apenas 21,4% das escolas do país em 2017. As escolas adequadas são caracterizadas pela existência de espaços como sala de professores, biblioteca, laboratório de informática e sanitário para Educação Infantil (EI), quadra esportiva e parque infantil, além de possuírem equipamentos complementares, como copiadora e acesso à internet.

Os autores apontam a necessidade de novos indicadores que considerem as especificidades dos itens para cada um dos segmentos de ensino da educação básica, pois a atual escala parte do pressuposto de que todas as variáveis³ possuem a mesma adequabilidade para todos os segmentos. Todavia, verificam-se possíveis desajustes entre a relevância dos itens e o nível de dificuldade na composição da escala de infraestrutura escolar, quando eles são analisados em consonância com o seu posicionamento. Em outras palavras, itens que caracterizam escolas com infraestrutura adequada, por exemplo, não necessariamente são representativos para todos os segmentos de ensino. Essa constatação indica, sobretudo, uma limitação para a movimentação das escolas na escala de infraestrutura escolar, considerando o princípio da cumulatividade pressuposto pela Teoria de Resposta ao Item (TRI). Nesse indicador, observa-se que a viabilidade de movimentação pode estar relacionada muito mais ao segmento de ensino que as escolas oferecem do que ao traço latente observado, que são os itens que compõem a sua infraestrutura.

No intuito de superar essa barreira, o presente estudo propõe três indicadores de infraestrutura escolar, específicos para cada um dos segmentos da educa-

³Das 34 variáveis inicialmente consideradas, cinco foram submetidas a um tratamento diferenciado: parque infantil, sanitário para EI e berçário (somente para escolas que ofertam EI); laboratório de ciências (somente para as escolas que ofertam Ensinos Fundamental e Médio); e cozinha (somente para as escolas públicas, uma vez que nelas a alimentação escolar é obrigatória). Para as escolas às quais essas variáveis não se adequavam, estas foram tratadas nas análises da TRI como “não se aplica”.

ção básica (EI, Ensino Fundamental – EF e Ensino Médio – EM), construídos com base no princípio da validade das medidas e focando especificamente nas evidências, com base no conteúdo, que definem o traço latente observado. O enfoque dado às evidências de validade dos indicadores propostos faz-se necessário diante da necessidade de verificar se as medidas por eles produzidas constituem uma representação adequada do construto mensurado, apontando se esse conceito está sendo representado por todas as suas definições constitutiva e operacional.

Assim, pretende-se identificar a especificidade do conceito infraestrutura escolar com base em teorias que sustentem as escalas específicas para cada um dos segmentos de ensino, superando a perspectiva até então estabelecida de que a infraestrutura é fundamentada exclusivamente por variáveis operacionais e pelas medidas produzidas por indicadores, fazendo com que a definição constitutiva do conceito seja mais ampla do que sua definição operacional. De posse dessas ferramentas, pretende-se identificar e mensurar as desigualdades e inequidades na infraestrutura escolar para o EI, o EF e o EM. Este estudo ainda objetiva verificar as diferenças dos níveis de infraestrutura dos diferentes segmentos de ensino da educação básica e demonstrar os problemas de equidade a que estão expostos os estudantes que têm acesso a escolas com poucos recursos.

Método

A análise dos dados para a construção dos indicadores de infraestrutura escolar de cada um dos segmentos de ensino da educação básica ocorreu em quatro etapas: (1) definição constitutiva e operacional da infraestrutura escolar e tratamento das variáveis, considerando a sua pertinência para cada um dos segmentos; (2) estimação dos parâmetros dos itens para cada segmento, por meio do Modelo de Respostas Graduais (MRG) da TRI, utilizando os bancos do Censo Escolar de 2017; (3) definição e descrição das escalas de infraestrutura escolar para a EI, o EF e o EM; e (4) análise descritiva para cada uma das escalas, considerando a região, a dependência administrativa e a localização das escolas de cada um dos segmentos de ensino.

Definição constitutiva e operacional da infraestrutura escolar e tratamento das variáveis

A literatura indica que o conceito de infraestrutura, de modo mais amplo, está associado ao *locus* privilegiado do processo educacional de transmissão de valores culturais e conhecimentos historicamente constituídos, independentemente se sua estrutura é planejada e adequada ou um espaço adaptado (SASTRE, 2010). No presente estudo, o conceito de infraestrutura escolar será

compreendido, de modo constitutivo, enquanto a materialidade da escola, o prédio e os recursos e materiais disponíveis na instituição escolar (KOWALTO-WSKI, 2011). Para a sua definição operacional, será considerada a matriz de especificação da infraestrutura escolar, apresentada no Quadro 1, composta por cinco fatores estruturais: infraestrutura básica; infraestrutura pedagógica, infraestrutura predial; equipamentos de apoio pedagógico; e adequação a necessidades especiais (CASTRO, 2018, no prelo).

Quadro 1 – Matriz de especificação da infraestrutura escolar

Variáveis do censo escolar	Definição do conceito (referencial teórico)
INFRAESTRUTURA BÁSICA	
Abastecimento de água Abastecimento de energia elétrica Esgoto sanitário Sanitários	Conexão da infraestrutura das escolas com serviços públicos básicos (eletricidade, água potável, sanitários, esgoto, sanitário dentro ou fora do prédio escolar etc.) (SÁTYRO; SOARES, 2007; DUARTE; GARGIULO; MORENO, 2011). Tipo de escola (pública ou privada) (CERQUEIRA; SAWYER, 2007). Tipo de escola (rural ou urbana) (NETO et al., 2013). Condições e instalações físicas das escolas (UNESCO; LLECE, 2008; JESUS; LAROS, 2004; DALLAGNOL, 2015). Prédios e instalações adequados (SOARES; 2007).
INFRAESTRUTURA PEDAGÓGICA	
Quadra esportiva Sala de leitura/biblioteca Laboratório de ciências Laboratório de informática	Áreas acadêmicas e pedagógicas (DUARTE; GARGIULO; MORENO, 2011). Dependências/benfeitorias (SÁTYRO; SOARES, 2007). Disponibilidade de espaços educativos e de atividades recreativas, desportivas e áreas de usos múltiplos (ginásio, auditório, quadra) (DUARTE; GARGIULO; MORENO, 2011; SÁ; WERLE, 2017). Recursos escolares (FLETCHER, 1998). Qualidade da infraestrutura escolar básica e do espaço físico (EI/EF) (MOTA, 2012). Infraestrutura, gestão escolar e desempenho em leitura e matemática (LIMA, 2012).
INFRAESTRUTURA PREDIAL	
Sala da diretoria Sala de professores Cozinha Área verde	Recursos tecno-pedagógicos (JESUS; LAROS, 2004). Áreas acadêmicas e pedagógicas (DUARTE; GARGIULO; MORENO, 2011). Dependências/benfeitorias (SÁTYRO; SOARES, 2007).
INFRAESTRUTURA PARA EQUIPAMENTOS DE APOIO PEDAGÓGICO	
Projetor multimídia (<i>datashow</i>) TV DVD Computador Aparelho de som Acesso à internet Impressora Copiadora Antena parabólica	Infraestrutura tecnológica (SÁ; WERLE, 2017). Inserção de novas práticas pedagógicas (ALLAN, 2011; ASSIS, 2011; CZESZAK, 2011; MENEZES, 2011).
INFRAESTRUTURA ADEQUADA ÀS NECESSIDADES ESPECIAIS	
Sala para atendimento especial Banheiro para Portadores de Necessidade Especiais (PNE) Dependências PNE	Espaços e objetos adaptados (LOPES; CAPELLINI, 2015). Infraestrutura física adequada e ambientes seguros e inclusivos (Marco de Ação da Educação 2030). Educação especial pela ótica da inclusão educacional (MARCÓCCIA, 2011).

Fonte: Castro (2018, no prelo), com adaptações.

Considerando essas dimensões e as variáveis que as compõem, para a composição das escalas, foram analisados todos os 55 itens dos blocos sobre caracterização e infraestrutura e sobre equipamentos do questionário de escola do Censo de 2017. Destaca-se que, nesses blocos de questões, identificou-se um conjunto de variáveis que não são consideradas pela matriz de especificação do conceito e que também não foram consideradas no presente estudo por não serem características da infraestrutura escolar, tais como: destinação do lixo; alojamento de alunos e de professor; lavanderia; retroprojetor; videocassete e internet banda larga.

Outros conjuntos de variáveis que não são considerados na matriz de especificação do conceito, mas que foram incluídos pelo estudo na composição operacional das escalas, estão compreendidos nas seguintes dimensões: infraestrutura básica (forma de ocupação do prédio, local de funcionamento da escola, água consumida pelos alunos); infraestrutura predial (quantidade de salas utilizadas como salas de aula) e infraestrutura pedagógica (pátio coberto e descoberto). O Quadro 2 apresenta todas as variáveis que efetivamente foram consideradas para as análises de calibração dos itens, em cada etapa de ensino.

Quadro 2 – Itens do censo escolar de 2017 utilizados para a composição das escalas dos segmentos de ensino

Item	Descrição	ei	ef	em
1	Forma de ocupação do prédio	Sim	Sim	Sim
2	Local de funcionamento da escola	Sim	Sim	Sim
3	Água consumida pelos alunos	Sim	Sim	Sim
4	Abastecimento de água	Sim	Sim	Sim
5	Abastecimento de energia	Sim	Sim	Sim
6	Esgoto sanitário	Sim	Sim	Sim
7	Sala de diretoria	Sim	Sim	Sim
8	Sala de professor	Sim	Sim	Sim
9	Sala de recursos multifuncionais para Atendimento Educacional Especializado (AEE)	Sim	Sim	Sim
10	Laboratório de informática	Não	Sim	Sim
11	Laboratório de ciências	Não	Sim	Sim
12	Cozinha	Sim	Não	Não
13	Biblioteca	Sim	Sim	Sim
14	Sala de leitura	Sim	Sim	Sim
15	Banheiro	Não	Sim	Sim
16	Banheiro de El	Sim	Não	Não
17	Banheiro acessível, adequado ao uso dos alunos com deficiência ou mobilidade reduzida	Sim	Sim	Sim
18	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida	Sim	Sim	Sim
19	Secretaria	Sim	Sim	Sim
20	Banheiro com chuveiro	Sim	Não	Não
21	Refeitório	Sim	Sim	Sim
22	Despensa	Sim	Não	Não
23	Almoxarifado	Sim	Sim	Sim
24	Auditório	Não	Sim	Sim
25	Área verde	Sim	Sim	Sim
26	Número de salas utilizadas como sala de aula (dentro e fora do prédio)	Sim	Sim	Sim
27	Aparelho de televisão	Sim	Sim	Sim
28	Aparelho de DVD	Sim	Sim	Sim
29	Copiadora	Sim	Sim	Sim
30	Impressora	Sim	Sim	Sim
31	Impressora multifuncional	Sim	Sim	Sim
32	Aparelho de som	Sim	Sim	Sim
33	Projetor multimídia (<i>datashow</i>)	Sim	Sim	Sim
34	Máquina fotográfica/Filmadora	Sim	Sim	Sim
35	Computador	Sim	Sim	Sim
36	Computadores para uso administrativo	Sim	Sim	Sim
37	Computadores para uso dos alunos	Sim	Sim	Sim
38	Acesso à internet	Sim	Sim	Sim
39	Pátio coberto e descoberto	Sim	Sim	Sim
40	Quadra de esporte coberta e descoberta	Não	Sim	Sim
41	Parque infantil	Sim	Não	Não
42	Berçário	Sim	Não	Não

Fonte: Elaboração da autora.

Essa análise inicial verificou a relevância de todas as variáveis para cada uma das etapas de ensino, bem como o tipo de tratamento necessário para aquelas que são politômicas, ou seja, que são de múltipla escolha ou que indicam a quantidade dos recursos avaliados. Para todas as essas variáveis de múltipla

escolha ou quantitativas, foram criadas categorias ou intervalos de valores como opções de respostas, a fim de considerar a qualidade ou a quantidade avaliada pelo item e não apenas o atributo da existência, como é estimado pelos itens dicotômicos.

Nas análises de respostas graduais pela TRI, os itens com maior desejabilidade foram classificados como nível 1, que representa o recurso ou o nível de qualidade desejável para todas as escolas. Por exemplo, é desejável que todas as escolas da educação básica, ao responderem ao questionário do Censo, tenham indicado que possuem abastecimento de água por meio da rede pública. Para as variáveis quantitativas, as categorias foram estabelecidas com base na análise de frequência das respostas, considerando que elas fossem quantitativamente comparáveis em cada uma das categorias e para cada um dos segmentos de ensino analisado.

Estimação dos parâmetros dos itens para cada segmento de ensino

Para a realização desta etapa do estudo, foi acessada a base de dados do Censo Escolar de 2017 com 285.540 escolas, incluindo públicas e privadas, rurais e urbanas e de todos os segmentos de ensino. O primeiro filtro feito nessa base foi na variável *Situação de Funcionamento*, considerando apenas as escolas “Em Atividade”. Após essa filtragem, a base ficou reduzida a 181.611 escolas.

Em seguida, foi aplicado um filtro na variável “Etapa de Ensino”, com a finalidade de considerar apenas as escolas que autodeclararam ofertar exclusivamente cada um dos segmentos de ensino. Esse procedimento foi adotado a fim de se garantir que as respostas em cada uma das alternativas das variáveis fossem coerentes com as diferentes realidades das escolas exclusivas daquela etapa de ensino e, conseqüentemente, permitissem uma melhor calibração dos parâmetros. Assim, chegou-se a três bancos específicos para cada etapa: EI, com 83.052 escolas; EF, com 100.899; e EM, com 24.825.

Feito esse tratamento, passou-se à segunda etapa da pesquisa, que é a análise inicial utilizando o *software* livre R, com o auxílio dos pacotes *psych* e *mirt*, versão 3.3.0 para estimação dos parâmetros dos itens, também conhecida como calibração. O processo de calibração dos itens foi realizado via TRI, para itens construídos em escala gradual, que consistiu em estimar conjuntamente os parâmetros de um item. Essas estimativas foram produzidas com base na distribuição de frequências dos padrões de respostas dos respondentes (SAMEJIMA, 1969). Para tanto, foi utilizado o MRG, que trabalha com itens que apresentam mais de duas opções de respostas em um formato gradativo e

que normalmente exigem do respondente uma avaliação sobre o assunto, para a escolha das alternativas.

O modelo MRG permite calcular a probabilidade de uma escola com determinado nível de infraestrutura responder a uma alternativa a partir dos parâmetros de discriminação dos itens e de dificuldade das alternativas. Ou seja, este modelo assume que os escores das categorias de um item i são arranjados em ordem do menor para o maior e denotados por $k = 0, 1, \dots, m_i$ onde $(m_i + 1)$ é o número de categorias do i – *ésimo* item. A probabilidade de um respondente j , com score θ_j , escolher uma categoria ou outra mais alta do item i é dada por:

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-D a_i(\theta_j - b_{i,k})}}$$

$b_{i,k}$: é o parâmetro de dificuldade da k – *ésima* categoria do item i . Esse parâmetro refere-se à probabilidade de um respondente com determinado score selecionar uma categoria de resposta (por exemplo, discordo totalmente) ou uma categoria mais alta ordenada (2, 3 ou concordo totalmente).

θ_j : score do j – *ésimo* respondente.

a_i : é o parâmetro de discriminação do item i , com o valor proporcional à inclinação da Curva Característica do Item (CCI) no ponto b_i . Este parâmetro representa o quanto um item discrimina entre os respondentes de diferentes scores.

D : é um fator de escala constante e igual a 1. Utiliza-se o valor 1,7 quando se deseja que a função logística forneça resultado semelhante ao da função ogiva normal.

Assim, nesse modelo, o parâmetro de discriminação “ a ” representa a capacidade de distinguir escolas com diferentes níveis de traço latente. Ou seja, diferencia as escolas que avaliaram os itens dos questionários como não existentes daquelas que apontam possuir os recursos, indicando, ainda, em que intensidade eles existem em suas dependências. Teoricamente, essa discriminação pode variar de $-\infty$ a $+\infty$. Porém, na prática, o valor da discriminação varia de 0 (zero), considerada uma discriminação muito baixa, a 4 (quatro), considerada uma discriminação muito alta.

O parâmetro de dificuldade “ b ”, no MRG, está associado à probabilidade das escolas com determinado nível de infraestrutura selecionar uma ou outra categoria/alternativa de resposta (por exemplo, “inexistente” ou “rede pública”).

É com base nos parâmetros de dificuldade dos itens, representados em uma escala que varia, em geral, de -4 a 4, que as categorias de resposta aos itens são estabelecidas. Para cada item do questionário, o número de parâmetros de dificuldade estimados é igual ao número de alternativas de resposta.

Por exemplo, no item 4 “Abastecimento de água”, os respondentes do questionário deveriam indicar, entre as cinco opções de respostas, qual melhor representava a proveniência do abastecimento de água para a escola. Considerando o nível de desejabilidade da resposta, no caso do item 4, as escolas que confirmam a proveniência do abastecimento de água via rede pública foram classificadas no nível 1. Quando elas não possuem abastecimento via rede pública, espera-se que tenham acesso a outros tipos de abastecimento (poço artesiano, cacimba/cisterna/poço ou de fonte/rio/igarapé/riacho/córrego), resposta classificada no nível 2 ou 3. Por fim, aquelas escolas que não têm abastecimento de água foram classificadas no nível 4. Assim, a probabilidade de uma escola optar pela alternativa “rede pública”, representada pelo nível 1, aumenta à medida que o seu traço latente ou nível de infraestrutura também aumenta.

As Tabelas 1 a 3 apresentam os resultados dessa análise, com os itens que irão compor as escalas de infraestrutura escolar específicas para cada um dos segmentos de ensino e seus parâmetros de discriminação e de dificuldade na escala (0,1), ou seja, em uma escala com média 0 e desvio padrão 1. Para a interpretação das tabelas, deve-se considerar os critérios definidos por Baker (2001), em que os parâmetros de discriminação “*a*” variam de 0,35 (considerados baixos) a 1,70 (considerados muito altos). Em relação aos parâmetros de dificuldade “*b*”, onde constar “-”, lê-se itens que não possuem categorias ou alternativas de respostas naquele nível.

Tabela 1 – Estimativas do parâmetro de discriminação “a” e de dificuldade “b” dos itens da escala de infraestrutura escolar da educação infantil

Item	Descrição	Parâmetro “a”	Parâmetro “b” 1	Parâmetro “b” 2	Parâmetro “b” 3	Parâmetro “b” 4	Parâmetro “b” 5
1	Forma de ocupação do prédio	0,37	-3,66	-2,78	-	-	-
2	Local de funcionamento da escola	1,21	-5,15	-	-	-	-
3	Abastecimento de água	1,30	-4,92	-2,54	-	-	-
4	Abastecimento de energia	1,52	-5,22	-4,38	-	-	-
5	Esgoto sanitário	0,89	-6,55	-1,21	-	-	-
6	Sala de diretoria	0,96	-1,23	-	-	-	-
7	Sala de professor	1,32	0,08	-	-	-	-
8	Sala de recursos multifuncionais para Atendimento Educacional Especializado (AEE)	0,97	3,46	-	-	-	-
9	Cozinha	0,81	-4,58	-	-	-	-
10	Biblioteca	0,91	1,93	-	-	-	-
11	Sala de leitura	0,70	2,19	-	-	-	-
12	Banheiro acessível, adequado ao uso dos alunos com deficiência ou mobilidade reduzida	1,18	0,84	-	-	-	-
13	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida	1,31	1,20	-	-	-	-
14	Secretaria	1,05	-0,49	-	-	-	-
15	Banheiro com chuveiro	0,87	-0,85	-	-	-	-
16	Refeitório	1,29	-0,92	-	-	-	-
17	Despensa	0,78	-1,15	-	-	-	-
18	Almoxarifado	1,19	0,16	-	-	-	-
19	Área verde	0,80	0,96	-	-	-	-
20	Número de salas utilizadas como sala de aula (dentro e fora do prédio)	1,08	1,24	2,71	4,11	5,12	-
21	Aparelho de televisão	1,17	-2,51	-0,20	0,78	1,86	-
22	Aparelho de DVD	1,03	-2,09	0,25	1,23	-	-
23	Copiadora	0,86	0,69	3,38	-	-	-
24	Impressora	0,99	-0,68	1,48	2,91	4,12	6,63
25	Impressora multifuncional	0,72	0,08	5,23	-	-	-
26	Aparelho de som	1,37	-1,38	0,11	1,06	-	-
27	Projetor multimídia (<i>datashow</i>)	1,50	0,50	2,39	3,21	4,29	-
28	Máquina fotográfica/Filmadora	1,28	-0,01	1,93	3,60	-	-
29	Computador	2,20	1,52	2,29	2,93	3,50	-
30	Computadores para uso administrativo	2,24	1,13	2,36	3,14	-	-
31	Computadores para uso dos alunos	1,65	1,74	2,77	3,90	-	-
32	Acesso à internet	1,73	-0,66	-	-	-	-
33	Parque infantil	1,50	-0,64	-	-	-	-
34	Berçário	0,75	0,01	-	-	-	-
35	Banheiro de El	1,14	-0,92	-	-	-	-

Fonte: Elaboração da autora.

Tabela 2 – Estimativas do parâmetro de discriminação “a” e de dificuldade “b” dos itens da escala de infraestrutura escolar do ensino fundamental

Item	Descrição	Parâmetro “a”	Parâmetro “b” 1	Parâmetro “b” 2	Parâmetro “b” 3	Parâmetro “b” 4	Parâmetro “b” 5
1	Forma de ocupação do prédio	0,76	-3,76	-2,18	-	-	-
2	Água consumida pelos alunos	0,41	-14,53	-4,65	-	-	-
3	Abastecimento de água	1,34	-6,16	-4,80	-	-	-
4	Abastecimento de energia	0,40	-11,42	-0,86	-	-	-
5	Esgoto sanitário	1,17	-2,33	-	-	-	-
6	Sala de diretoria	1,66	-1,95	-	-	-	-
7	Sala de professor	2,21	-0,92	-	-	-	-
8	Laboratório de informática	1,23	0,08	-	-	-	-
9	Laboratório de ciências	0,57	3,10	-	-	-	-
10	Sala de recurso multifuncionais para Atendimento Educacional Especializado (AEE)	1,31	-0,91	-	-	-	-
11	Biblioteca	0,40	2,27	-	-	-	-
12	Sala de leitura	1,14	-0,51	-	-	-	-
13	Banheiro	0,82	0,23	-	-	-	-
14	Banheiro acessível, adequado ao uso dos alunos com deficiência ou mobilidade reduzida	1,43	-1,70	-	-	-	-
15	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida	0,57	0,67	-	-	-	-
16	Secretaria	0,94	-0,41	-	-	-	-
17	Refeitório	1,13	0,72	-	-	-	-
18	Almoxarifado	0,70	0,95	-	-	-	-
19	Auditório	0,89	-0,62	0,41	1,61	2,65	-
20	Área verde	1,00	-2,08	-0,63	0,34	1,34	-
21	Número de salas utilizadas como sala de aula (dentro e fora do prédio)	0,69	-1,29	0,79	2,41	-	-
22	Aparelho de televisão	0,53	-0,45	2,00	-	-	-
23	Aparelho de DVD	0,96	-1,46	-0,37	0,39	1,08	2,02
24	Copiadora	0,78	-0,36	1,70	-	-	-
25	Impressora	0,80	-1,28	0,54	2,57	-	-
26	Impressora multifuncional	1,26	-1,41	-0,44	0,22	0,95	-
27	Aparelho de som	1,10	-0,51	1,08	2,72	-	-
28	Projetor multimídia (<i>datashow</i>)	8,80	-0,71	-0,19	0,16	0,55	-
29	Máquina fotográfica/Filmadora	2,83	-0,57	0,23	0,71	-	-
30	Computador	6,01	-0,74	-0,09	0,29	-	-
31	Computadores para uso administrativo	1,96	-1,69	-	-	-	-
32	Computadores para uso dos alunos	0,76	-3,76	-2,18	-	-	-
33	Acesso à internet	0,41	-14,53	-4,65	-	-	-

Fonte: Elaboração da autora.

Tabela 3 – Estimativas do parâmetro de discriminação “a” e de dificuldade “b” dos itens da escala de infraestrutura escolar do ensino médio

Item	Descrição	Parâmetro “a”	Parâmetro “b” 1	Parâmetro “b” 2	Parâmetro “b” 3	Parâmetro “b” 4	Parâmetro “b” 5
1	Forma de ocupação do prédio	0,76	-3,76	-2,18	-	-	-
2	Abastecimento de água	0,41	-14,53	-4,65	-	-	-
3	Abastecimento de energia	1,34	-6,16	-4,80	-	-	-
4	Esgoto sanitário	0,40	-11,42	-0,86	-	-	-
5	Sala de diretoria	1,17	-2,33	-	-	-	-
6	Sala de professor	1,66	-1,95	-	-	-	-
7	Laboratório de informática	2,21	-0,92	-	-	-	-
8	Laboratório de ciências	1,23	0,08	-	-	-	-
9	Sala de recurso multifuncionais para Atendimento Educacional Especializado (AEE)	0,57	3,10	-	-	-	-
10	Biblioteca	1,31	-0,91	-	-	-	-
11	Sala de leitura	0,40	2,27	-	-	-	-
12	Banheiro acessível, adequado ao uso dos alunos com deficiência ou mobilidade reduzida	1,14	-0,51	-	-	-	-
13	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida	0,82	0,23	-	-	-	-
14	Secretaria	1,43	-1,70	-	-	-	-
15	Refeitório	0,57	0,67	-	-	-	-
16	Almoxarifado	0,94	-0,41	-	-	-	-
17	Auditório	1,13	0,72	-	-	-	-
18	Área verde	0,70	0,95	-	-	-	-
19	Número de salas utilizadas como sala de aula (dentro e fora do prédio)	0,89	-0,62	0,41	1,61	2,65	-
20	Aparelho de televisão	1,00	-2,08	-0,63	0,34	1,34	-
21	Aparelho de DVD	0,69	-1,29	0,79	2,41	-	-
22	Copiadora	0,53	-0,45	2,00	-	-	-
23	Impressora	0,96	-1,46	-0,37	0,39	1,08	2,02
24	Impressora multifuncional	0,78	-0,36	1,70	-	-	-
25	Aparelho de som	0,80	-1,28	0,54	2,57	-	-
26	Projetor multimídia (<i>datashow</i>)	1,26	-1,41	-0,44	0,22	0,95	-
27	Máquina fotográfica/Filmadora	1,10	-0,51	1,08	2,72	-	-
28	Computador	8,80	-0,71	-0,19	0,16	0,55	-
29	Computadores para uso administrativo	2,83	-0,57	0,23	0,71	-	-
30	Computadores para uso dos alunos	6,01	-0,74	-0,09	0,29	-	-
31	Acesso à internet	1,96	-1,69	-	-	-	-
32	Banheiro	0,46	-8,88	-2,00	-	-	-

Fonte: Elaboração da autora.

Definição e descrição das escalas de infraestrutura escolar para todos os segmentos de ensino

Adotando como base o estudo de Beaton e Allen (1992)⁴, foram definidos os itens-âncora de cada um dos níveis da escala de desempenho, ou seja, aqueles que são característicos de um ponto ou um nível da escala em virtude da grande maioria das escolas situadas naquele nível indicar que possuem o recurso. O posicionamento dado pela proporção empírica, tende a se assemelhar ao posicionamento previsto pela TRI, alcançando-se uma distribuição dos itens e das escolas semelhante ao de uma curva normal na escala de infraestrutura.

A fim de se estabelecer os pontos de corte da escala, o valor da média, do desvio padrão, dos parâmetros dos itens e dos scores das escolas, que foram calculados originalmente na métrica (0;1), foram transformados para uma escala cuja média é 50 e o desvio padrão, 10. É esperado que a maior concentração de escolas esteja no intervalo de 40 a 60 pontos da escala, ou seja, no intervalo entre um desvio padrão abaixo e um desvio padrão acima da média. De modo a se aproximar ao máximo dessa distribuição normal, os intervalos foram estabelecidos com variações de 1 em 1 desvio padrão, com menor variabilidade nos níveis intermediários e maior variação nos extremos das escalas. Assim, definiu-se que as escalas de todos os segmentos seriam descritas em quatro níveis, assim denominados: elementar (*score* na escala de infraestrutura maior ou igual a 0 e menor que 30); básica (*score* na escala de infraestrutura maior ou igual a 30 e menor que 50); adequada (*score* na escala de infraestrutura maior ou igual a 50 e menor que 70); e avançada (*score* na escala de infraestrutura maior ou igual a 70).

⁴Essa metodologia estabelece três critérios para que o item possa ser classificado como âncora: (1) se a maioria das escolas situadas naquele nível da escala (mais de 65%) responde positivamente ao item; (2) menos da metade das escolas situadas no nível imediatamente inferior também o responde positivamente; (3) a diferença entre os percentuais de resposta no ponto âncora e no ponto imediatamente anterior é igual ou superior a 30%.

RESULTADOS E DISCUSSÃO

Analisando os resultados de calibração dos itens, observa-se que os itens comuns a todos os segmentos de ensino apresentam parâmetros “*b*” de dificuldade muito distintos, a depender do segmento analisado, reforçando a necessidade de especificação das escalas de infraestrutura escolar. Destaca-se o resultado dos itens que representam a dimensão da infraestrutura escolar básica (abastecimento de água, energia elétrica e esgoto), que, contrariamente às expectativas iniciais da pesquisa – por serem concebidos como itens indispensáveis –, apresentaram parâmetros de dificuldade muito altos e discrepantes ao comparar as três escalas.

Ainda em relação a esses itens da dimensão de infraestrutura básica, é desejável que, no processo de construção de medidas resumidas e padronizadas para o monitoramento de políticas públicas, o seu provimento seja garantido via rede pública. No entanto, observa-se que diversas escolas brasileiras podem não se encaixar nesse cenário em que a prestação de serviços públicos é a alternativa mais viável. Isso ocorre porque, considerando os contextos de escolas localizadas em áreas de assentamento, comunidades remanescentes de quilombolas e indígenas e unidades de uso sustentável, o abastecimento desses serviços via fontes alternativas (energia via geradores, eólica, solar, abastecimento de água via nascentes e poços e tratamento de esgoto via fossa séptica) parece ser ecologicamente mais coerente.

Diferentemente desses recursos básicos, os itens que compõem as demais dimensões operacionais da infraestrutura escolar (predial, equipamentos de apoio pedagógico, pedagógica e adequada às necessidades especiais) apresentaram parâmetros de dificuldade medianos e bem próximos, embora distintos, ao se analisar as escalas de cada um dos segmentos de ensino. Esses resultados corresponderam às expectativas iniciais, tendo em vista que, em geral, esses itens são considerados medianos ou fáceis, dada a sua relevância para cada uma das etapas de ensino.

Assim, na construção da interpretação das escalas de infraestrutura escolar específicas para cada um dos segmentos de ensino da educação básica, observa-se um maior nível de detalhamento quando são abordadas as variáveis que avaliam a qualidade da infraestrutura básica e as variáveis que avaliam quantitativamente os equipamentos de apoio pedagógico. Além disso, destaca-se o fato de ter sido possível, na composição das escalas, adequar a relevância dos itens que compõem a dimensão da infraestrutura pedagógica e o posicionamento que eles assumem, conforme a dificuldade calculada pelo parâmetro “*b*” da TRI.

Como consequência, nas escalas da EI, do EF e do EM, os itens característicos de cada uma dessas etapas que representam a dimensão da infraestrutura pedagógica foram posicionados no nível de infraestrutura adequada, conforme sua necessidade para o desenvolvimento educacional esperado nas diferentes etapas de ensino. No nível de infraestrutura avançada, foram posicionados apenas os itens que caracterizam a escola em relação ao seu porte, à quantidade de materiais e ao bem-estar oferecido aos estudantes que a frequentam, além dos itens relacionados à adequação ao atendimento de estudantes com deficiência ou mobilidade reduzida. Desse modo, as escalas tornam-se mais justas, pois consideram as especificidades e os diversos contextos dos diferentes segmentos de ensino e que o posicionamento das escolas no nível avançado, representa um avanço no porte da escola e não necessariamente na qualidade da sua infraestrutura. Nos Quadros 3 a 5, são apresentadas as propostas de descrições das escalas de infraestrutura escolar para a EI, o EF e o EM, assim como o percentual de escolas do Censo de 2017 em cada intervalo da escala.

No Quadro 3, a média de desempenho das escolas brasileiras que autodeclararam no Censo Escolar de 2017 ofertar exclusivamente a EI foi de 44,37 pontos na escala, apresentando uma infraestrutura escolar básica. Evidencia-se que essa média representa 36,40% das escolas desse segmento, que, provavelmente, não possuem nos níveis desejáveis os itens essenciais à dimensão da infraestrutura básica, pois ocupam prédios alugados, são atendidas via rede pública apenas por água e, em relação à infraestrutura predial, apresentam somente cozinha e até 7 salas utilizadas como salas de aula. Elas não possuem infraestrutura pedagógica para a EI e, em relação à infraestrutura de apoio pedagógico, contam somente com 5 computadores, sendo apenas 1 deles para uso dos estudantes, além de 1 aparelho de TV, DVD e de som.

Também é possível verificar a existência de um percentual muito similar (35,73%) de escolas que apresentam uma infraestrutura escolar adequada. Segundo a escala da EI, provavelmente, essas escolas possuem os requisitos desejáveis ao se considerar uma infraestrutura básica, como abastecimento de água, energia elétrica e esgoto sanitário provenientes de rede pública. Em relação à infraestrutura predial, elas dispõem, em suas dependências, de até 7 salas utilizadas como salas de aula, sala de diretoria e de professor, cozinha, secretaria, refeitório, despensa e almoxarifado. No que se refere à infraestrutura pedagógica, elas possuem todos os itens desejáveis a uma escola de EI: banheiro com chuveiro e adequados à EI, parque infantil e berçários. Na dimensão da infraestrutura de apoio pedagógico, essas escolas contam com um quantitativo maior de equipamentos se comparado com as escolas posi-

cionadas no nível básico. Elas têm 2 aparelhos de televisão e de DVD; de 2 a 3 aparelhos de som; de 1 a 2 impressoras, sendo pelo menos 1 multifuncional; 1 equipamento de multimídia e 1 máquina fotográfica/filmadora; até 5 computadores, sendo apenas 1 deles para uso dos estudantes; e acesso à internet.

Ainda é preocupante o fato de que 22,88% das escolas de EI apresentam uma infraestrutura escolar elementar, visto que não apresentam nenhum dos requisitos desejáveis para a dimensão da infraestrutura escolar básica, da infraestrutura predial e de equipamentos pedagógicos. No lado oposto da escala, da infraestrutura escolar avançada, encontram-se apenas 4,99% das escolas de EI que se diferenciam das escolas do nível adequado apenas pela maior quantidade de equipamentos pedagógicos e pela dimensão predial que possuem. Destaca-se que essas escolas também já são capazes de atender estudantes com necessidades especiais.

Quadro 3 – Descrição e percentuais de escolas por intervalo de proficiência do indicador de infraestrutura escolar para a educação infantil

Níveis da escala	Intervalos da escala	Percentuais de escolas	Descrição dos intervalos da escala de infraestrutura da EI
1 – Elementar	Maior que 0 e menor que 10	1,24%	Neste intervalo, considerando a dimensão da infraestrutura básica , as escolas são caracterizadas por, provavelmente, ocupar espaços cedidos, cujos locais de funcionamento são adversos (salas de empresa, templo de igrejas, galpão, prédios compartilhados e unidades prisionais ou de atendimento socioeducativo, salas de aulas em outras escolas e casa de professor); o abastecimento de água é inexistente ou realizado por fontes alternativas (cacimba, cisterna, poço, rio, igarapé, riacho ou córrego); o abastecimento de energia elétrica é inexistente ou advém de fontes alternativas; e o esgotamento sanitário é inexistente ou realizado através de fossa séptica. Em relação à infraestrutura predial , essas escolas, provavelmente, possuem até 7 salas utilizadas como salas de aula. Na dimensão da infraestrutura para equipamentos de apoio pedagógico , provavelmente, possuem até 5 computadores, sendo apenas 1 deles para uso dos estudantes.
	Maior ou igual a 10 e menor que 30	21,64%	Além das condições citadas acima, as escolas neste intervalo, considerando a dimensão da infraestrutura básica , provavelmente, já funcionam em prédios escolares, possuem abastecimento de água por meio de poços artesianos, a energia elétrica é proveniente de geradores e o esgoto sanitário, via fossa séptica. Em relação à infraestrutura predial , essas escolas, provavelmente, contam com cozinha em suas dependências.
2 - Básica	Maior ou igual a 30 e menor que 40	16,17%	Além dos itens citados no nível anterior, as escolas neste nível, considerando a dimensão da infraestrutura básica , provavelmente, ocupam espaços alugados e possuem abastecimento de água proveniente da rede pública. Em relação à infraestrutura para equipamentos de apoio pedagógico , provavelmente, possuem aparelho de televisão, mas ainda não contam com itens característicos da dimensão pedagógica e adequada às necessidades especiais .
	Maior ou igual a 40 e menor que 50	20,23%	Além das condições citadas acima, as escolas neste intervalo, considerando a dimensão da infraestrutura básica , provavelmente, já ocupam espaços próprios e, em relação à infraestrutura para equipamentos de apoio pedagógico , provavelmente, possuem aparelho de DVD e de som.

Níveis da escala	Intervalos da escala	Percentuais de escolas	Descrição dos intervalos da escala de infraestrutura da EI
3 – Adequada	Maior ou igual a 50 e menor que 60	22,18%	Além dos itens citados nos níveis anteriores, as escolas neste nível, considerando a dimensão da infraestrutura básica , provavelmente, possuem esgoto sanitário proveniente da rede pública. Considerando a infraestrutura predial e pedagógica , possuem sala de diretoria, secretaria, banheiro com chuveiro e adequado à EI, refeitório, despensa e parque infantil. Em relação à infraestrutura para equipamentos de apoio pedagógico , as escolas neste nível têm uma impressora.
	Maior ou igual a 60 e menor que 70	13,55%	Além das condições citadas acima, as escolas neste intervalo, considerando a infraestrutura predial e pedagógica , provavelmente, têm sala de professor, almoxarifado e berçário. Considerando a infraestrutura para equipamentos de apoio pedagógico , possuem uma maior quantidade de aparelhos de televisão e de DVD (2) e de som (2 a 3), além de contarem com impressoras multifuncionais, equipamento multimídia, máquina fotográfica/filmadora e acesso à internet.
4 – Avançada	Maior ou igual a 70 e menor que 80	4,03%	Além dos itens citados nos níveis anteriores, as escolas neste nível, provavelmente, atendem à dimensão de infraestrutura adequada às necessidades especiais , com banheiros e dependências acessíveis ao uso dos alunos com deficiência ou mobilidade reduzida. Em relação à infraestrutura predial , possuem área verde e um número maior de salas utilizadas como salas de aula (8 a 10). Na dimensão de infraestrutura para equipamentos de apoio pedagógico , as escolas têm uma maior quantidade de impressoras (2), aparelhos de televisão (3 a 4), DVD (3), som (mais de 4), copiadora e computadores (6 a 15), sendo pelo menos 3 computadores de uso administrativo e de 2 a 11 destinados para uso dos alunos.
	Maior ou igual a 80 e menor que 90	0,85%	Além das condições citadas acima, as escolas neste intervalo, provavelmente, possuem em sua infraestrutura pedagógica biblioteca e sala de leitura. Na dimensão de infraestrutura para equipamentos de apoio pedagógico , têm uma maior quantidade de aparelhos de televisão (5), equipamentos de multimídia (2), máquinas fotográficas/filmadoras (2 a 3) e computadores (16 a 23), sendo pelo menos 6 computadores para uso administrativo.
	Maior ou igual a 90	0,10%	Além das condições citadas acima, as escolas neste intervalo, provavelmente, atendem à dimensão infraestrutura adequada às necessidades especiais , com sala de atendimento especial. Em relação à infraestrutura predial , possuem um número maior de salas utilizadas como salas de aula (mais de 11). Na dimensão de infraestrutura para equipamentos de apoio pedagógico , as escolas têm uma maior quantidade de copiadoras (2), impressoras (3), equipamentos de multimídia (mais de 3), máquinas fotográficas/filmadoras (mais de 4) e computadores (mais de 24), sendo pelo menos 10 computadores para uso administrativo e 12 para uso dos alunos.

Fonte: Elaboração da autora.

No Quadro 4, é apresentada a escala de infraestrutura escolar do EF e é que apresenta a menor quantidade de variáveis operacionais ou características da dimensão pedagógica dessa etapa de ensino, pois todos os itens considerados na sua composição são comuns e essenciais para os outros segmentos, como biblioteca e laboratórios. A média na escala do indicador foi de 49,08 pontos, indicando que as escolas brasileiras que oferecem essa etapa de ensino possuem uma infraestrutura escolar básica, mas quase no nível adequado e desejado da escala. Nessa situação, estão 49,13% das escolas brasileiras.

Já as escolas posicionadas no nível adequado representam um percentual de 49,56% e, provavelmente, contam com todos os recursos relacionados à dimensão da infraestrutura básica via rede pública, possuem em suas dependências uma estrutura predial de banheiros dentro e fora da escola, têm mais de 8 salas utilizadas como salas de aula, além de diretoria e de professor, secretaria, refeitório e almoxarifado. Elas ainda dispõem de itens específicos da dimensão de infraestrutura adequada ao atendimento especial, com salas e banheiro acessível aos estudantes com deficiência. Considerando a dimensão da infraestrutura pedagógica, essas escolas possuem laboratório de informática, biblioteca e sala de leitura e contam com uma quantidade maior de equipamentos de apoio pedagógico: pelo menos 2 aparelhos de TV, DVD, som, foto, multimídia, copiadora, impressora (sendo pelo menos 1 multifuncional), de 6 a 23 computadores, sendo pelo menos 6 de uso administrativo e 11 destinados aos estudantes, e acesso à internet.

Nas extremidades da escala de infraestrutura específica para o EF, posiciona-se um percentual pouco expressivo de escolas (4,39%), resultado que pode ser considerado, em certa medida, positivo, pois indica que, pelo menos no nível elementar, há poucas escolas desse segmento de ensino. Vale lembrar que todos os recursos escolares essenciais dessa etapa estão posicionados no nível adequado da escala de infraestrutura escolar.

Por fim, em relação ao EM, as escolas que autodeclararam ofertar exclusivamente essa etapa de ensino apresentaram a maior média de 52,28 pontos na escala de infraestrutura escolar. Essas escolas estão no nível de infraestrutura adequada e compreendem 61,22% das escolas brasileiras. Elas, provavelmente, contam com todos os recursos relacionados à dimensão da infraestrutura básica via rede pública. Possuem, em suas dependências, uma estrutura predial com banheiros dentro e fora da escola e acessíveis aos estudantes com deficiência, além de sala de diretoria e de professor, secretaria, refeitório, almoxarifado, área verde, auditório e mais de 11 salas utilizadas como salas de aula. Em relação à infraestrutura pedagógica, essas escolas, provavelmente, possuem laboratórios de informática e de ciências, assim como biblioteca, e, na dimensão da infraestrutura adequada às necessidades especiais, elas contam com sala de atendimento especial. Essas escolas, provavelmente, ainda possuem uma quantidade maior de equipamentos pedagógicos: pelo menos 2 aparelhos de TV, DVD, som, foto, multimídia, copiadora, impressora (sendo pelo menos 1 multifuncional) e de 6 a 23 computadores, sendo pelo menos 6 de uso administrativo e 11 destinados aos estudantes, além de acesso à internet. Para o ensino médio, os resultados ainda demonstram que 36,78% das escolas ainda não possuem uma infraestrutura mínima para ofertar esse

seguimento de ensino, pois estão nos níveis elementar e básico da escala. Indicando as condições desiguais em que estas escolas se encontram, especialmente no que se refere a uma infraestrutura pedagógica mínima desejável.

Destaca-se ainda, que os itens biblioteca e sala de leitura, que compõem a dimensão de infraestrutura pedagógica para todos os segmentos de ensino, se comportaram de modo inesperado ao se posicionaram no nível avançado das escalas da EI e do EM (apenas sala de leitura). Esse posicionamento pode estar relacionado a uma não compreensão das escolas, no processo de auto declaração ao questionário do Censo Escolar, de que a sala de leitura também pode não ser compreendida como um espaço similar ou substitutivo da biblioteca ou a uma não usabilidade, de fato, desses espaços. Outro item relacionado à dimensão da infraestrutura pedagógica, posicionado no nível avançado da escala é o laboratório de ciências para o EF, indicando que há poucas escolas que contam que esse ambiente em suas dependências.

Ademais, destaca-se que, embora as três escalas não sejam metodologicamente comparáveis, os segmentos de ensino só passam a contar com todos os itens da infraestrutura básica a partir do nível adequado. Os itens relacionados à dimensão da infraestrutura adequada às necessidades especiais também são características de escolas que apresentam nível adequado e avançado, indicando, portanto, que escolas que fazem este tipo de atendimento pedagógico possuem melhores condições de infraestrutura.

A partir da interpretação dessas escalas, espera-se que todas as escolas da educação básica sejam posicionadas no nível adequado, onde todas as variáveis de infraestrutura pedagógica estão posicionadas. Esse nível passa a ser considerado como o patamar mínimo para o funcionamento escolar, independentemente do segmento. Desconstrói-se, assim, a ideia de que uma infraestrutura escolar ideal é representada pelo nível avançado, uma vez que esse nível é composto apenas por variáveis que indicam um avanço no porte das escolas.

Quadro 4 – descrição e percentuais de escolas por intervalo de proficiência do indicador de infraestrutura escolar para o ensino fundamental

Níveis da escala	Intervalos da escala	Percentuais de escolas	Descrição dos intervalos da escala de infraestrutura do EF
1 - Elementar	Maior que 20 e menor que 30	3,07%	Neste nível, considerando a dimensão da infraestrutura básica , as escolas, provavelmente, são caracterizadas por ocupar espaços próprios, onde os alunos consomem água filtrada, o abastecimento de água é via poços artesianos, o fornecimento de energia elétrica advém de geradores e o esgotamento sanitário é através de fossa séptica. Em relação à infraestrutura predial , elas, provavelmente, possuem banheiros dentro e fora da escola e até 7 salas utilizadas como salas de aula. Considerando a infraestrutura para equipamentos de apoio pedagógico , elas, provavelmente, contam com até 5 computadores, sendo apenas 1 deles para uso dos estudantes.
2 - Básica	Maior ou igual a 30 e menor que 40	19,84%	Além dos itens citados no nível anterior, as escolas neste nível, considerando a dimensão da infraestrutura básica , provavelmente, possuem abastecimento de energia elétrica via rede pública.
	Maior ou igual a 40 e menor que 50	26,21%	Além das condições citadas acima, considerando a dimensão da infraestrutura básica , as escolas neste intervalo, provavelmente, possuem abastecimento de água também via rede pública. Em relação à infraestrutura predial , provavelmente, possuem sala de diretoria e de professor e secretaria. Considerando a infraestrutura para equipamentos de apoio pedagógico , essas escolas, provavelmente, têm aparelho de TV, DVD, som, multimídia e impressora, além de acesso à internet.
3 - Adequada	Maior ou igual a 50 e menor que 60	35,72%	Além dos itens citados nos níveis anteriores, considerando a dimensão da infraestrutura básica , as escolas neste nível, provavelmente, possuem esgotamento sanitário via rede pública e, em relação à infraestrutura predial , possuem uma quantidade maior de salas utilizadas como salas de aula (8 a 10), além de refeitório e almoxarifado. No que se refere à infraestrutura pedagógica , provavelmente, contam com laboratórios de informática e biblioteca. Considerando as dimensões pedagógica e adequada às necessidades especiais , as escolas neste nível, provavelmente, contam com salas de atendimento especial e banheiro acessível aos estudantes com deficiência. Em relação à infraestrutura para equipamentos de apoio pedagógico , provavelmente, possuem copiadora e uma maior quantidade de aparelhos de TV (2) e DVD (2), equipamentos de som (2 a 3), impressoras (2) – sendo pelo menos 1 multifuncional –, equipamentos de foto e computadores (6 a 15), sendo pelo menos 3 para uso administrativo e de 2 a 11 para os alunos.
	Maior ou igual a 60 e menor que 70	13,84%	Além das condições citadas acima, as escolas neste intervalo, considerando a infraestrutura predial , provavelmente, possuem uma quantidade maior de salas utilizadas como salas de aula (11 a 14). Em relação à infraestrutura pedagógica , provavelmente, possuem sala de leitura e, em relação à dimensão adequada às necessidades especiais , contam com dependências adequadas para alunos com deficiência. No que se refere à infraestrutura para equipamentos de apoio pedagógico , essas escolas, provavelmente, possuem uma maior quantidade de copiadoras (2), equipamentos de foto (2 a 3), impressoras (mais de 3) e aparelhos de TV (3 a 4), multimídia (4) e som (mais de 4), além de computadores (16 a 23), sendo pelo menos 6 destinados ao uso administrativo e 12 ao uso dos estudantes.

Níveis da escala	Intervalos da escala	Percentuais de escolas	Descrição dos intervalos da escala de infraestrutura do EF
4 - Avançada	Maior ou igual a 70 e menor que 80	1,28%	Além dos itens citados nos níveis anteriores, considerando a infraestrutura predial , as escolas neste nível, provavelmente, contam com auditório e área verde em suas dependências. Em relação à infraestrutura pedagógica , possuem laboratório de ciências. No que se refere à infraestrutura para equipamentos de apoio pedagógico , essas escolas possuem uma maior quantidade de impressoras multifuncionais (2), equipamentos de multimídia (mais 5) e computadores (mais de 24).
	Maior ou igual a 80	0,03%	Além das condições citadas acima, considerando a infraestrutura predial , as escolas neste intervalo, provavelmente, possuem uma quantidade maior de salas utilizadas como salas de aula (mais de 15) e, em relação à infraestrutura de apoio pedagógico , elas possuem mais de 3 impressoras multifuncionais.

Fonte: Elaboração da autora.

Quadro 5 – descrição e percentuais de escolas por intervalo de proficiência do indicador de infraestrutura escolar para o ensino médio

Nível	Intervalo	Percentual de escolas	Descrição dos intervalos da escala de infraestrutura do EM
1 - Elementar	Maior que 10 e menor que 20	0,24%	Neste nível, considerando a dimensão da infraestrutura básica , as escolas são caracterizadas, provavelmente, por ocupar espaços cedidos, onde o abastecimento de água é realizado por fontes alternativas (cacimba, cisterna, poço, rio, igarapé, riacho ou córrego) ou poço artesiano, o fornecimento de energia elétrica advém de geradores ou fontes alternativas e o esgotamento sanitário é por meio de fossa séptica. Em relação à infraestrutura predial , provavelmente, possuem banheiros dentro do prédio e até 7 salas utilizadas como salas de aula. No que se refere à infraestrutura de apoio pedagógico , essas escolas, provavelmente, possuem até 5 computadores, sendo apenas 1 deles para uso dos estudantes.
	Maior ou igual a 20 e menor que 30	1,81%	Além das condições citadas acima, considerando a dimensão da infraestrutura básica , as escolas neste intervalo, provavelmente, ocupam espaços alugados e contam com abastecimento de água e energia elétrica por meio da rede pública. Em relação à infraestrutura predial , provavelmente, possuem banheiros dentro e fora do prédio.
2 - Básica	Maior ou igual a 30 e menor que 40	5,18%	Além dos itens citados no nível anterior, considerando a dimensão da infraestrutura básica , as escolas neste nível, provavelmente, ocupam espaços próprios e, considerando a dimensão da infraestrutura predial , elas possuem sala de diretoria e de professores e secretaria. Em relação à infraestrutura de apoio pedagógico , provavelmente, contam com aparelho de TV e multimídia, impressora e acesso à internet.
	Maior ou igual a 40 e menor que 50	29,44%	Além das condições citadas acima, considerando a infraestrutura predial , as escolas neste intervalo, provavelmente, possuem uma quantidade maior de salas utilizadas como salas de aulas (8 a 10). Essas escolas, provavelmente, também passam a ter itens específicos da dimensão de infraestrutura pedagógica , como laboratório de informática e biblioteca, bem como uma infraestrutura adequada às necessidades especiais , com banheiros adequados aos alunos com deficiência. Em relação à infraestrutura de apoio pedagógico , provavelmente, contam com equipamento de DVD, som, máquina fotográfica, aparelho de TV e multimídia e uma quantidade maior de computadores (16 a 23), sendo pelo menos 3 de uso administrativo e de 2 a 19 destinados aos alunos.

Nível	Intervalo	Percentual de escolas	Descrição dos intervalos da escala de infraestrutura do EM
3 - Adequada	Maior ou igual a 50 e menor que 60	46,22%	Além dos itens citados nos níveis anteriores, considerando a dimensão da infraestrutura básica , as escolas neste nível, provavelmente, possuem esgotamento sanitário via rede pública e, em relação à infraestrutura predial , elas possuem uma quantidade maior de salas utilizadas como salas de aula (11 a 14), além de almoçarifado. No que se refere à infraestrutura pedagógica , essas escolas, provavelmente, possuem laboratório de ciências e, na dimensão da infraestrutura adequada às necessidades especiais , contam com dependências adequadas a alunos com deficiência. Considerando a infraestrutura de apoio pedagógico , elas, provavelmente, possuem copiadora e uma quantidade maior de aparelhos de TV (2) e impressoras (2), sendo pelo menos 1 multifuncional.
	Maior ou igual a 60 e menor que 70	15,01%	Além das condições citadas acima, considerando a infraestrutura predial , as escolas neste intervalo, provavelmente, possuem auditório e área verde. Em relação à infraestrutura de apoio pedagógico , provavelmente, contam com uma quantidade maior de aparelhos de TV (3 a 4) e de DVD (2), além de impressoras (3).
4 - Avançada	Maior ou igual a 70 e menor que 80	2,03%	Além dos itens citados nos níveis anteriores, considerando a infraestrutura predial , as escolas neste nível, provavelmente, possuem uma quantidade maior de salas utilizadas como sala de aula (15 a 18). Em relação à infraestrutura de apoio pedagógico , provavelmente, contam com uma quantidade maior de copiadoras (2), de aparelhos de DVD (mais de 3) e impressoras (mais de 3), sendo pelo menos 2 multifuncionais.
	Maior ou igual a 80	0,07%	Além das condições citadas acima, considerando a dimensão da infraestrutura adequada às necessidades especiais , as escolas neste intervalo, provavelmente, possuem sala de atendimento especial e, considerando a infraestrutura pedagógica , elas contam com sala de leitura. Na dimensão da infraestrutura predial , provavelmente, têm uma quantidade maior de salas utilizadas como salas de aula (mais de 19).

Fonte: Elaboração da autora.

Análises descritivas de cada uma das escalas dos segmentos de ensino, por região, dependência administrativa e localização das escolas

Após a descrição das escalas de infraestrutura, deu-se início às análises descritivas, estabelecendo-se, primeiramente, as médias e os percentuais de escolas em todas as escalas, por região, unidade da federação, dependência administrativa e localização. A Tabela 4 mostra, em termos absolutos e relativos, a distribuição das escolas nos intervalos e nos níveis das escalas.

Tabela 4 – Quantitativo e percentuais de escolas nos intervalos de proficiência da escala de infraestrutura escolar, por segmento de ensino

Níveis de infraestrutura	Intervalos da escala	EI		EF		EM	
		N.	%	N.	%	N.	%
Elementar	0 < 10	1031	1,24	0	0,00	0	0,00
	10 < 20	6132	7,38	0	0,00	59	0,24
	20 < 30	11838	14,25	3098	3,07	450	1,81
Básica	30 < 40	13433	16,17	20022	19,84	1287	5,18
	40 < 50	16802	20,23	26448	26,21	7309	29,44
Adequada	50 < 60	18417	22,18	36037	35,72	11474	46,22
	60 < 70	11256	13,55	13966	13,84	3725	15,01
Avançada	70 < 80	3351	4,03	1295	1,28	503	2,03
	80 < 90	709	0,85	33	0,03	18	0,07
	90 < 100	83	0,10	0	0,00	0	0,00

Fonte: Elaboração da autora.

Nos intervalos que incluem as escolas com infraestrutura elementar em cada um dos segmentos de ensino, destaca-se a diferença entre os percentuais de escolas da EI (22,88%) e do EF e do EM (3,07% e 2,05%, respectivamente). Esse comportamento evidencia que o investimento na infraestrutura escolar da primeira infância deve ser prioridade na agenda das políticas públicas, dado o quantitativo de escolas que estão posicionadas três desvios abaixo do ponto considerado como o patamar mínimo da escala de infraestrutura escolar. O EF é o segmento com o maior percentual de escolas no nível básico (46,06%) e o EM com o maior percentual de escolas no nível adequado (61,22%). Todos os segmentos de ensino apresentaram um número pouco expressivo de escolas no nível avançado (4,99%; 1,32% e 2,10%).

Na Tabela 5, são apresentadas as porcentagens de escolas de todos os segmentos de ensino, em cada um dos níveis da escala, por região geográfica do Brasil. Em qualquer uma das regiões, conforme esperado pela descrição das escalas, a porcentagem de escolas em todos os segmentos de ensino no nível avançado é sempre menor que 2%. Observa-se que a região Sudeste apresenta o maior percentual de escolas no nível adequado para todos os segmentos.

No nível básico, a região Nordeste conta com o maior percentual de escolas da EI (21,99%) e do EF (29,11%), enquanto a região Sudeste possui o maior percentual de escolas de EM no mesmo nível (14,18%). Esse fato corrobora as vantagens em se analisar a infraestrutura escolar por escalas específicas, pois é possível observar que regiões como a Sudeste também apresentam um percentual expressivo de escolas que não possuem infraestrutura adequada para determinado segmento de ensino. Assim, são superados os prejulgamentos

de que apenas as regiões Norte e Nordeste do país apresentam situações mais críticas de infraestrutura escolar, e endossa-se a tese de que o investimento em recursos escolares deve priorizar as regiões em que os segmentos de ensino apresentam níveis mais baixos de infraestrutura escolar. Por fim, no nível elementar das escalas, a região Norte apresenta o maior percentual de escolas do EF (2,34%) e do EM (0,81%), e a região Nordeste, da EI (14,27%).

Tabela 5 – Percentuais de escolas nos níveis da escala de infraestrutura escolar, por região e segmento de ensino

Regiões	Níveis de infraestrutura escolar											
	Elementar			Básica			Adequada			Avançada		
	EI	EF	EM	EI	EF	EM	EI	EF	EM	EI	EF	EM
Norte	7,55%	2,34%	0,81%	4,20%	8,27%	4,02%	1,97%	3,28%	3,52%	0,21%	0,06%	0,11%
Nordeste	14,27%	0,70%	0,36%	21,99%	29,11%	9,71%	9,99%	12,47%	9,03%	0,83%	0,20%	0,36%
Sudeste	0,69%	0,00%	0,52%	6,86%	5,50%	14,18%	13,57%	19,98%	29,76%	1,98%	0,52%	0,73%
Sul	0,16%	0,01%	0,26%	2,20%	1,99%	3,48%	6,99%	9,37%	13,60%	1,48%	0,42%	0,71%
Centro-Oeste	0,20%	0,03%	0,10%	1,15%	1,19%	3,25%	3,21%	4,47%	5,31%	0,49%	0,11%	0,19%
Brasil	22,88%	3,07%	2,05%	36,40%	0,00%	34,63%	35,73%	49,56%	61,22%	4,99%	1,32%	2,10%

Fonte: Elaboração da autora.

A Tabela 6 apresenta em detalhes os resultados das escalas em relação à dependência administrativa, ou seja, se as escolas são federais, estaduais, municipais ou privadas, para cada um dos segmentos de ensino. Observa-se que as escolas privadas, embora que em percentuais pouco expressivos, são as que apresentam uma maior quantidade de escolas da EI (3,15%) e do EM (1,52%) no nível avançado das escalas. Nos níveis de infraestrutura escolar adequada e básica, as escolas municipais se destacam tanto para a EI (28,41% e 19,61%), quanto para o EF (37,99% e 19,08%) e as escolas estaduais, para o EM (24,31% e 39,94%).

Esses resultados refletem o atual regime de parceria entre União, estados, municípios e Distrito Federal na oferta de determinados segmentos de ensino, em que EI e EF ficam sob a responsabilidade majoritariamente dos municípios, enquanto o EM fica a cargo dos estados. Logo, considerar que escolas municipais apresentam melhor infraestrutura escolar para a EI e para o EF e que as escolas estaduais oferecem melhores infraestruturas apenas para o EM pode induzir e acirrar as desigualdades educacionais a depender do tipo de rede de escolas que os estudantes brasileiros frequentam, principalmente se considerarmos as trajetórias de estudantes que, por exemplo, cursam a educação básica toda na rede estadual de ensino.

Tabela 6 – Percentuais de escolas nos níveis da escala de infraestrutura escolar, por dependência administrativa e segmento de ensino

Dependência administrativa	Níveis de infraestrutura escolar											
	Elementar			Básica			Adequada			Avançada		
	EI	EF	EM	EI	EF	EM	EI	EF	EM	EI	EF	EM
Federal	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%	0,00%	0,02%	0,38%	0,01%	0,03%	0,13%
Estadual	0,19%	0,14%	1,97%	0,30%	2,38%	24,31%	0,36%	14,57%	39,94%	0,11%	0,32%	0,43%
Municipal	22,54%	2,93%	0,06%	28,41%	37,99%	0,91%	19,61%	19,08%	0,64%	1,72%	0,08%	0,02%
Privada	0,14%	0,00%	0,02%	7,69%	5,68%	9,40%	15,75%	15,89%	20,27%	3,15%	0,89%	1,52%
Brasil	22,88%	3,07%	2,05%	36,40%	46,06%	34,63%	35,73%	49,56%	61,22%	4,99%	1,32%	2,10%

Fonte: Elaboração da autora.

A Tabela 7 apresenta detalhadamente a distribuição das escolas nas escalas da EI, do EF e do EM, em relação à sua localização, ou seja, se são urbanas ou rurais. Observa-se, ainda, o inexpressivo percentual de escolas rurais posicionadas nos níveis adequado e avançado das escalas para todos os segmentos de ensino analisados. Nota-se uma concentração dessas escolas nos níveis mais baixos das escalas em todos os segmentos de ensino, destacando-se a EI, com 22,46% no nível elementar, e o EF, com 35,10% no nível básico.

Tabela 7 – Percentuais de escolas nos níveis da escala de infraestrutura escolar, por localização e segmento de ensino

Localização	Níveis de infraestrutura escolar											
	Elementar			Básica			Adequada			Avançada		
	EI	EF	EM	EI	EF	EM	EI	EF	EM	EI	EF	EM
Urbana	0,41%	0,01%	0,62%	14,42%	10,96%	28,40%	30,62%	43,36%	59,26%	4,92%	1,31%	2,07%
Rural	22,46%	3,06%	1,43%	21,98%	35,10%	6,23%	5,11%	6,20%	1,96%	0,07%	0,01%	0,02%
Brasil	22,88%	3,07%	2,05%	36,40%	46,06%	34,63%	35,73%	49,56%	61,22%	4,99%	1,32%	2,10%

Fonte: Elaboração da autora.

CONCLUSÕES

A presente pesquisa apresenta-se como uma inovação para a área de construção de medidas no cenário nacional, pois, além de fazer uso de uma metodologia moderna de análise a TRI para MRG, estabeleceu três escalas de infraestrutura escolar com quatro níveis que conseguem precisar, além da existência, a qualidade ou a quantidade de determinados itens nos contextos escolares para cada um dos segmentos de ensino da educação básica. Essas informações podem subsidiar um direcionamento equitativo dos recursos públicos, priorizando, por exemplo, aquelas escolas que não contam com a infraestrutura básica e pedagógica.

Além disso, pode influenciar no desenvolvimento de estratégias que viabilizem a adequação e a revitalização da infraestrutura escolar que seja capaz de ofertar condições de implementação do currículo mínimo, garantido pela BNCC, de modo a apontar, ainda, que a falta de uma infraestrutura adequada das escolas brasileiras pode ser determinante para o aumento das desigualdades educacionais no país impeditiva para implementação de um currículo nacional. Assim, faz-se necessário considerar como meta que todas as escolas da educação básica sejam posicionadas por esses indicadores no nível adequado das escalas que, por sua vez, sumariza os elementos fundamentais para o processo de ensino-aprendizagem de cada uma das etapas de ensino, estabelecendo-se este como o patamar mínimo de infraestrutura para a EI, o EF e o EM.

Contudo, este estudo apresenta como limitação a falta de um tratamento específico para variáveis relacionadas ao porte e à localização das escolas, na composição desses indicadores. Pesquisas futuras devem levar em consideração essas variáveis, verificando se de fato a ponderação delas na construção das medidas pode estar influenciando nos parâmetros de discriminação e de dificuldade das alternativas dos itens, em especial para a definição dos níveis avançados das escalas. Continua também sendo uma das limitações do indicador a quantidade de itens utilizada para a construção da escala, pois, quanto maior for, maior será a precisão das medidas e a confiabilidade das informações disponibilizadas.

Salienta-se que essas escalas não são comparáveis e que, como avaliam a qualidade ou quantidade de alguns dos itens que compõem a medida, ainda serão necessárias descrições mais específicas e detalhadas entre os intervalos, e não apenas entre os níveis, considerando ainda a existência ou não das cinco dimensões operacionais que compreendem o conceito infraestrut-

tura escolar. A proposta feita nesta pesquisa é apenas a primeira a adotar o MRG, mostrando que ainda cabem aprimoramentos para o indicador. Dessas lacunas, advém a necessidade de mais estudos e instrumentos que abordem as definições constitutivas e operacionais, abarcando todas as dimensões do traço latente observado.

REFERÊNCIAS

ALBANEZ, A.; FERREIRA, F.; FRANCO, F. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e Planejamento Econômico**, v. 32, n. 3, p. 453-475, dez. 2002.

ANDRADE, M. et al. Gênero e desempenho em matemática ao final do ensino médio: quais as relações? **Estudos em Avaliação Educacional**, São Paulo, v. 27, p. 77-95, jan./jun. 2003.

BAKER, F. **The basics of item response theory**. ERIC Clearinghouse on Assessment and Evaluation, University of Wisconsin, 2001.

BEATON, A. E.; ALLEN, N. L. Interpreting scales through scale anchoring. **Journal of Educational Statistics**, v. 17, p. 191-204, 1992.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Secretaria de Educação das Unidades de Federação. **Censo Escolar 2017**. Brasília: MEC, 2018.

CASTRO, E. S. **Análise da validade dos indicadores de infraestrutura escolar brasileira**. Estudo em andamento, na composição da tese. 2018

_____. **Estudo do resultado na Prova Brasil de 2011 das escolas com estudantes beneficiários do programa bolsa família**. Brasília, 2014. 120 f. Dissertação (Mestrado em Desenvolvimento, Sociedade e Cooperação Internacional) – Centro de Estudos Avançados Multidisciplinares, Universidade de Brasília, 2014.

CASTRO, E. S. et al. **A série histórica da infraestrutura escolar brasileira**. Estudo em andamento, na composição da tese. 2018

KARINO, C. A. **Avaliação da igualdade, equidade e eficácia no sistema educacional brasileiro**. 2016. 147 f. Tese (Doutorado em Psicologia Social, do Trabalho e das Organizações) – Universidade de Brasília. Brasília, 2016.

KOWALTOWSKI, D. **Arquitetura escolar**: o projeto do ambiente de ensino. São Paulo: Oficina de Textos, 2011.

NETO, J. J. S. et al. Uma escala para medir a infraestrutura escolar. **Estudos em Avaliação Educacional**, São Paulo, v. 24, n. 54, p. 78-99, jan./abr. 2013.

REYNOLDS, D. et al. Educational Effectiveness Research (EER): a state-of-the-art review. **School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice**, v. 25, n. 2, p. 197-230, 2014.

SAMEJIMA, F. A. Estimation of latent ability using a response pattern of graded scores. **Psychometrika**, v. 35, n. 1, p. 139, mar. 1970.

SAMMONS, P. et al. **Key characteristics of effective schools**: a review of school effectiveness research. London: Office for Standards in Education [OFSTED], 1995.

SASTRE, J. R. de S. **O edifício escolar nas políticas públicas de educação**: Rio Grande do Sul 1999–2006. 2010. 472 f. Tese (Doutorado em Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. **Revista Electronica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 2, n. 2, p. 83-104, jul./dez. 2004.

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

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

Recebido em: 15 de dezembro de 2019

Aceito em: 15 de junho de 2020

Publicado em: 30 de junho de 2020

MENSURAÇÃO EM SISTEMAS ADMISSIONAIS NO ENSINO SUPERIOR E AVALIAÇÕES DE IMPACTO

Tufi Machado Soares¹

Mariana Calife Nóbrega Soares²

¹Professor Titular do Departamento de Estatística da UFJF e pesquisador da Fundação CAEd/UFJF, Juiz de Fora, Minas Gerais, Brasil. E-mail: tufi@caed.ufjf.br.

²Doutora em Educação pela PUC-Rio e analista de avaliação da Fundação CAEd/UFJF. Juiz de Fora, Minas Gerais, Brasil. E-mail: mariana.soares@caed.ufjf.br .

Resumo

Este trabalho apresenta um estudo que revisa os principais sistemas de admissão ao ensino superior no mundo. Em particular, revisa os diferentes processos de admissão e os exames e provas usados além dos métodos de mensuração empregados. Além disso, aborda a necessidade de se avaliar a qualidade desses testes e suas medidas e revisa alguns desses métodos. Finalmente, uma análise contextualizada da prova do ENEM, assim como seu uso no SISU, é realizada utilizando os resultados encontrados no estudo.

Palavras-chave: Avaliação de Impacto. Medidas. ENEM. Qualidade. SISU.

Abstract

This work presents a study that reviews the main admission systems to higher education Institutions in the world. In particular, it reviews the different admission processes and the exams and tests used in addition to the measurement methods employed. Moreover, it addresses the need to evaluate the quality of these tests and their measures and reviews some of these methods. Finally, it performs a contextualized analysis of the ENEM test, as well as its use in SISU using the results found in the study.

Keywords: Impact Evaluation. Measures. ENEM. Quality. SISU.

INTRODUÇÃO

A admissão no ensino superior no mundo tem características bastante variadas, não havendo padrões específicos que são seguidos exatamente, mesmo por grupos de países política e culturalmente parecidos. Além disso, mesmo dentro de um país, variações no processo de admissão dos candidatos são observadas. Assim, cada país cria ou adapta os critérios de admissão conforme suas necessidades e tradições.

Alguns padrões gerais serão discutidos com vistas a embasar a discussão da prova do ENEM. No contexto dos países da OCDE (SARGEANT et. al., 2012), uma primeira classificação dos sistemas de admissão é a separação entre **sistemas abertos** e **sistemas seletivos de admissão**: um sistema aberto existe quando resultados ou certificados obtidos no ensino médio garantem o acesso automático à educação superior desde que certas credenciais sejam alcançadas pelos candidatos; já em um sistema seletivo outros critérios mais rígidos são utilizados para a admissão.

Não significa dizer, portanto, que um sistema aberto garante a Educação Superior a todos que desejam fazê-la, mas que a seleção/credenciamento pode ocorrer em diferentes momentos da educação básica. A Alemanha, por exemplo, tem um sistema considerado aberto para a admissão no ensino superior, no entanto, o sistema pode ser considerado altamente seletivo ao longo da educação básica. De fato, a taxa de graduação é de 30% na Alemanha, contra 38% nos Estados Unidos, que apresenta um sistema mais seletivo na admissão (MCGRATH et. al., 2014).

Por outro lado, mesmo em um sistema aberto, para determinadas universidades de grande prestígio e para certos cursos mais disputados, ou que requeiram habilidades especiais, como é comum no caso dos cursos de Medicina, por exemplo, há processos seletivos complementares aos resultados obtidos no ensino médio. Além disso, muitos sistemas abertos utilizam os resultados de uma prova de credenciamento ao final do ensino médio, como é o caso do exame Baccalaureate, na França.

Segundo uma tipologia adaptada de Palmer, Bexley e Jame (2011) os diferentes sistemas de admissão ou não utilizam exames ou testes, como os que utilizam os resultados do ensino médio, como é o caso da Noruega e do Canadá, ou usam testes ou exames: exames de saída do ensino médio³; exames para en-

³ *Secondary leaving examinations.*

trada no ensino superior⁴, que em geral medem o conhecimento adquirido em certas disciplinas ou áreas de conhecimento; testes padronizados de aptidões ou habilidades⁵, que incluem medidas de habilidades ou competências cognitivas, compreensão e outras habilidades gerais; múltiplos exames realizados ao longo do ensino médio e/ou diferentes exames realizados em diferentes momentos no processo de entrada no ensino superior⁶. De qualquer forma, mesmo dentro de um país, a existência de um sistema unificado é exceção.

Edwards, Coates e Friedman (2012) classificam os sistemas de admissão, segundo o uso dos testes da seguinte forma: sistemas que utilizam testes como o único determinante para a entrada e citam, especificamente, China, Grécia, Portugal e Coréia do Sul; sistemas que utilizam os testes como um critério adicional para a admissão e citam Japão, África do Sul, Suécia, Turquia e Estados Unidos; e sistemas que utilizam menos frequentemente os testes no processo admissional citam, como exemplos, a Austrália, Canadá, México e Reino Unido. Nestes últimos, os testes são mais comuns para determinadas universidades e cursos específicos.

Em um estudo⁷ sistemático realizado por McGrath et. al. (2014), os sistemas de ingresso (ou admissão) à educação superior de 10 países foram analisados e comparados detalhadamente, entre eles: França, Itália, Alemanha, Eslovênia, Suécia e Reino Unido, Turquia, Austrália, Japão e Estados Unidos. Três aspectos principais usados para a comparação dos sistemas de ingresso foram a capacidade de produzir equidade, a qualidade dos candidatos selecionados e a capacidade de gerar mobilidade dentro da União Europeia. Alguma evidência foi achada de que a forma de ingresso está associada a essas três características do sistema de educação superior, mas o mais interessante, nesse caso, é que o parlamento europeu entende que essas são características importantes para a composição de um sistema de educação superior. Normalmente, nos EUA, como se verá na seção 3, um aspecto importante dos testes aplicados aos alunos é que seus resultados estejam altamente correlacionados ao sucesso acadêmico futuro, como o êxito na conclusão total ou parcial de determinadas etapas do curso selecionado. Mais importante do que um ranqueamento inicial dos candidatos, portanto, nesses países, é a previsibilidade do êxito na conclusão do ensino superior.

⁴ *Entrance examinations.*

⁵ *Standardized aptitude tests.*

⁶ *Multiple examinations.*

⁷ Estudo patrocinado e apoiado pelo *Policy Department B: Structural cohesion policies – Culture and Education* do Parlamento Europeu.

Outra caracterização importante dos sistemas de admissão é o grau de centralização do processo. Centralização quer dizer que existe um ou mais processos unificados para organizar o ingresso em um grupo de universidades. De fato, portanto, podemos dizer que um sistema é mais centralizado ou mais descentralizado que outro. Dificilmente existem sistemas totalmente centralizados em um único sistema admissional regidos pelos mesmos critérios. Sistemas mais centralizados podem ainda dar pouca autonomia às universidades ou mais autonomia no processo de escolha, mas, em geral, sistemas mais descentralizados dão autonomia para a escolha dos candidatos, a escolha dos critérios de admissão e a operacionalização da aplicação.

OS PAÍSES E SEUS SISTEMAS DE ADMISSÃO

São vários os tipos de instrumentos de admissão existentes para auxiliar as universidades a escolherem seus estudantes, incluindo exames de saída do ensino médio ou exames de certificação ou exames de entrada e testes padronizados de aptidões. Em alguns casos, os alunos podem ser admitidos sem qualquer exame, mas, ainda assim, selecionados por meio de demonstração de requisitos obtidos ao longo de sua educação básica formal. Por outro lado, mesmo quando se considera um mesmo país, há diferentes formas de requisitos e seleção, dependendo se o curso ou a universidade possui uma grande procura, prestígio ou requisitos específicos. Apresentam-se, sucintamente, as principais características dos sistemas de admissão de alguns países:

- V. Na **Austrália**, existem dois tipos de admissão, a mais comum ocorre por meio do “*Year12 Applicants*” (um sistema centralizado de admissão, destinado aos alunos do último ano do ensino médio) e o “*Mature Applicants*” destinado a alunos mais velhos. Normalmente, quando a entrada acontece por meio do “*Year12 Applicants*”, os candidatos apresentam um certificado que os habilita para a educação superior denominado “*Australian Tertiary Admissions Rank*” (ATAR). Os escores do **ATAR** são calculados comparativamente aos estudantes do mesmo ano em dado estado ou território e são calculados em termos de níveis de percentis. Algumas universidades ou centros utilizam ainda testes padronizados denominados “*Special Tertiary Admissions Test*” (STAT), e outras exigem exames da própria instituição ou outros exames específicos, geralmente para medição de habilidades ou conhecimentos específicos em determinadas áreas.

- VI. Na **Inglaterra**, a forma mais comum de admissão é a submissão dos resultados dos exames do GCE A-Levels, ou equivalentes, realizados ao final do ensino médio para as universidades que, em sua maioria, estabelecem os níveis mínimos de aceitação. Além disso, um grande número de universidades realiza seus próprios exames.
- VII. Na **França**, os candidatos devem ter o certificado de nível médio obtido por meio do exame Baccalaureate, ou o diploma de nacional (*Diplôme d'accès aux études universitaires – DAEU*) ou equivalentes internacionais. Note-se que nem todos os que concluíram o ensino médio francês se habilita, assim, para o sistema universitário. O sistema francês é em sua maioria um sistema aberto para acesso ao ensino superior, desde que os candidatos tenham o certificado. No entanto, outros exames de entrada e/ou entrevistas são exigidos para cursos e instituições de grande prestígio, como as chamadas *grandes écoles*.
- VIII. Na **Alemanha**, encontra-se um sistema aberto de admissão, os candidatos devem possuir um certificado de qualificação para a entrada na educação superior chamado Abitur. Alternativas existem para candidatos mais velhos e experientes, que podem entrar por meio de outros tipos de testes ou qualificação.
- IX. Na **Itália**, geralmente, a admissão transcorre a partir da qualificação obtida com a conclusão do ensino médio e algum certificado emitido pelas escolas. Alguns cursos possuem a admissão regulada nacionalmente por meio de testes padronizados, como medicina e cirurgia, odontologia, medicina veterinária, enfermagem e arquitetura.
- X. Todas as universidades no **Japão**, públicas ou privadas, devem seguir o guia para a implementação da seleção dos alunos determinado pelo Ministério da Educação, Cultura, Esportes e Ciência e Tecnologia (MEXT) no final do mês de maio. Apesar disso, todas as universidades têm liberdade para se decidirem pelos critérios de seleção, que geralmente são baseados em exames de entrada geral das universidades, recomendação dos diretores das escolas de ensino médio, testes padronizados produzidos pelo Centro Nacional de Testagem para Admissão nas Universidades ou combinação desses requisitos. O sistema de seleção é quase todo descentralizado.

- XI. Os requisitos de entrada na **Suécia** são baseados nos *school-leaving grades*, resultados escolares do final do ensino médio, para a maioria dos cursos e por meio dos resultados do *Swedish Scholastic Aptitude Test* (SweSAT). Requisitos adicionais são exigidos para algumas áreas em particular como saúde, direito e áreas específicas que exigem habilidades artísticas.
- XII. O sistema educacional Norte Americano (**EUA**) é provavelmente o mais descentralizado de todos. As instituições definem os critérios, selecionam e controlam a aplicação dos alunos. No entanto, para alguns cursos de algumas universidades, existem alguns sistemas unificados. Entre os requisitos de entrada nas universidades norte-americanas são muito comuns os resultados dos testes de aptidões, sendo os mais comuns o *Scholastic Aptitude Tests* (SATs) e os testes de *Subject American College Testing* (ACTs), realizados independentemente das universidades por organizações ou empresas especializadas em testagem. Outros testes de aptidões e/ou conhecimentos são requisitados para determinados cursos, como direito e medicina. Os resultados dos testes, geralmente, são parte dos requisitos, mas não constituem o único requisito de entrada.
- XIII. Na **Turquia**, o sistema é totalmente centralizado. O sistema nacional seleciona os critérios, os alunos e controla a aplicação. A entrada na universidade é baseada no ÖSS, um exame compulsório baseado em testes para habilidades verbais e quantitativas. O teste abrange uma variedade de disciplinas, incluindo ciências, matemática, língua turca, língua estrangeira e ciências sociais. Os resultados do ÖSS são combinados com as médias do certificado *Lise Diploması*, de saída do Eesino básico. O *Yükseköğretim Kurulu* (YÖK), conselho de educação superior, é responsável por coordenar e determinar os critérios e realizar a seleção dos candidatos em si.
- XIV. No **Chile**, existe a prova de seleção Universitária (*La Prueba de Selección Universitaria* – PSU), cujos resultados são usados para a admissão por muitas das Universidades do país. Os testes de Matemática e Linguagem e comunicação são obrigatórios, enquanto os testes de ciências (biologia, física, química e técnico-profissional) e história, geografia e ciências sociais são eletivos. Os alunos escolhem as eletivas de acordo com os requisitos das universidades. Existe um sistema único de admissão, simultâneo e integrado, para as Universidades Chilenas Públicas do *Consejo de Rectores* (CRUCH) e Universidades privadas associadas. O sistema é

coordenado pelo *Departamento de Evaluación, Medición Y Registro Educacional* (DEMRE) da *Univesidad de Chile* que também coordena a prova PSU (cf. UNIVERSIDAD DE CHILE, [c2017]b). As Universidades consideram para a admissão, além das pontuações na PSU, o ranking de pontuações obtidas pelos colegas na escola de ensino médio (medida do contexto) e as notas obtidas nas escolas de nível médio.

Variações no processo seletivo dentro de um mesmo país ocorrem, frequentemente, entre os cursos. A disciplina de ciências médicas representa o principal exemplo, em muitos países, em que testes de aptidão complementam os certificados do ensino médio. Na Inglaterra, os estudantes têm que realizar testes como o *Bio-Medical Admissions Test* (BMAT), o *The Helth Professions Admission Test* (HPAT) ou o *UK Clinical Aptitude Test* (UKCAT). Na Itália, os cursos de medicina e cirurgia, veterinária e odontologia estão entre os que têm sua entrada regulada a nível nacional por meio de testes padronizados que focam conhecimentos gerais (história, raciocínio lógico etc.) e conhecimentos específicos relevantes para a área em questão. Na Austrália, o teste *Undergraduate Medical Admissions Test* (UMAT) é requisito para muitos cursos de odontologia e médicos. Nos EUA, o *Law School Admission Test* (LSAT) é muito usado na admissão dos candidatos nas escolas de Direito e o *Medical College Admission Test* (MCAT) para as escolas de medicina.

OS TESTES E SEUS SISTEMAS DE MENSURAÇÃO

Os testes, dentro das Teorias dos Testes de Avaliação, incluindo os de seleção, são construídos para medir (avaliar) algum constructo: o aprendizado em uma disciplina, a proficiência em física, o conhecimento em matemática, a aptidão em relações humanas, a habilidade de raciocínio quantitativo, a habilidade de interpretação de textos em língua portuguesa etc. Genericamente, utilizaremos o termo **proficiência cognitiva** para o constructo a ser medido, tendo em vista que nos exames admissionais para o ensino superior esses constructos geralmente representam alguma dimensão da cognição humana. Portanto, os testes são, essencialmente, instrumentos de medição de uma característica do ser humano que não pode ser observada diretamente por outro meio (por isso ela é denominada, então, de latente).

Os testes têm usos diversos. Testes de avaliação educacional em larga escala, como o PISA⁸, o NAEP⁹, o SAEB¹⁰ e a Prova Brasil, são normalmente desenhados para avaliar sistemas educacionais. Por outro lado, há os testes que necessitam ser desenhados para avaliar uma ou mais características de um indivíduo especificamente. É o caso de testes de diagnósticos psicológicos, psiquiátricos e, naturalmente, aqueles que orientam a seleção para ingresso no ensino superior. Essa diferença é substancial, pois a composição do teste, o conjunto de habilidades a serem medidas, as dificuldades dos itens entre outras características dos itens dos testes devem ser calibradas para atender às suas finalidades específicas.

Os testes de admissão são, geralmente, divididos entre os testes de realização ou êxito (*achievement*), mais comuns no Brasil, também denominados de testes de conhecimento (*subject*), como os que medem o conhecimento adquirido em uma disciplina em uma etapa escolar, por exemplo, e que naturalmente se referem a uma realização passada do estudante; os testes de habilidades (*abilities*) e competências cognitivas que medem a capacidade de realizar determinadas tarefas como, por exemplo, o raciocínio quantitativo e que, em geral, representam uma habilidade atual do indivíduo; e os testes de aptidão (*aptitude*) que procuram revelar uma capacidade do indivíduo para uma realização exitosa no futuro. Evidentemente, nem sempre os testes são exclusivamente de um tipo somente, mas costumam transitar entre essas três características, mesmo que na maioria dos casos uma delas predomine sobre as demais.

Definido o constructo ou constructos que são relevantes, os testes de seleção, por exemplo, são construídos para escolher, entre os candidatos, aqueles que apresentam os maiores valores de proficiência para o constructo ou, ainda, para distinguir aqueles que alcançaram determinados níveis de habilidades mínimas requisitadas para o curso daqueles que não alcançaram essas habilidades.

Como não é possível observar diretamente o constructo de interesse, usa-se o teste como instrumento para se obter uma avaliação do constructo. Chama-se a esse processo de medição, sendo o teste o instrumento de medição, e o resultado que se extrai dele, no caso uma pontuação calculada com base nos resultados das questões, uma medida do constructo. Assim, a pontuação do

⁸ *Program for International Student Assessment (Pisa)* – OECD.

⁹ *National Assessment of Educational Progress (NAEP)* – USA.

¹⁰ **Sistema de Avaliação da Educação Básica.**

resultado de um teste respondido por um candidato é, então, uma avaliação (medição) do constructo que se quer conhecer. Na próxima subseção, as técnicas de pontuação serão discutidas e suas implicações analisadas.

Formas de pontuação e medição dos testes

Infelizmente, toda medida é uma representação parcial do valor que se quer de fato conhecer, denominado por alguns autores de **valor verdadeiro** do constructo, embora não haja consenso sobre essa interpretação. Assim, a toda medida se associa uma incerteza, que é inerente a qualquer processo de medição. Esse argumento vale para medidas em todas as áreas do conhecimento e suas aplicações: desde a física até as ciências humanas e sociais. As diferentes teorias das medidas, nas diferentes áreas e aplicações procuram, então, quantificar essas incertezas para que novos métodos de medição possam ser propostos com o objetivo de produzir resultados mais acurados e as medidas possam ser usadas de forma mais adequada.

A incerteza da medida é usualmente denominada no campo da Psicometria e da Estatística como o **erro da medida**. O termo erro é, infelizmente, mal entendido muitas vezes. Ele não quer dizer que a medida esteja errada em si mesma, mas que toda medida tem associada uma incerteza em relação ao valor verdadeiro do que se quer de fato conhecer e que não pode ser observado diretamente.

Na Psicometria, que é a ciência que estuda os processos de medição de constructos latentes por meio de testes psicométricos, dos quais os testes de avaliação e seleção são exemplos, duas teorias principais norteiam a construção e a verificação da qualidade das medidas. Com a primeira, denominada de Teoria Clássica dos Testes (TCT), procura-se analisar os erros das medidas calculadas pelos métodos clássicos de pontuação dos testes e, assim, sugerir formas de pontuação mais acuradas e fidedignas dos valores verdadeiros, além de preconizar métodos de pontuação que permitam a comparabilidade de resultados quando os instrumentos (testes) são diferentes, por exemplo.

A segunda teoria mais comum no campo da Psicometria é a Teoria da Resposta ao Item (TRI) que, apontando determinadas limitações nas formas clássicas de pontuação, preconiza formas diferentes de se pontuar os testes e construir medidas para os constructos latentes. Particularmente, ela é muito útil na comparabilidade de resultados de testes diferentes aplicados a diferentes alunos.

A seguir, analisadas no campo das duas teorias, serão discutidas as principais formas de pontuação, suas limitações e vantagens.

Forma Clássica de Pontuação e Teoria Clássica do Teste (TCT)

Os escores reportados como resultados dos testes são derivados do desempenho obtido pelo candidato no teste por meio de um processo estatístico chamado pontuação (*scaling*). O exemplo mais simples é aquele em que um candidato que responde corretamente a 55 itens, em um teste de 60, recebe uma pontuação (*score*) de 55. Nesse caso, a pontuação é chamada de pontuação bruta (*raw score*). Outros tipos de *raw scores* são o percentual de acerto no teste ou uma pontuação ponderada dos itens, por exemplo.

Para testes padronizados, e testes admissionais no mundo, a pontuação bruta quase nunca é utilizada diretamente. Isso ocorre porque os pontos brutos são diretamente dependentes dos itens presentes em uma forma¹¹ particular de um teste. Por exemplo, uma determinada forma pode conter itens mais fáceis que outra forma e, portanto, a mesma pontuação, em ambas as formas do mesmo teste, não expressa o mesmo nível de habilidade dos candidatos. Principalmente por isso, quase sempre, os pontos brutos são transformados para uma escala específica. Este processo é denominado **escalamento** (*scaling*).

Grosso modo, a TCT lida com a incerteza da medida da seguinte forma: admita que seja possível construir formas paralelas de um teste, isto é, formas de um teste cujos resultados sejam indistinguíveis. Ainda assim, estudos empíricos mostram que os pontos obtidos não são exatamente os mesmos quando o mesmo aluno responde a essas formas. Portanto, o mesmo aluno, respondendo a formas absolutamente equivalentes, ainda assim apresenta resultados diferentes.

A teoria clássica, então, admite que o **escore verdadeiro (T)** do candidato é a pontuação média obtida em todas as possíveis formas paralelas do teste; na prática, é a média dos pontos brutos que o aluno teria ao responder a uma quantidade grande de formas paralelas. Assim, matematicamente, pode-se relacionar os pontos brutos obtidos em um teste (**O**), chamado de escore observado, como sendo $O = T + e$, onde **e** é denominado de erro de medida observado na pontuação de um teste.

¹¹ Formas de um mesmo teste são conjuntos de itens (geralmente organizados em um caderno de teste), pontuados segundo uma escala específica, medindo a mesma habilidade cognitiva, portanto, referenciados aos mesmos conteúdos. As formas são denominadas **paralelas**, quando apresentam as mesmas características psicométricas, como a dificuldade dos itens, de tal forma que sejam indistinguíveis do ponto de vista da medida produzida. Na prática, formas paralelas são muito difíceis de serem construídas e, por isso, faz-se necessário algum processo de equalização que estabeleça a comparabilidade das pontuações.

A pontuação bruta obtida no teste pelo candidato é, nesse sentido, uma estimativa para o escore verdadeiro desconhecido. Sob determinadas hipóteses, admitidas sobre a natureza probabilística do erro e sobre o escore verdadeiro, é possível se estimar, por meio de técnicas estatísticas, as propriedades estatísticas do erro **e**, como seu desvio-padrão, denominado de **erro padrão** da medida. Para uma explicação mais pormenorizada desses procedimentos, sugerem-se as seguintes referências: Lord e Novick (1968), Thissen e Wainer (2001).

Para sistemas de testagem que necessitam ter múltiplas aplicações, com diferentes formas do teste, é necessário manter a comparabilidade dos resultados para que possam ser usados adequadamente nos processos de admissão, por exemplo. Quando se trata do mesmo teste aplicado, a comparabilidade dos resultados obtidos pelas diferentes formas é produzida por meio de um processo estatístico chamado de **equalização** (*equating*). O conceito exato de equalização pode variar um pouco, conforme o autor. Mas, na tradição Norte Americana, conforme Kolen e Brennan (2004), a equalização é um procedimento de comparabilidade que ajusta os escores nas diferentes formas de um mesmo teste, que são construídas para serem similares em dificuldade e conteúdo. O escalonamento, portanto, geralmente está associado a um processo de equalização utilizado para manter comparáveis os resultados de diferentes formas de testes aplicadas a diferentes grupos de candidatos. Quando as formas não são similares, ou os testes não são exatamente os mesmos em conteúdo, os métodos de comparabilidade ganham outras denominações como *linking*, moderação, projeção etc. Também é comum o termo “equalização horizontal”, quando os grupos de alunos que realizam as diferentes formas do teste são similares em níveis de proficiências, e “equalização vertical”, quando os grupos de alunos não são similares.

Naturalmente, ao se fazer uma equalização, assim como ao se realizar qualquer outro método de comparabilidade, novas incertezas são introduzidas nas medidas construídas. Por esse motivo, busca-se construir formas de testes que sejam as mais parecidas, do ponto de vista psicométrico e do conteúdo, manter a padronização na aplicação dos testes e empregar os métodos de equalização mais precisos que se disponham. Infelizmente, muitas vezes não se sabe qual é exatamente o melhor método que possa ser empregado em uma situação ou outra. Na teoria, alguns métodos foram comparados em situações ideais e, geralmente, os psicometristas procuram seguir uma fonte ou outra, além da sua própria *expertise*.

Existem muitas maneiras de se produzir a equalização dos resultados. Não se vai aqui discuti-las, tendo em vista a necessidade de conhecimento técnico

específico para isso. Uma boa referência para esses métodos é Kolen e Brennan (2004), que apresentam os métodos de equalização tanto para a forma clássica de pontuação quanto para a Teoria da Resposta ao Item.

Modelos de Resposta ao Item

A Teoria da Resposta ao Item (ver Lord, 1980) surge da necessidade de superar as limitações de se pontuar os resultados dos alunos em testes educacionais na forma clássica, como os percentuais de acertos ou escores brutos dos testes, e da dificuldade de se comparar esses resultados para as diferentes formas de um teste aplicadas em diferentes situações, como em diferentes datas, por exemplo. Na teoria clássica, para se contornar esses problemas, hipóteses muito restritivas são impostas, muitas das quais difíceis de serem comprovadas empiricamente.

A TRI muda o foco de análise do teste como um todo para a análise de cada item. A ideia básica consiste no emprego de modelos para o funcionamento dos itens com relação aos valores do constructo que se deseja medir. Geralmente, esses modelos são paramétricos, de tal forma que os parâmetros podem representar características importantes dos itens e, nesse sentido, podem até ser interpretáveis.

A suposição de que, fixada a escala de proficiências, os parâmetros dos itens não variam mesmo quando aplicados a diferentes grupos de alunos, **se observada empiricamente**, garante que as medidas de proficiências construídas por meio desses modelos, para os alunos nos diferentes grupos, estejam na mesma escala e, portanto, sejam comparáveis. Essa propriedade é conhecida como **invariância dos parâmetros dos itens**.

Assim, parte-se do pressuposto de que o modelo do item seja invariante, independente do grupo de alunos a que esteja sendo submetido o teste – suposição esta que pode ou não ser verificada empiricamente –, a TRI permitirá a comparabilidade dos resultados produzidos para grupos de indivíduos diferentes, mesmo quando os testes aplicados são diferentes. É frequente a afirmativa dos autores de que a medida de habilidade produzida por meio da TRI é independente do instrumento aplicado e independente do grupo de indivíduos ao qual é aplicado. Claro que certos cuidados técnicos devem ser tomados para garantir essa propriedade, além da verificação dos pressupostos sob os quais os modelos dos itens são propostos.

Existem muitos tipos de modelos, dependendo de como os itens dos testes são pontuados, do número de constructos a serem medidos, entre outras características. Consideram-se aqui apenas itens pontuados segundo uma regra

dicotômica, certo ou errado, por exemplo, e modelos para medir um único constructo latente, chamados de unidimensionais. Nesse caso, o modelo associa à probabilidade de uma resposta correta no item a proficiência cognitiva e determinados parâmetros dos itens, que representam certas propriedades do item, por meio de uma função matemática específica. Pela importância que possuem nos testes admissionais em questão, vai-se apresentar aqui muito superficialmente o modelo de 3 parâmetros.

Modelo de três parâmetros

Em geral, a maioria dos itens de múltipla escolha em avaliações educacionais em larga escala, pelo menos no Brasil, apresenta um comportamento tal que mesmo alunos com baixa proficiência apresentam uma pequena probabilidade de acertá-los. Claro que, em parte, esse comportamento do item depende da forma como os especialistas elaboram os itens do teste e constroem as alternativas de respostas não corretas (chamadas de distratores¹²). Alguns testes recomendam que o aluno, caso não saiba a resposta correta, não responda ao item, numa tentativa de evitar esse acerto casual, conhecido popularmente como “chute”. No entanto, infelizmente, na prática, essa recomendação dificilmente é observada e, mesmo que o seja, muitos itens provavelmente manterão o mesmo comportamento. Esse fenômeno foi interpretado sob diversas formas: Birnbaum (1968), exatamente nas linhas dos argumentos apresentados acima, sugeriu que os alunos de baixa habilidade estariam acertando este item devido a escolhas casuais, por isso sugeriu introduzir o parâmetro c_i no modelo que será apresentado a seguir; Lord (1980) notou que, em geral, o percentual de acerto nesses níveis de habilidades muito baixas era menor do que o inverso do número de alternativas, provavelmente devido à forma como os itens eram elaborados, fazendo com que um dos distratores se tornasse mais atraente para esses alunos. Por esses e outros motivos, o modelo da TRI para os itens utilizados nas avaliações em larga escala de sistemas, no Brasil e em boa parte do mundo, é o chamado modelo logístico de 3 parâmetros, cuja expressão matemática na representação logística é dada por:

$$P(y_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-D a_i (\theta_j - b_i)}}$$

O modelo acima, $P(y_{ij} = 1 | \theta_j, a_i, b_i, c_i)$, representa a probabilidade do aluno j acertar ao item i tendo uma proficiência θ_j , sendo que o item apresenta um índice de discriminação a_i , uma dificuldade b_i e um parâmetro de *guessing* c_i , que representa a probabilidade de um aluno de proficiência mui-

¹² Tradução literal do inglês *distractors*.

to baixa acertar o item. Para ver mais detalhes dessa representação, incluindo a justificativa para a interpretação dos parâmetros, sugere-se, por exemplo, Baker e Kim (2004).

Pontuação por Meio da Teoria da Resposta ao Item

A forma de se pontuar os resultados dos testes usando a TRI é chamada de métodos de estimação das proficiências. Se conhecidos os parâmetros dos itens, os principais métodos de estimação das proficiências são os métodos de verossimilhança máxima, de máximo da distribuição a posteriori e a média da distribuição a posteriori. O mais usado, na prática, é o método da média da distribuição a posteriori (EAP¹³), que é o método usado no ENEM.

Uma estatística importante que caracteriza a incerteza da medida é chamada de **erro padrão da medida de proficiência**. Os principais *softwares* usados na produção das medidas de proficiências na TRI fornecem estimativas do erro padrão que indicam a acurácia da medida produzida, e inclusive por meio deles pode-se construir intervalos de confiança para o valor verdadeiro da proficiência. Esses intervalos são construídos como forma de indicar a incerteza da medida produzida, fornecendo não uma estimativa pontual, mas um intervalo de valor onde provavelmente o valor verdadeiro está.

As proficiências podem, ainda, ser estimadas conjuntamente com os parâmetros dos itens. Os métodos mais comuns são os de máxima verossimilhança conjunta, máxima verossimilhança conjunta marginalizada, máxima distribuição a posteriori conjunta marginalizada (MMAP), e variações. Em todos os casos, os erros de medida são fornecidos tanto para os parâmetros estimados quanto para as proficiências estimadas, de tal forma que a acurácia das medidas pode ser verificada.

Infelizmente, os métodos de pontuação nesses casos não são baseados em formas explícitas e fáceis de serem apresentadas. Quase sempre, constituem-se de cálculos realizados por métodos iterativos e computacionalmente intensivos. No entanto, os cálculos são totalmente objetivos, uma vez definidos os métodos e os pressupostos básicos. Para uma revisão e explicação pormenorizada desses métodos sugere-se, por exemplo, Baker e Kim (2004). Assim, conhecidos os métodos que são empregados, pode-se reproduzir muito aproximadamente o cálculo de forma independente, desde que se tenham todas as informações e bases necessárias.

Comparação Entre as Formas Clássicas e da TRI de Pontuar

¹³ *Expected A Posteriori*.

Frequentemente, a literatura aponta uma série de vantagens teóricas da pontuação obtida pelos métodos da TRI sobre a pontuação tradicional. Ver, por exemplo, Hambleton e Jones (1993) para uma comparação das duas teorias. No entanto, como afirmam Thissen e Orlando (2001), ambas as teorias (TCT e TRI) apresentam formas de avaliar a precisão das pontuações e podem ser usadas de forma complementar.

A vantagem mais apontada das pontuações obtidas por meio da TRI é que elas permitem comparabilidade de resultados de testes diferentes, respondidos por candidatos diferentes, por procedimentos mais diretos e, na maioria das vezes, com maior precisão do que os artifícios que necessitam ser usados quando se está pontuando os testes na forma tradicional. Isto é, em geral, afirma-se que é mais simples e preciso produzir a comparabilidade de resultados quando se usa a TRI.

O desempenho dos testes computadorizados e adaptativos, por exemplo, que produzem medidas muito mais precisas do que as produzidas pelos métodos tradicionais de papel e lápis é totalmente potencializado pelo emprego da TRI (Thissen; Orlando, 2001). Assim, devido à sua versatilidade, justifica-se o emprego da TRI, quando é necessário aplicar sistematicamente diferentes formas de teste.

Estudos, no entanto, apontam que aparentemente alguns métodos de comparabilidade desenhados para as técnicas clássicas de pontuação, **com os correspondentes processos de escalonamento**, podem alcançar desempenho similar em algumas situações. Porém, ainda há muita controvérsia na literatura. Alguns estudos demonstram pequena superioridade da pontuação obtida pela TRI em algumas aplicações, mas não em todas as situações.

Em um trabalho recente, Jabrayilov, Emons e Sijtsma (2016) mostraram, em um estudo empírico, a superioridade da pontuação obtida pela TRI na avaliação do ponto de mudança em tratamentos psicológicos clínicos¹⁴, quando os testes contêm ao menos 20 itens, mas não quando os testes são inferiores.

Especificamente em testes admissionais, Fan (1998) utilizou os dados do *Texas Assessment of Academic Skills (TAAS) tests*, aplicado em outubro de 1992 aos estudantes do 11º ano escolar. Essa avaliação é obrigatória no estado do Texas, USA, e se constitui de testes de leitura, matemática e escrita. Os testes de leitura (com 48 itens) e matemática (com 60 itens) são todos de múltipla escolha e dicotomicamente pontuados como certo ou errado. A base de dados se constitui das repostas de 193.000 respondentes aos testes de leitura e

¹⁴ *Change assessment in clinical settings.*

matemática. Os achados indicam que tanto as estatísticas dos itens quanto as proficiências derivadas da TCT e da TRI são bastante comparáveis. Os resultados indicaram também que a invariância das estatísticas dos itens ao longo das formas é tão boa para os métodos tradicionais de escalonamento quanto para os métodos da TRI. Por outro lado, Salle (2009) compara, em um teste de química aplicado aos alunos do ensino médio, a pontuação clássica e a obtida pela TRI por meio de duas amostras equivalentes de alunos e duas formas do teste. Os resultados foram comparados considerando as dificuldades dos itens, consistência interna (fidedignidade) dos testes e erros das medidas de proficiências. O modelo da TRI utilizado foi o modelo de Rasch. Apontando menor erro de medida, portanto, nas pontuações da TRI.

Fan (1998), em seu estudo baseado em dados reais, achou alta correlação (maior que 0,96) entre as pontuações dos respondentes obtidas por métodos tradicionais e as pontuações obtidas pela TRI e, da mesma forma, alta correlação entre as medidas de dificuldade dos itens obtida pelos métodos tradicionais e pela TRI (maior que 0,90). Ele também não encontrou grande evidência de uma grande invariância dos parâmetros dos itens calculados por meio da TRI com relação ao da teoria clássica. Resultado similar foi encontrado por Courville (2005) em outro estudo empírico.

Uma limitação dos estudos empíricos é que os valores dos parâmetros não podem ser manipulados e os valores verdadeiros não são conhecidos. MacDonald e Paunonen (2002) realizaram um estudo de Monte Carlo (simulação intensiva) no qual eles controlaram a variação das dificuldades e as discriminações dos itens. De forma similar a Fan (1998) e Courville (2005), eles encontraram alta correlação entre as dificuldades dos itens estimadas pelos métodos da TRI. No entanto, no caso da discriminação, uma superioridade na TRI foi observada.

Grosso modo, então, os estudos indicam que, dependendo do contexto e das características dos testes, os métodos tradicionais podem até ter desempenho similar aos métodos da TRI. Estes autores desconhecem resultados em contrário, isto é, em que a forma de pontuação clássica possa ser superior à obtida pela TRI, a não ser aqueles em que se ficou demonstrado que os modelos e os métodos da TRI empregados foram inadequadamente utilizados. Por isso, tem-se aqui a convicção de que utilizar os métodos preconizados pela TRI num teste com a dimensão e a importância do ENEM é mais apropriado porque conduz a resultados mais justos no geral, embora menos transparentes para a sociedade.

Qualidade das medidas produzidas

Uma preocupação importante na Psicometria, tanto no que se refere à TCT quanto à TRI, é garantir a qualidade das medidas produzidas. No Brasil, não existe previsão legal que regule amplamente um sistema de testagem de tal forma que abarque o uso da TRI.

Tampouco existe no Brasil uma regulação ou um conjunto de sugestões propostas por alguma entidade científica ou profissional Nacional. Nos EUA, existem guias ou *standards* que constituem recomendações de algumas entidades científicas para garantir a qualidade das medidas e dos processos de testagem. O mais importante são os *standards* organizados pela *American Psychological Association (APA)* e pela *American Educational Research Association (AERA)*, sendo o mais recente de 2014.

Aqui reportamos algumas características importantes do *Educational Testing Service (ETS) Standards for Quality and Fairness*, por estar disponível para consulta livre e por ser muito semelhante ao da APA e AERA. Segundo o documento (EDUCATIONAL TESTING SERVICE, 2015), os *standards* proveem padrões de excelência que são usados pelo *staff* da ETS ao longo do processo de planejamento, desenvolvimento e entrega das medidas, de modo a prover resultados de testes “justos”, “válidos” e “fidedignos”. O último documento (de 2014) é relativamente extenso, composto de 13 capítulos, 76 *standards*, cada um correspondente a um aspecto relevante, desde a construção, aplicação e validação de um teste.

Foram selecionados pelo autor deste texto alguns *standards* que ele considera como muito relevantes para a presente discussão, por serem mais críticos na questão da pontuação do ENEM e porque servirão para orientar documentos futuros.

Fidedignidade dos Resultados dos Testes, Validade e Equalização/*Linking*

A fidedignidade é a propriedade de um teste produzir medidas consistentes, com níveis aceitáveis de erros de medida. Tanto a TRI quanto a Teoria Clássica proveem medidas da fidedignidade dos testes para atestar a acurácia dos resultados.

A título de exemplificação, os *standards* preconizados pela ETS preconizam que se usem métodos adequados, que sejam analisadas as suas implicações, **mas, sobretudo, que se informe ao usuário do teste sobre a imprecisão das medidas calculadas e como ele pode lidar com ela.**

Uma maneira de quantificar a quantidade do erro de medida de uma estimativa de proficiência é usar o chamado erro padrão da medida. O erro padrão provê uma estimativa do erro que está presente nos escores obtidos pelo método de pontuação do teste devido à natureza da imperfeição de qualquer medida na natureza e, em particular, na medida da proficiência cognitiva obtida por meio de qualquer tipo de teste.

No caso do LSAT, por exemplo, o erro padrão é usado para construir intervalos (*score bands*) que podem ser usados para reportar os escores, de tal forma que a incerteza associada à medida é reportada juntamente com os resultados. Isto será visto em uma exemplificação dos resultados de um teste na seção 4.3 (Law School Admission Council, 2017b).

Muitos outros fatores podem afetar os resultados nos testes em um dia em particular, como a motivação, a saúde física e mental, a ansiedade produzida por diferentes fatores, algum transtorno produzido por uma obra ou pelo trânsito etc. Esses outros fatores, normalmente, não têm como serem levados em consideração no cálculo do erro padrão.

A fidedignidade de todo um teste pode ser estimada por diferentes medidas como, no caso de pontuações clássicas, pelo alfa de Cronbach e, no caso de pontuações obtidas pela TRI, a função de informação do teste.

Grosso modo, a validade de um teste consiste em estabelecer que aquilo que o teste mede é, de fato, o que se deseja que ele meça. Especificamente, existem muitos tipos de validades, citam-se aqui alguns tipos mais importantes: Validade de Constructo (o teste mede as habilidades e competências que deveriam medir); Validade de Conteúdo (o teste mede os conteúdos apropriadamente); Validade Preditiva (o teste prevê o sucesso do candidato razoavelmente, no caso de testes de admissão, por exemplo, o sucesso na conclusão do curso); Validade de consequências (o teste apresenta consequências adversas mínimas); Validade externa (o teste tem associação elevada com outras medidas do mesmo constructo).

Mesmo que seja impossível realizar estudos de validade amplos para todos esses tipos, em pouco tempo, deve ser uma preocupação constante a realização desses estudos ao longo do tempo, informando aos usuários das medidas sobre eles.

Outro *standard* da ETS (4.3: *Obtaining and Documenting the Evidence*) enfatiza a importância de se investigar e documentar a validade do teste.

Estudos de validade podem ser conduzidos tanto pelos responsáveis pelos testes, até para justificar a qualidade do seu produto, quanto pelos usuários (no caso as Universidades), para determinar se a escolha é apropriada para o sucesso acadêmico do indivíduo.

Já se apresentaram os conceitos de equalização e de *linking*. Vai-se, apenas, fazer a apresentação dos *standards* 8.1 a 8.5¹⁵, sobre a importância de se documentar os métodos de equalização utilizados.

Embora não preconize especificamente os métodos de comparabilidade que devem ser usados, os *standards* deixam clara a necessidade de explicitá-los para que possam, inclusive, ser criticados. Além disso, enfatizam a necessidade de explicitar o erro de medida para que o usuário faça uso apropriado da medida.

Finalmente, os *standards* abordam outras questões importantes como sigilo, segurança na aplicação e nos resultados, documentação, conferência da pontuação obtida, entre outras.

Exemplos de testes de admissão ao ensino superior

Alguns testes de aptidão e conhecimentos gerais

Nesta seção selecionaram-se alguns exames e testes para serem analisados mais detalhadamente. Nos EUA, o uso de testes padronizados de aptidão (*aptitude*) e/ou de conhecimentos gerais são os mais comuns usados nos processos de admissão das universidades. Há, ainda, testes de aptidão ou conhecimentos específicos (*subject*) muito usados em disciplinas ou áreas específicas adicionalmente aos testes mais gerais. Existem muitos, mas destacamos como os testes de aptidão e/ou conhecimentos gerais nos EUA mais tradicionais os seguintes:

- I. O *Scholastic Aptitude Test (SAT)*, administrado pelo *College Entrance Examination Board* desde 1926, é um teste que tem o seu foco na habilidade de raciocínio (Liu; Harris; Schmidt, 2007). Segundo o *College Board*, o SAT mede as habilidades e os conhecimentos que os pesquisadores mostram que são os mais importantes para o sucesso no curso superior

¹⁵ *Standard 8.1: Using Appropriate Equating or Linking Methodologies; Standard 8.2: Documenting Equating or Linking — Population and Comparability; Standard 8.3: Documenting Equating or Linking; Linking — Data Collection Design Describe the data collection design for the equating or linking study and state explicitly the assumptions implied by the use of that design.; Standard 8.4: Documenting Equating or Linking — Statistical Procedures; Standard 8.5: Documenting Equating or Linking — Results* (EDUCATIONAL TESTING SERVICE, 2015).

e na carreira¹⁶. Ele inclui as seguintes seções: Leitura e escrita, baseadas em evidências, e Matemática. Além disso, possui seções de ensaios (ou redações) opcionais, que medem habilidades de leitura, análise e escrita.

Existem, ainda, os chamados “SAT *Subjects*”, que medem conhecimentos e habilidade em disciplinas e áreas específicas. Há, no momento, 20 testes destinados a diferentes disciplinas ou áreas do conhecimento divididas em cinco grandes áreas: Inglês, História, Linguagens e Matemáticas. Cada teste tem uma hora de duração e são todos de múltipla escolha e pontuados numa escala de 200-800 pontos (The College Board, c2017a). Os testes são realizados em diferentes datas e, assim como no caso do SAT, o candidato pode realizar o teste sempre que quiser e pode escolher a sua pontuação que julgar mais conveniente para ser enviada às Universidades¹⁷. Também existem testes destinados a alunos de séries inferiores da educação básica, conhecidos como *Preliminar SAT tests* (PSATs).

O SAT é pontuado utilizando-se **a teoria clássica do teste**, no entanto, os resultados das diferentes seções e testes são escalonados (*scaled*) para escalas comparáveis (isto é, são equalizados) de tal forma que, segundo seus organizadores, os resultados das diferentes formas (isto é, dos testes realizadas em diferentes datas) possam ser utilizados pelas universidades indistintamente. Os métodos de equalização empregados pelo SAT são genericamente descritos por Liu, Harris e Schmidt (2007). Portanto, não são os pontos brutos que são apresentados, **mas os escores escalonados para a escala do SAT**.

Os alunos não respondem a uma única forma de teste. Isto é, diferentes grupos de itens são apresentados a diferentes grupos de alunos. No entanto, um procedimento espiralar de aplicação do teste (*spiraling procedure*) garante que os grupos sejam equivalentes em proficiências. Assim, os escores de uma forma escolhida são equalizados para a escala do SAT por meio de um método conhecido como *Nonequivalent Anchor Test Design* (NEAT), sendo os escores das demais formas equalizadas com esta por meio de um *Random/Equivalent Goup Design* (EG).

Tabelas de conversão entre os pontos brutos alcançados em um teste específico e os pontos equivalentes das escalas são fornecidas de tal forma que o candidato pode conferir a sua pontuação. No Quadro 1, apresenta-se a relação dos resultados que são reportados pelo SAT.

¹⁶ *The SAT measures the skills and knowledge that research shows are the most important for success in college and career.*

¹⁷ Ver: The College Board, 2017.

Quadro 1 – Relação dos resultados reportados pelo SAT

SAT Score Reported	Details	Score Range
Total score	Sum of the two section scores.	400-1600
Section scores	Evidence-Based Reading and Writing, and Math.	200-800
Test scores	Reading, Writing and Language, and Math.	10-40
SAT Essay scores*	Reading, Analysis, and Writing.	2-8
Cross-test scores	Analysis in History/Social Studies and Analysis in Science. Based on selected questions in the Reading, Writing and Language, and Math Tests. These scores show how well you use your skills to analyze texts and solve problems in these subject areas.	10-40
Subscores	Reading and Writing and Language: Command of Evidence and Words in Context. Writing and Language: Expression of Ideas and Standard English Conventions. Math: Heart of Algebra, Problem Solving and Data Analysis, and Passport to Advanced Math.	1-15

* The SAT Essay is optional.

Fonte: The College Board (c2017c).

Os escores são apresentados na escala do SAT e, em alguns casos, comparativamente em percentis (*percentile ranks*), para as distribuições dos escores em relação aos alunos que fizeram os testes nos últimos anos.

Quando não concorda com a pontuação recebida, o candidato pode recorrer, solicitando uma verificação, conforme procedimento específico.

Para se ter uma ideia, cerca de 2,4 milhões de indivíduos realiza o SAT e 1,2 milhões o ACT todos os anos.

II. O *Achievement College Test (ACT)* é um teste de avaliação com características similares ao SAT, realizado pela ACT Inc. Porém, segundo Liu, Harris e Schimidt (2007), ele tem foco na realização (*achievement*)¹⁸ e inclui quatro seções obrigatórias de testes de múltipla escolha de Inglês, Matemática, Leitura e Ciências. Assim como o SAT, o ACT é pontuado por meio de estatísticas clássicas e suas pontuações brutas são escalonadas (por meio de um procedimento de equalização) para permitir a comparabilidade de resultados; o método é genericamente descrito em Liu, Harris e Schimidt (2007). Devido ao método de equalização utilizado, é possível produzir uma tabela que mostra a equivalência entre os pontos brutos alcançados pelos candidatos no teste e os pontos equivalentes na escala do ACT. Existem também tabelas de comparação entre os escores do ACT e SAT, possivelmente construídas por meio de uma projeção entre as duas escalas.

¹⁸ They describe essential skills and knowledge students need to become ready for college and career.

III. O *General Test (GRE)* é realizado pela ETS e é um teste normalmente exigido para a admissão em programas de pós-graduação nos EUA ou em cursos de graduação considerados de 2º ciclo. Possui duas versões: impressa e computadorizada. O GRE mede raciocínio verbal, raciocínio quantitativo, pensamento crítico e competências em escrita. É organizado em três baterias: *Verbal Reasoning*, *Quantitative Reasoning* e *Analytical Writing*. Em geral, o GRE é utilizado pelas universidades de forma complementar às informações da graduação. O teste é pontuado da seguinte forma: as duas primeiras baterias são divididas em duas seções, com 20 itens cada, e os pontos brutos convertidos numa escala de 130-170, com incrementos de 1 ponto, e a última seção em uma escala de 0-6 pontos, com incremento de 0,5 pontos. Para as duas primeiras, os escores brutos (*raw scores*) são computados diretamente das respostas corretas e convertidos para a escala específica do GRE por meio de um processo de equalização (*equating*).

Publicações de autores da ETS afirmam que até a década de 1980 se utilizava o método de Tucker (*Tucker Equating Procedure*), ver, por exemplo, Mckinley e Kingston (1987). A partir da década de 1990, provavelmente com a introdução da versão computadorizada tradicional em *Computer Adaptive Test (CAT)*, passaram-se a utilizar os métodos da TRI para a pontuação e a equalização dos escores (MISLEV et al, 2006). Com a mudança introduzida em 2011, com a substituição do CAT tradicional por uma versão *Multistage Adaptive Test (MST)*, não só a escala como o método de equalização mudou. Pelo que é apresentado no site da ETS, e como primeiro são calculados os pontos brutos e depois convertidos para escala do GRE, é muito provável que, a exemplo dos demais testes de admissão realizados pela ETS e outros nos EUA, se utilize um método de equalização baseado num procedimento do tipo *Item Response Theory (IRT) True Score Equating* (KOLEN; BRENNAN, 2004).

Os procedimentos de equalização acima utilizam a Teoria da Resposta ao Item para calcular os parâmetros dos itens e as proficiências dos examinandos, conforme uma determinada escala de referência. No entanto, os escores calculados pela TRI não são utilizados diretamente na pontuação dos alunos. Dada uma forma do teste, eles são usados, juntamente com os parâmetros dos itens presentes na forma, para estabelecer uma relação funcional entre as proficiências calculadas pela TRI e o chamado *true score* correspondente àquela forma. Como essa mesma relação já é conhecida para o caso de uma forma de referência ou de origem da escala, constrói-se uma tabela associando os *true scores* das duas formas (a forma que se quer equalizar e a forma de referência). Por meio dessas tabelas, os escores brutos dos examinandos

que responderam à forma presente do teste são transformados para a escala da forma de referência.

Assim, embora se utilize a TRI, isto é feito de forma indireta no processo de escalonamento. Por outro lado, esse procedimento garante que o ordenamento dos alunos em uma determinada forma, segundo os pontos brutos, não se altere em relação à pontuação convertida. A razão pela qual a maior parte dos testes de admissões nos EUA utiliza esse método de equalização não está absolutamente clara. Há estudos que procuram mostrar que a precisão dessas medidas é muito próxima da precisão de medidas obtidas diretamente pela TRI para diferentes tipos de modelos. Além disso, dada a tabela de conversão, é possível ao examinando conferir sua pontuação em função do número de acertos no teste e, assim, recorrer, se for o caso de eventuais discrepâncias. Pode-se também supor que esse tipo de procedimento atenda à legislação sobre testagem para admissão no ensino superior presente em vários estados daquele país.

Para a seção de escrita analítica, cada ensaio recebe um escore 0 a 6 com base na tarefa especificada no teste por um corretor humano. O ensaio é também submetido a um programa corretor. Se os escores coincidem aproximadamente, então a média de ambos é tomada como a média final, caso contrário, um segundo corretor humano avalia o ensaio e o escore final é o escore médio dos dois corretores humanos (EDUCATIONAL TESTING SERVICE, c2017).

Existem ainda os GRE *subject tests* que são usados para medir o conhecimento em áreas específicas: Biologia, Química, Literatura em Inglês, Matemática, Física e Psicologia.

Alguns testes de aptidão e conhecimentos específicos

Entre os testes de aptidões e conhecimentos específicos, destacam-se:

- I. O *Test of English as a Foreign Language (TOEFL)* é a avaliação em língua inglesa mais usada em todo o mundo com finalidade admissional de alunos estrangeiros e é aceito por Universidades e organizações em 130 países diferentes, incluindo EUA, Reino Unido, Austrália, Nova Zelândia (Educational Testing Service, 2011). A versão mais recente do TOEFL iBT foi lançada em 2005. O TOEFL é de responsabilidade do importante centro de testagem ETS e é realizado desde 1965. O teste é organizado em quatro seções que são pontuadas em uma escala, conforme descrição na tabela a seguir:

Tabela 1 – Escala TOEFL

Score	Scale	Reliability Estimate	SEM
Reading	0-30	0.85	3.35
Listening	0-30	0.85	3.20
Speaking	0-30	0.88	1.62
Writing	0-30	0.74	2.76
Total	0-120	0.94	5.64

Fonte: *Educational Testing Service*, 2011.

As escalas para as medidas dos testes do TOEFL iBT foram estabelecidas para um intervalo de 0 a 30 para as 4 seções, conforme descrito na Tabela 1, e todas têm o mesmo peso na medida total da habilidade em Inglês. O escore total é simplesmente a soma de todos os escores obtidos nas diferentes seções.

Assim, os escores brutos das seções são computados e, posteriormente, são convertidos também por meio do método do escore verdadeiro, *IRT True Score Equating* (Kolen; Brennan, 2004), para estabelecer a relação entre os escores brutos do último teste com os escores brutos de uma forma de referência e, como existe uma equivalência dos escores brutos da forma de referência com os níveis da escala, é possível fornecer uma equivalência entre os escores brutos da forma atual com os pontos de 0 a 30 da escala original, aproximadamente.

Especificamente, Way e Reese (1990) afirmam que até então o modelo da TRI utilizado seria o Modelo Logístico de 3 Parâmetros.

II. O **LSAT** é, segundo o site da organização *Law School Admission Council* (LSAC), responsável pelo teste requerido na maioria das escolas de direito dos EUA (Law School Admission Council, c2017a). Ainda segundo a organização, atualmente, 221 escolas de direito (*law schools*) nos EUA, Canadá e Austrália são membros do conselho e utilizam o LSAT. O teste consiste de cinco seções de 35 minutos com questões de múltipla escolha. Quatro dessas seções contribuem para a pontuação total no teste: uma que avalia a compreensão em leitura, uma que avalia o raciocínio analítico e duas que avaliam o raciocínio lógico. A seção, cujos resultados não são usados na pontuação, é normalmente usada para testar novos itens ou pré-equalizar novas formas do teste. Uma seção de 35 minutos que avalia a escrita é aplicada ao final do teste, e uma cópia do texto é enviada diretamente à Universidade. Segundo os organizadores, o LSAT é projetado para medir as competências que são consideradas essenciais na escola de direito: a leitura e compreensão de textos complexos com acurácia e *insights*; a

organização e administração das informações e a habilidade de construir inferências razoáveis a partir delas; a habilidade de pensar criticamente e a análise e avaliação do raciocínio e dos argumentos dos outros. Os escores de cada seção são pontuados em uma escala de 120-180 pontos, no entanto, além do último escore, são apresentados também os últimos resultados de todos os testes realizados pelo examinando em diferentes momentos (até o limite de 12), a partir de uma determinada data (atualmente, junho de 2012), incluindo cancelamentos e não comparecimentos. Além disso, é apresentado um intervalo de escores denominado *score band*. Segundo o LSAT, o *score band* reflete a precisão da medida do LSAT e é apresentado para enfatizar que o escore LSAT é uma estimativa para o verdadeiro e desconhecido nível de proficiência. Trata-se de um intervalo de confiança construído dentro de um nível de confiança de 68%, admitindo-se normalidade para os escores. Como o erro padrão do *score* também é fornecido, intervalos com níveis de confiança maiores podem ser construídos com base na hipótese de normalidade do estimador de proficiências. Também é apresentado o percentil do escore alcançado, calculado a partir dos escores dos candidatos que fizeram o teste nos três anos anteriores.

Na sua versão em papel e lápis, o LSAT equaliza os escores da seguinte forma: os pontos brutos são transformados na escala LSAT usando IRT e *true-scoring equating* (cf. Van der Linden; Pashley, 2013). Segundo Weissman (2011), em artigo patrocinado pelo LSAC, o modelo usado na equalização é o modelo logístico de 3 parâmetros.

III. O *Medical College Admission Test (MCAT)* é administrado e desenvolvido pela *Association of American Medical Colleges (AAMC)* e utilizado para ingresso nos cursos de medicina. Seu processo de pontuação também produz uma transformação dos escores brutos para uma escala por meio de um processo de equalização. Os métodos de pontuação e equalização são muito parecidos com os empregados no LSAT. O teste é organizado em quatro seções: *Biological and Biochemical Foundations of Living Systems*; *Chemical and Physical Foundations of Biological Systems*; *Psychological, Social, and Biological Foundations of Behavior*; e *Critical Analysis and Reasoning Skills*. Assim, o número de respostas corretas em cada seção é transformado para uma escala entre 118 e 132 pontos. Assim como no LSAT, os percentis dos escores são apresentados sendo atualizados anualmente em função do desempenho dos candidatos nas últimas avaliações. O Anexo II contém um exemplo da apresentação do resultado de

um indivíduo. É apresentado o escore estimado, o intervalo de confiança de 68% do escore verdadeiro e o percentil correspondente na distribuição de proficiências nas últimas avaliações. Um detalhe importante é que os escores são apresentados sem casas decimais.

Grandes exames nacionais

Por sua semelhança com o ENEM, em forma e contexto, apresentam-se dois exemplos de grandes provas nacionais.

- I. La Prueba Saber, na Colômbia, é realizada pelo Ministério da Educação (MINEDUCACIÓN), se iniciou em 1968 e tem sido reformulada desde então. Atualmente, se presta aos seguintes objetivos: a) selecionar os alunos para a educação superior; b) monitorar a qualidade da formação dos estabelecimentos que oferecem educação de nível médio; e c) produzir medidas para a estimação do valor agregado da educação superior. Atualmente, o exame de Estado Saber 11º se compõe de cinco provas: Matemática, Leitura Crítica, Sociais e cidadania, Ciências Naturais e inglês. Além dos escores nestas competências, se produz medidas de duas subcompetências: cidadania e raciocínio quantitativo. Os escores são computados por meio do uso de modelos de Rasch, da TRI. Cada prova é computada numa escala de 0 a 100 e tem como referência os participantes do teste em 2014, para os quais a média foi estabelecida em 50 pontos e o desvio padrão em 10 pontos da escala. Também são apresentados os percentis em que as pontuações dos alunos se encontram (Instituto Colombiano para la Evaluación de la Educación, 2014).

- II. La Prueba de Selección Universitaria (**PSU**), no Chile, contém testes de Matemática e Linguagem e comunicação, que são obrigatórios, e testes de ciências (biologia, física, química e técnico-profissional) e história, geografia e ciências sociais, que são eletivos. Os alunos escolhem as eletivas de acordo com os requisitos das universidades. A prova é coordenada pelo *Departamento de Evaluación, Medición Y Registro Educacional* (DEMRE) da *Univesidad de Chile*. A prova é pontuada classicamente da seguinte forma: calcula-se a pontuação bruta e esta é transformada (escalamento) para uma escala padrão com média 500 e desvio padrão 110, por meio da padronização de cada escore, tomando como valores, para a padronização, a média e o desvio-padrão dos pontos brutos dos que fizeram a prova naquele ano. Um cálculo adicional não explicitado faz com que os valores sejam fixados entre 150 e 850 pontos. Percentis dos

escores padronizados também são fornecidos às universidades (UNIVERSIDAD DE CHILE, [c2017]a). O processo de equalização aqui empregado pressupõe que os grupos que realizam os testes em cada ano são muito semelhantes segundo sua distribuição de escores de proficiências.

O ACESSO AO ENSINO SUPERIOR NO BRASIL E O ENEM

O sistema de ingresso na educação superior no Brasil é, naturalmente à semelhança de países com igual nível de desenvolvimento econômico, um sistema bastante seletivo. No entanto, até pouco tempo atrás, era bastante descentralizado.

Em 1998, o governo federal do Brasil criou o Exame Nacional do Ensino Médio como um instrumento para avaliar o desempenho dos estudantes no término da educação básica. A finalidade do exame era auxiliar o Ministério da Educação (MEC) na elaboração de políticas pontuais e estruturais de melhoria do ensino brasileiro por meio dos Parâmetros Curriculares Nacionais (PCNs) do ensino médio e fundamental. Durante mais de dez anos, este exame foi usado única e exclusivamente para avaliar as habilidades e competências de concluintes do ensino médio, sem o objetivo de selecionar para o ensino superior. Os exames de seleção para o ensino superior, os concursos vestibulares, eram formulados por equipes locais país a fora e formatos diferentes ocorriam nas diversas universidades e centros de ensino superior.

O ingresso no ensino superior no Brasil, particularmente o ingresso nas escolas públicas, sofreu mudanças substanciais a partir de 2009, quando o ENEM passou por uma reformulação metodológica, com a finalidade de se incentivar a sua utilização como exame de seleção unificada nos processos seletivos, principalmente, das universidades públicas federais. Além disso, o nível de centralização do sistema de ingresso se elevou significativamente com a implantação do Sistema de Seleção Unificada (Sisu), que consiste em uma plataforma online por meio da qual as instituições de ensino superior ofertam vagas em cursos de graduação a estudantes que serão selecionados com base na nota obtida no último Enem.

Assim, em poucos anos, o ensino superior brasileiro transitou de um sistema de seleção bastante descentralizado, em que alunos se aplicavam diretamente e de forma independente para cada instituição, para um sistema razoavelmente centralizado, em que parte expressiva deles se submete a algumas

opções de curso para um regulador central e o critério de seleção é baseado nos resultados de um exame de ingresso comum, ainda que parcialmente, no caso de algumas instituições.

Genericamente, o objetivo da política apresentado pelo Ministério da Educação seria a democratização do acesso ao ensino superior de qualidade, possibilitando igualdade na diversificação de escolha por parte de todos os candidatos e acesso a um rol maior de instituições. Apontava-se, ainda, por conta da redução de diversos custos incorridos pelo aluno com a implantação do novo ENEM e do SISU que uma das consequências dessa mudança do sistema de seleção seria o aumento da mobilidade estudantil interna.

Por outro lado, esperava-se a melhoria dos níveis de habilidades dos alunos ingressantes nos diferentes cursos e uma repercussão positiva no ensino médio em geral, pois traria, em certo sentido, um padrão de qualidade a ser perseguido pelas escolas. Em contrapartida, como consequência adversa, esperava-se um aumento do dispêndio com o exame por parte do governo federal e o aumento da evasão nos primeiros anos do Ensino Superior. Há alguma evidência de que houve aumento nas taxas de mobilidade estudantil, mas também nos níveis de evasão de diversos cursos (ver, LI, 2016). No entanto, até agora, realizaram-se poucos estudos para uma avaliação mais abrangente dos impactos do ENEM.

CONCLUSÃO

A atual prova do ENEM, em parte, copia os testes dos vestibulares tradicionais que o antecederam, no que se refere à concepção do que se deve medir, com foco na realização ou êxito (*achievement*) relacionado a conhecimentos específicos (*subjects*) organizados nas quatro disciplinas, avaliadas por questões objetivas de múltipla escolha. Além disso, contém uma prova de redação mensurada a partir de critérios e procedimento padronizados por avaliadores.

Embora o ENEM original fosse concebido para ser uma prova que mensurava mais habilidades cognitivas do que conhecimentos específicos, com o tempo ele se tornou mais parecido com os testes dos vestibulares tradicionais. Outras características psicométricas também se alteraram ao longo dos anos, como o nível de dificuldade dos itens. Essas modificações, em nossa opinião, ocorrem, pelo menos, por dois fatores importantes. O primeiro diz respeito à concepção e à prática avaliativa dos profissionais de certos cursos

das Universidades, mais acostumados em focar a avaliação no conhecimento do que nas habilidades. Logo, para facilitar a adesão ao uso dos resultados dos testes, nada mais natural que aproximar os testes do ENEM dos testes de vestibular tradicionais. Em segundo, também para atender à necessidade das universidades que utilizam o ENEM, e tendo em vista a acirrada disputa por uma vaga em determinados cursos, especialmente nas universidades públicas, o teste passa a introduzir itens mais difíceis numa tentativa de aumentar a precisão das medidas nos níveis mais elevados de proficiências, o que de fato vem ocorrendo e se reflete nas características dos testes que serão discutidas oportunamente neste e no outro documento a ser elaborado.

No entanto, por ser um grande exame nacional, com uma abrangência territorial imensa, o ENEM não poderia utilizar o sistema de pontuação habitualmente utilizado nos vestibulares tradicionais. Isso ocorre, por exemplo, como já foi inclusive observado em edições de anos anteriores do ENEM, pela ocorrência de determinados eventos não previsíveis, como intempéries e catástrofes, que naturalmente demandam a aplicação de novos testes ao grupo de indivíduos afetados pelos eventos, quer seja porque não tenham podido realizar as provas como os demais, quer seja porque tivessem suas provas anuladas por algum motivo. Naturalmente, pelos custos envolvidos, a anulação de todo o ENEM é, senão infactível, pelo menos irracional. Portanto, o ENEM, como grande prova nacional que é, necessita naturalmente de um sistema de pontuação que permita a comparabilidade das medidas produzidas em testes diferentes, do contrário se tornaria inviável a sua realização.

A experiência brasileira em mensuração e escalonamentos de testes admissionais era pouco expressiva até o novo ENEM. Por outro lado, havia considerável experiência na mensuração de testes de avaliação educacional em larga escala, sendo o exemplo mais paradigmático, a mensuração dos resultados do SAEB e da Prova Brasil, pelo menos desde meados da década de 90.

Em contrapartida, embora exista legislação sobre os concursos públicos de admissão, que de certa forma abrange os concursos vestibulares, não há especificidade da lei no que tange ao processo de mensuração em si. Por outro lado, o escopo da lei existente sobre concursos parece não atingir plenamente um teste com a dimensão, necessidades e tecnologia do ENEM.

O sistema de pontuação dos resultados do novo ENEM, a partir de 2009, então deriva dos sistemas de mensuração utilizados nas avaliações em larga escala empregados no Brasil e no mundo para a avaliação em larga escala de sistemas educacionais. A escolha, com pouco debate, recaiu sobre os mesmos métodos de cálculo das medidas e equalização dos resultados empre-

gados no caso da mensuração do SAEB/Prova Brasil e de outras avaliações de sistemas educacionais no país, muito provavelmente porque essa era a *expertise* dos especialistas contratados pelo INEP. Quando se analisam os testes utilizados na admissão ao ensino superior no mundo (ver seção 5,) verifica-se a singularidade dos métodos utilizados na pontuação do ENEM.

Isso não significa, *a priori*, que as medidas do ENEM são piores ou melhores do que as empregadas no SAT ou no LSAT, por exemplo. Ambos os sistemas, SAT e LSAT, utilizam métodos diferentes para o escalonamento de seus resultados. Mas ambos utilizam métodos que permitem a transformação de seus escores brutos para uma escala comparável pelo método conhecido como *IRT True Score Equating*, conforme apresentado na seção 5. A tabela de transformação é fornecida aos candidatos de tal forma que, por meio dela, se pode conferir sua pontuação final. Não está explícito, mas parece que essa metodologia tem sido utilizada, em parte, para satisfazer a legislação americana para testes admissionais de alguns estados (Greer, 1984), como a dos estados da Califórnia e Nova York. De qualquer forma, esse expediente permite maior controle do candidato de sua pontuação final.

REFERÊNCIAS

BAKER, F. B.; KIM, S. **Item Response Theory**. 2 ed. Statistics: Textbooks and Monographs, v. 129. New York, USA: Marcel Dekker, Inc., 2004.

BIRNBAUM, A. Some Latent Trait Models and Their Use in Inferring an Examinee's ability. In: LORD, F. M.; NOVICK, M. R. **Statistical Theories of Mental Test Scores**. Reading, MA: Addison- Wesley Pub, 1968. p. 397-472.

COURVILLE, T. G. **An empirical comparison of item response theory and classical test theory item/person statistics**. (Unpublished doctoral dissertation). Texas: A&M University, 2005. Disponível em: <<http://txspace.tamu.edu/bitstream/handle/1969.1/1064/etd-tamu-2004B-EPSY-Courville-2.pdf?sequence=1>>. Acesso em: 05 nov. 2007.

EDUCATIONAL TESTING SERVICE. **2014 ETS Standards for Quality and Fairness**. Princeton, NJ: ETS, 2015.

_____. **How the Test Is Scored**. c2017. Disponível em: <https://www.ets.org/gre/revised_general/scores/how/>. Acesso em: 04 out. 2017.

_____. Reliability and Comparability of TOEFL iBT™ Scores. **TOEFL iBT Research**, Princeton, NJ, Series 1, v. 3, 2011.

EDWARDS, D.; COATES, H.; FRIEDMAN, T. A survey of international practice in university admissions testing. **Higher Education Management and Policy**, OECD, v. 24/1, 2012.

FAN, X. Item response theory and classical test theory: an empirical comparison of their item/person statistics. **Educational and Psychological Measurement**, Santa Barbara, CA, v. 58, n. 3, p. 357-381, June 1998.

GREER, D. G. "**Thuth-in-testing legislation**" An Analysis of political and legal consequences, and prospects. Houston TX: Institute for Higher Education Law and Governance, 1984.

HAMBLETON, R. K.; JONES, R. W. Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. **Educational Measurement: Issues and Practice**, v. 12(3), p. 38-47, 1993.

INSTITUTO COLOMBIANO PARA LA EVALUACIÓN DE LA EDUCACIÓN. Sistema Nacional de Evaluación Estandarizada de la Educación. Alineación del examen SABER 11°. **Lineamientos generales 2014 – 2**. 2014. Disponível em: <<http://www.icfes.gov.co/docman/instituciones-educativas-y-secretarias/saber-11/novedades/650-guia-lineamientos-generales-saber-11-2014-2/file?force-download=1>>. Acesso em: 15 out. 2017.

JABRAYILOV, R.; EMONS, W. H. M.; SIJTSMA, K. Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment. **Applied Psychological Measurement**, v. 40(8), p. 559-572, 2016.

KOLEN M.; BRENNAN R. L.; **Test Equating, scaling, and linking**. New York: Springer-Verlag, 2004.

LAW SCHOOL ADMISSION COUNCIL. Law School Admission Test (LSAT). Score Bands. c2017b. Disponível em: <<https://www.lsac.org/jd/lSAT/your-score/score-band>>. Acesso em: 04 out. 2017.

LI, D. L. **O novo Enem e a plataforma Sisu**: efeitos sobre a migração e a evasão estudantil. 2016. Dissertação (Mestrado em Ciências) – Programa de Pós-Graduação em Economia do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2016.

LIU, J.; HARRIS, D. J.; SCHMIDT, A. Statistical Procedures Used in College Admissions Testing. In: RAO, C. R.; SINHARAY, S. **Handbook of Statistics 26: Psychometrics**. Amsterdam: Elsevier, 2007, p. 1057-91.

LORD, F. M. **Applications of item response theory to practical testing problems**. Hillsdale: Lawrence Erlbaum, New York, 1980.

LORD, F. M.; NOVICK, M. R. **Statistical Theories of Mental Test Scores**. Reading, MA: Addison-Wesley, 1968.

MACDONALD, P.; PAUNONEN, S. V. A Monte Carlo comparison of item and person parameters based on item response theory versus classical test theory. **Educational and Psychological Measurement**, v. 62, p. 921-43, 2002.

MCGRATH, C. H. et al. **Higher Education Entrance Qualifications and Exams in Europe: A Comparison**. Brussels: European Union, 2014.

McKINLEY, R.; KINGSTON, N. **Exploring The Use of IRT Equating for the GRE Subject Test in Mathematics**. ETS Report, RR-87-2, 1987.

MISLEV R. et al. Concepts, Terminology and Basic Models of Evidence-Centered Design. In: WILLIAMSON D.; MISLEV R.; BEJAR I. **Automated Scoring of Complex Tasks in Computer-based Testing**. London: Ed. Lawrence Erlbaum, 2006, p. 15-48.

PALMER, N.; BEXLEY, E.; JAME, R. **Selection and participation in higher education university selection in support of student success and diversity of participation**. Melbourne: CSHE, 2011.

SALLE, C. M. Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data. **The International Journal of Educational and Psychological Assessment**, v. 1, Issue 1, p. 1-11, April 2009.

SARGEANT, C. et al. **INCA comparative tables**. London: International Review of Curriculum and Assessment Frameworks Internet Archive, 2012. Disponível em: <https://www.nfer.ac.uk/what-we-do/information-and-reviews/inca/INCA-comparativetablesMarch2012.pdf>. Acesso em: ?????

THE COLLEGE BOARD. **At a glance**. c2017a. Disponível em: <<https://collegereadiness.collegeboard.org/sat-subject-tests/about/at-a-glance>>. Acesso em: 04 out. 2017.

_____. **Score Choice**. c2017b. Disponível em: <<https://collegereadiness.collegeboard.org/sat-subject-tests/scores/sending-scores/score-choice>>. Acesso em: 04 out. 2017.

_____. **Score Structure**. c2017c. Disponível em: <<https://collegereadiness.collegeboard.org/sat/scores/understanding-scores/structure>>. Acesso em: 04 out. 2017.

_____. **Understanding Scores**. c2017d. Disponível em: <https://collegereadiness.collegeboard.org/sat/scores/understanding-scores/interpreting>. Acesso em: 04 out. 2017.

THISSEN, D.; WAINER, H. **Test Scoring**. Mahwah, New Jersey: Lawrence Erlbaum Associates Pub., 2001.

UNIVERSIDAD DE CHILE. Departamento de evaluación, medición y registro educacional. Prueba de Selección Universitaria. **¿Cómo se calcula el puntaje?** [c2017]a. Disponível em: <http://www.psu.demre.cl/la-prueba/que-es-la-psu/calculo-puntaje-psu>. Acesso em: 20 out. 2017.

UNIVERSIDAD DE CHILE. Departamento de evaluación, medición y registro educacional. Prueba de Selección Universitaria. **El proceso de admisión**. [c2017]b. Disponível em: <http://www.psu.demre.cl/proceso-admision/>. Acesso em: 24 out. 2017.

VAN DER LINDEN, W. J.; PASHLEY, P. J. Item Selection and Ability Estimation in Adaptive Testing. In: VAN DER LINDEN, W. J.; GLAS, C. A. W. (Eds.). **Elements of adaptive testing**. London: Kluwer Academic Pub., 2013. p. 2-26.

WAY, C.; REESE, C. An Investigation of the Use of Simplified IRT Models for Scaling and Equating the TOEFL Test. **ETS Research Report**, RR-90-29, 1990.

WEISSMAN, A. Optimizing Information Using the Expectation-Maximization Algorithm in Item Response Theory. **Law School Admission Council Research Report**, 11-01, March 2011.

Recebido em: 10 de dezembro de 2019

Aceito em: 19 de junho de 2020

Publicado em: 30 de junho de 2020

Pareceristas (Avaliadores ad hoc) 2020

Alexandre Chibebe Nicolella, Universidade de São Paulo, São Paulo, São Paulo, Brasil

Alexandre José Cadilhe, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Ana Rosa Picanço Moreira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Antonio Ferreira Colchete Filho, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Beatriz de Basto Teixeira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Clécio da Silva Ferreira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Daniel Eveling da Silva, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Denise Rangel Miranda, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Ilka Schapper Santos, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Isabela de Mattos Ferreira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Isabela Monken Velloso, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Josiane Toledo Ferreira Silva, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Juliane Fonseca Figueiredo, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Leonardo Ostwald Vilardi, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Luciana da Silva de Oliveira, Instituto Federal de Minas Gerais, Bambuí, Minas Gerais, Brasil

Márcio Fagundes Alves. Instituto Metodista Granbery, Juiz de Fora, Minas Gerais, Brasil

Patrícia Maria Reis Cestaro, Prefeitura Municipal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Rafaela Reis Azevedo de Oliveira, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Roberto Perobelli de Oliveira, Universidade Federal do Espírito Santo, Vitória, Espírito Santo, Brasil

Vanda Mendes Ribeiro, Universidade da Cidade de São Paulo, São Paulo, São Paulo, Brasil

Vítor Fonseca Figueiredo, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

Víviam Carvalho de Araujo, Prefeitura Municipal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

■ PESQUISA APLICADA

AVALIAÇÃO DA QUALIDADE NA EDUCAÇÃO INFANTIL: IMPASSES E PERSPECTIVAS NO BRASIL

Maria Malta Campos

AVALIAÇÃO DA APRENDIZAGEM NA EDUCAÇÃO INFANTIL: ESTUDO DE CASO EM UM MUNICÍPIO BRASILEIRO

Bruna Ribeiro

EDUCAÇÃO INFANTIL E AVALIAÇÃO: PROCESSOS E PERSPECTIVAS

Rita de Cássia de Freitas Coelho

Alice de Paiva Macário

A RESTITUIÇÃO COMO PROCESSO CENTRAL NA AVALIAÇÃO DE CONTEXTO

Catarina Moro

AVALIAÇÕES DA EDUCAÇÃO INFANTIL EM MUNICÍPIOS PAULISTAS: LIMITES E POTENCIALIDADES PARA CONTRIBUIR COM A GARANTIA DO DIREITO À EDUCAÇÃO DAS CRIANÇAS PEQUENAS

Cláudia Oliveira Pimenta

AVALIAÇÃO DE CONTEXTO NA EDUCAÇÃO INFANTIL: UMA ESCOLHA PARA PENSAR A MELHORIA DAS PRÁTICAS EDUCATIVAS

Patrícia Corsino

Jordanna Castelo Branco

INFRAESTRUTURA ESCOLAR E SATISFAÇÃO PROFISSIONAL: PERCEPÇÃO DE PROFESSORES DA EDUCAÇÃO BÁSICA BRASILEIRA

Lívia Maria Fraga Vieira

Edmilson Antônio Pereira Junior

OBRIGATORIEDADE DA PRÉ-ESCOLA EM PERSPECTIVA: AVALIAÇÃO DO ATENDIMENTO NO ESTADO DO RIO DE JANEIRO

Bruno Tovar Falciano

Maria Fernanda Rezende Nunes

AVALIAÇÕES EM LARGA ESCALA DE LÍNGUA PORTUGUESA: UMA PESQUISA SOBRE A COMPLEXIDADE DOS TEXTOS QUE DÃO SUPORTE A ITENS QUE AVALIAM A LEITURA

Begma Tavares Barbosa

Hilda Linhares Micarello

Rosângela Veiga Ferreira

A AVALIAÇÃO EXTERNA E OS NOVOS SUJEITOS DA EDUCAÇÃO

Marcelo Baumann Burgos

AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA E ACCOUNTABILITY: UMA BREVE ANÁLISE DA EXPERIÊNCIA BRASILEIRA

Reynaldo Fernandes

Amaury Patrick Gremaud

AVALIAÇÃO EDUCACIONAL: FATORES CONTEXTUAIS DE EFICÁCIA ESCOLAR EM CENÁRIOS DE ALTA VULNERABILIDADE SOCIAL

Edivaldo Cesar Camarotti Martins

Adolfo-Ignacio Calderón

ESCALAS ESPECÍFICAS PARA MEDIR A INFRAESTRUTURA ESCOLAR DA EDUCAÇÃO BÁSICA

Joaquim José Soares Neto

Elianice Silva Castro

MENSURAÇÃO EM SISTEMAS ADMISSIONAIS NO ENSINO SUPERIOR E AVALIAÇÕES DE IMPACTO

Tufi Machado Soares

Mariana Calife Nóbrega Soares