

IMPLICAÇÕES DAS AVALIAÇÕES EM LARGA ESCALA NAS ESCOLAS DE EDUCAÇÃO BÁSICA: UMA REVISÃO DE LITERATURA

IMPLICATIONS OF LARGE-SCALE ASSESSMENTS IN BASIC EDUCATION SCHOOLS: A LITERATURE REVIEW

IMPLICACIONES DE LAS EVALUACIONES A GRAN ESCALA EN LAS ESCUELAS DE EDUCACIÓN BÁSICA: UNA REVISIÓN DE LA LITERATURA

CRISTIANE MACHADO*

ALINE GABRIELE PEREIRA**

Resumo: Este artigo apresenta resultados de estudo de revisão bibliográfica que teve o objetivo de analisar as implicações das avaliações em larga escala nas escolas de Educação Básica. Foram selecionados 12 artigos em um levantamento realizado no Portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes, onde foram localizados 118 artigos. Concluiu-se que evidências da centralidade da avaliação em larga escala no cotidiano escolar podem ser notadas, dentre outras, por iniciativas pedagógicas de treino dos alunos para as provas, pela intensificação do trabalho docente e pela imposição às escolas de responsabilização dos seus resultados. Além disso, há a ocultação das desigualdades educacionais produzidas pelos contrastes econômicos e sociais próprios da histórica ocupação territorial desordenada dos municípios brasileiros e pelos processos de exclusão de alunos de inclusão.

Palavras-chave: Avaliação em larga escala; Escolas de Educação Básica; Revisão bibliográfica.

Abstract: This paper presents results of a bibliographic review study that aimed to analyze the implications of large-scale assessments in Basic Education schools. Twelve papers were selected in a survey carried out on the Journal Portal of the Coordination for the Improvement of Higher Education Personnel - CAPES, where 118 papers were found. It was concluded that evidence of the centrality of large-scale assessments in the school routine can be noticed, among others, by pedagogical initiatives to train students for the tests, by intensifying the teaching work and imposing on the schools responsibility for their results. In addition, there is the concealment of educational inequalities produced by the economic and social contrasts inherent in the historical disorderly territorial occupation of Brazilian municipalities and by the processes of exclusion of inclusion students.

Keywords: Large-scale assessment; Basic Education schools; Literature review.

Resumen: Este artículo presenta resultados de estudio de revisión bibliográfica que tuvo como objetivo analizar las implicaciones de las evaluaciones a gran escala en las escuelas de Educación Básica. Fueron seleccionados 12 artículos en una recopilación realizada en el Portal de periódicos de la Coordinación de Perfeccionamiento de Personal de Nivel Superior - Capes, donde se encontraron 118 artículos. Se concluyó que evidencias de la centralidad de la evaluación a gran escala en el cotidiano escolar pueden ser observadas, entre otras, por iniciativas pedagógicas de entrenamiento de los alumnos para las pruebas, por la intensificación del trabajo docente e por la imposición a las escuelas de responsabilización de sus resultados. Además, hay ocultación de las desigualdades

* Doutora em Educação pela Universidade de São Paulo (USP). Docente no Departamento de Políticas, Administração e Sistemas Educacionais (DEPASE) na Faculdade de Educação da Universidade Estadual de Campinas (UNICAMP). E-mail: crimacha@unicamp.br.

** Doutoranda na Faculdade de Educação da Universidade Estadual de Campinas (UNICAMP) e Docente na Universidade do Estado de Minas Gerais (UEMG). E-mail: alinesisb@gmail.com.

educativas produzidas por los contrastes económicos y sociales inherentes a la ocupación territorial histórica desordenada de los municipios brasileños e por los procesos de exclusión de alumnos de inclusión.

Palabras clave: Evaluación a gran escala; Escuelas de Educación Básica; Revisión bibliográfica.

INTRODUÇÃO

A implementação progressiva e difusa de políticas de avaliação em larga escala, no Brasil, tem estimulado a ampliação de pesquisas que buscam conhecer, apreender, interpretar e analisar as consequências desse movimento nas diversas dimensões que ele reverbera, especialmente nas instituições educacionais. As investigações, em geral, debruçam-se na compreensão dos fundamentos e dos desenhos dessas avaliações, seus objetivos, seus sentidos políticos e sua capacidade de impactar positivamente na qualidade do ensino, dentre outros. Observa-se, mais recentemente, um deslocamento das preocupações das pesquisas que passaram a adotar como foco privilegiado as implicações dessas avaliações para e nas escolas de Educação Básica.

Machado (2012) aponta dois importantes momentos de inflexão para esse deslocamento que aproximaram as pesquisas das escolas, notadamente do interior dessas instituições. O primeiro, em 2005, com a criação da Prova Brasil, que começou a aferir censitariamente o desempenho dos/as estudantes, o que até então era feito de forma amostral. O segundo, em 2007, como consequência do primeiro, que trata da instauração, pelo Governo Federal, de um indicador da qualidade da educação no país, o Índice de Desenvolvimento da Educação Básica (Ideb). Tanto o primeiro evento como o segundo ampliaram a visibilidade sobre o trabalho escolar e, principalmente, sobre o que passou a ser compreendido como resultado desse trabalho, o Ideb, uma vez que esse índice é representado por meio de uma nota obtida a partir da proficiência dos/as estudantes na Prova Brasil.

Esse movimento de proximidade das avaliações em larga escala com o cotidiano das escolas tem evidenciado a necessidade de estudos que investiguem o potencial dessas políticas contribuírem (ou não!) com avanços na gestão e na organização do trabalho escolar e no ensino oferecido. Ao examinarem pesquisas nessa temática, Bauer, Alavarse e Oliveira (2015, p. 1377) repercutem constatações de um grupo de autores que, por meio de suas pesquisas, ponderam que “[...] algumas implicações das avaliações sobre as redes e escolas não são, em essência, boas ou ruins, mas que podem ter influências diferenciadas, a depender de cada contexto escolar”.

De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2020), o objetivo da avaliação em larga escala federal, principal referência para a proposição, elaboração e formatação das avaliações desse tipo em níveis estaduais e municipais (MACHADO, 2020), é produzir informações educacionais que permitam, por um lado, a realização de “[...] um diagnóstico da educação básica brasileira e de fatores que podem interferir no desempenho do estudante” (INEP, 2020, n.p.) para os órgãos governamentais acompanharem o desenvolvimento da educação nacional. Por outro lado, possibilitar a avaliação “[...] da qualidade da educação oferecida aos estudantes” (INEP, 2020, n.p.) e o fornecimento de “[...] subsídios para a elaboração, o monitoramento e o aprimoramento de políticas educacionais com base em evidências” (INEP, 2020, n.p.) pelas escolas e pelas redes municipais e estaduais de ensino de todo o país.

Nesse contexto, cabe indagar: Quais têm sido as implicações das avaliações em larga escala no cotidiano das escolas de Educação Básica no Brasil? Ao buscar responder a essa questão e compreender as interfaces das avaliações em larga escala com as escolas, este artigo tem o objetivo de examinar as implicações dessas avaliações no cotidiano das instituições de Educação Básica. Trata-se de uma pesquisa de revisão bibliográfica que consultou textos depositados no Portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

O levantamento¹ foi realizado na última quinzena do mês de abril de 2020, com a utilização do descritor “avaliação em larga escala” e a aplicação dos filtros “artigos” e “education”. O resultado exibiu

¹ Realizado por Henrique de Souza e Silva, Héri Anai Wavrita e Lo-ruama Kemily Felipe, Bolsistas do Serviço de Apoio ao Estudante (SAE) da Unicamp.

118 manuscritos. Após leitura criteriosa dos resumos de todos os textos, foi detectado que 56 analisavam as avaliações em larga escala em interface com implicações na escola. Destes, foram descartados três por se referirem ao Ensino Superior (2) e ao contexto internacional (1), restando 53 artigos.

Ao assumir-se o fundamento metodológico que a partir de “[...] um contato direto do pesquisador com a situação pesquisada, torna-se possível reconstruir os processos e as relações que configuram a experiência escolar diária” (ANDRÉ, 2008, p. 141), privilegiou-se o exame de artigos que se debruçaram em investigar empiricamente as implicações das avaliações em larga escala no cotidiano das escolas. São, desse modo, textos que evidenciam resultados de pesquisas que recorreram ao campo como forma de compreensão de seus objetivos. Assim, dentre os 53 textos da temática desta pesquisa, implicações das avaliações em larga escala nas escolas de Educação Básica, foram separados 41 manuscritos que se caracterizam como ensaio por abordarem a temática na perspectiva teórica, restando 12 que, além do debate conceitual, trazem à luz elementos de empiria, de pesquisa de campo, se constituindo como amostra deste estudo.

Antes da análise dos 12 artigos, serão pontuados, a seguir, de forma breve, marcos da trajetória de expansão da avaliação em larga escala no Brasil como forma de contribuir com a contextualização do debate encetado nos textos examinados.

APONTAMENTOS DA EXPANSÃO DA AVALIAÇÃO EM LARGA ESCALA NO BRASIL

O amplo crescimento de políticas de avaliação em larga escala dos sistemas educativos, exaustivamente demonstrado e debatido no contexto das reformas educativas como tendência mundial (BAUER; ALAVARSE; OLIVEIRA, 2015), vincula-se ao movimento de transformações no papel do Estado que teve início na Europa e nos Estados Unidos com a crise da década de 1970. O processo foi marcado pelas críticas e pelo questionamento das políticas sociais, pela disseminação da concepção de que os investimentos realizados em tais políticas eram demasiadamente altos em relação aos resultados alcançados, ou seja, à qualidade dos serviços oferecidos. Nessa lógica, a implantação de sistemas de avaliação surge como forma de mensurar a qualidade resultante dos recursos financeiros investidos, convertendo-se em mecanismo de controle e justificando a decisão, a escolha e o direcionamento das verbas orçamentárias.

No Brasil, a reforma do Estado teve início na década de 1990, mais precisamente no governo de Fernando Henrique Cardoso, o qual elegeu a qualidade como eixo das políticas educacionais. No período, marcado por forte expansão das políticas de avaliação nos diversos níveis da educação nacional, foram implementados o Exame Nacional de Cursos (ENC), que visava avaliar o Ensino Superior; o Exame Nacional do Ensino Médio (Enem), destinado a avaliar estudantes do Ensino Médio; e o Sistema de Avaliação da Educação Básica (Saeb), existente desde 1990, foi reformulado metodologicamente e ampliado.

Nova onda de expansão pode ser observada em 2005, mas agora com foco no contingente avaliado. O Saeb foi reformulado e dividido em duas avaliações: a Avaliação Nacional da Educação Básica (Aneb), que manteve o formato de aferição amostral; e a Avaliação Nacional do Rendimento Escolar (Anresc), conhecida como Prova Brasil, que passou a avaliar estudantes censitariamente. Essa reorganização, de acordo com Bonamino e Sousa (2012, p. 379), significou avaliar mais de três milhões de alunos distribuídos em aproximadamente 45.000 escolas urbanas de 5.398 municípios, indo muito além dos 300.000 alunos avaliados pelo Saeb.

O renovado formato do Saeb possibilitou a criação, em 2007, do Ideb, índice que articula o resultado da proficiência dos/as estudantes em língua portuguesa e matemática, obtido com a Prova Brasil, com taxas de aprovação informadas no Censo Escolar. O Ideb, divulgado publicamente a cada dois anos (mesma frequência de realização da Prova Brasil), apresenta resultados para o país, estados, municípios e, também, por escola, em uma escala de 0 a 10, apreendido socialmente como a “nota” da qualidade da educação de redes e escolas.

Esse contexto de criação da Prova Brasil e a instauração do Ideb impulsionaram outra onda de expansão das avaliações em larga escala no país, agora de forma capilar pelos entes federados. Embora

alguns estados tenham adotado avaliações em larga escala próprias antes de 2005, como Ceará (1992), Minas Gerais (1992), Paraná (1995), São Paulo (1996), Alagoas (2001), Mato Grosso do Sul (2003), Espírito Santo (2004) e Sergipe (2004), foi após esse período que houve uma disseminação generalizada de estados que desenvolveram sistemas de avaliação: Bahia (2007), Rio de Janeiro (2008), Amazonas (2008), Pernambuco (2008), Acre (2009), Piauí (2011), Distrito Federal (2011), Goiás (2011), Rio Grande do Sul (2011), Tocantins (2011), Rondônia (2012), Paraíba (2012), Pará (2013) (MACHADO; ALAVARSE; ARCAS, 2015). Também entre os municípios, o planejamento e a implementação de tais políticas têm se expandido e já atingem pelo menos 2.478 municípios nacionais (BAUER *et al.*, 2015).

Essa breve contextualização da trajetória das avaliações em larga escala no Brasil permite constatar duas dimensões na expansão das políticas avaliativas. A primeira diz respeito à ampliação quantitativa de sistemas de ensino que praticam essas políticas como tentativa de mensurar a qualidade do ensino ofertado, tanto em relação aos níveis da educação nacional como também em relação à adoção crescente pelos diferentes entes federados do país. A segunda dimensão, que pode ser compreendida como qualitativa, refere-se ao processo de aprofundamento das implicações das avaliações sobre a instituição escolar. Inicialmente, foi de caráter amostral e diagnóstico sem fortes impactos nas escolas; e, posteriormente, com aferição censitária e divulgação de *rankings* de redes e de escolas, estimulando a comparação e a competição, além da responsabilização pelos resultados alcançados e pressão por melhores índices (BONAMINO; SOUSA, 2012).

Protagonistas nesse movimento de expansão nas últimas três décadas, as escolas passaram a participar de um número cada vez maior de avaliações e a perceber um aprofundamento das implicações, diretas e indiretas, dos resultados dessas avaliações sobre o cotidiano escolar. É nesse contexto que se insere este estudo com a questão inicial: Quais têm sido as implicações das avaliações em larga escala no cotidiano das escolas de Educação Básica no Brasil?

Cabe ressaltar que não é objetivo deste artigo esgotar o assunto ou abordar a totalidade dos efeitos dos processos avaliativos nas escolas, mas, sim, examinar, a partir de pesquisas empíricas transformadas em textos e reunidas no Portal de periódicos da Capes, eventuais implicações das avaliações em larga escala que já puderam ser percebidas no cotidiano escolar.

AVALIAÇÃO EM LARGA ESCALA E IMPLICAÇÕES NAS ESCOLAS

Neste estudo de revisão bibliográfica, foram investigados 12 artigos localizados no Portal de periódicos da Capes em um levantamento que encontrou 118 resultados. Optou-se por textos que evidenciassem resultados de pesquisas que investigaram empiricamente o tema das avaliações em larga escala e suas implicações nas escolas. Essa escolha partiu do princípio de que o contato direto do/a pesquisador/a com a escola pesquisada possibilita uma apreensão *in loco* do cotidiano escolar e das eventuais implicações das avaliações em larga escala no desenvolvimento e na proposição das atividades educativas da instituição.

Entretanto, vale destacar que artigos com outras características, como ensaios e conclusões de pesquisas que se utilizaram de dados secundários, coletados em geral pela própria aferição das avaliações em larga escala, também têm potencial singular de contribuir significativamente com o entendimento dos processos e dos efeitos das avaliações em larga escala nas escolas e nos sistemas educacionais. Por exemplo, Fontanive (2013), ao analisar os possíveis impactos positivos que os resultados das avaliações em larga escala podem ter na prática docente na perspectiva de promover ganhos na aprendizagem dos/as estudantes, salienta consideráveis aspectos que merecem reflexão das equipes dirigentes da educação e das equipes pedagógicas das escolas.

A autora apresenta algumas hipóteses para a não mudança na relação professor-aluno-aprendizagem, apesar da enorme quantidade de dados coletados e divulgados há mais de 20 anos pelas avaliações em larga escala. Ela também cita iniciativas realizadas por órgãos públicos para colaborar na melhoria na aprendizagem, como, por exemplo, a criação de materiais estruturados para alunos e professores, incluindo questões de avaliação e exercícios para os alunos. Fontanive (2013) conclui, dentre outras afirmações, que é necessário investimentos na formação e na capacitação de professores e que o acompanhamento do desempenho dos alunos deve ser realizado frequentemente pelos docentes.

Já o artigo de Silva e Meletti (2014), ao apresentar estudo que utilizou dados secundários do Inep para conhecer a relação entre a matrícula de estudantes com necessidades educacionais especiais na rede regular de ensino do município de Londrina (PR) e a participação na Prova Brasil, traz à tona o relevante debate sobre eventuais encaminhamentos pedagógicos destinados aos estudantes de inclusão das escolas públicas. Os resultados da pesquisa indicam que a participação dos alunos com necessidades educacionais especiais na Prova Brasil foi mínima, o que, para a autora, reforça a ideia de ambiguidade entre discurso e prática inclusiva, desvelando aspectos que podem representar práticas excludentes, apesar da existência de uma retórica legal de inclusão.

Também em Vieira, Vidal e Nogueira (2015) encontram-se sistematizações relevantes de estudo que teve o objetivo de compreender o posicionamento de professores em relação às dificuldades de aprendizagem dos alunos, com base no estudo de respostas do questionário do professor de Ensino Fundamental, 150.972 docentes de 5º ano e 153.440 de 9º ano, na Prova Brasil de 2011. Mesmo com os limites detectados na amostra utilizada, como a alta porcentagem de respostas em branco, 40% em relação aos professores de 5º ano e 22,3% no caso dos professores de 9º ano, os achados são oportunos para entender as condições de trabalho docente.

Após o exame dos dados, as autoras demonstram que há diferenças na percepção dos docentes de 5º e 9º anos do Ensino Fundamental sobre fatores que podem ser associados aos resultados de aprendizagem dos alunos. Elas salientam que, ao que parece, os professores do 9º ano são os que menos identificam relação entre seu trabalho e as dificuldades de aprendizagem de seus alunos. Para 40,7% do total de professores de 9º ano que responderam ao questionário da Prova Brasil em 2011, os alunos são “[...] os grandes responsáveis pelos problemas de aprendizagem com destaque para falta de aptidão (46,6%), baixa autoestima (69,7%), desinteresse e falta de esforço (94,3%) e indisciplina (73,1%)” (VIEIRA; VIDAL; NOGUEIRA, 2015, p. 100).

Ao problematizarem essa constatação, Vieira, Vidal e Nogueira (2015) questionam se é possível concluir que os docentes não se reconhecem como agentes da gestão de aprendizagem de seus alunos e aventam ser afirmativa a hipótese. Nesse sentido, asseveram que “[...] não reconhecendo seu próprio protagonismo, os docentes renunciam a um papel que deles é esperado, pelo menos como mediadores do processo de ensino-aprendizagem [...]” (VIEIRA; VIDAL; NOGUEIRA 2015, p. 101) e advertem que essa “[...] é uma questão a ser melhor aprofundada pela política e pela gestão educacional” (VIEIRA; VIDAL; NOGUEIRA 2015, p. 101).

Neste estudo, os 12 artigos, selecionados após criterioso procedimento de triagem, foram lidos na íntegra e, a seguir, serão apresentados e analisados de acordo com o ano de publicação de cada texto, em ordem crescente. Dentre os manuscritos selecionados, dois são de autoria de Ivan Amaro (2013, 2016). O primeiro deles, publicado em 2013, discute pesquisa realizada por meio de entrevistas com um gestor da Secretaria Municipal de Educação (SME) de um município na região metropolitana do Rio de Janeiro, gestores escolares, orientadoras pedagógicas e professores de três escolas municipais que tiveram aumento significativo no Ideb entre as edições de 2005 e 2011.

O estudo apontou dificuldade de compreensão/apropriação dos dados produzidos pela Prova Brasil, não só na escola, onde os resultados do Ideb parecem estar vinculados apenas ao desempenho dos alunos nos exames, cujos gestores, orientadores e professores não demonstram conhecimento acerca das matrizes de referência, descritores e escalas de proficiência, mas também na SME, onde o corpo técnico apresentou dificuldades em compreender e interpretar os resultados. Outro fator que se mostrou comum entre todos os entrevistados foi o contexto de responsabilização e o clima de pressão pela melhoria dos resultados. Segundo Amaro (2013, p. 34), gestores municipais “[...] precisam dar respostas ao MEC, gestores escolares que necessitam de dar respostas à Secretaria e, por fim, professores que precisam de dar respostas ao gestor escolar”.

Entre as implicações apresentadas pela pesquisa, além da pressão sentida pelos profissionais da educação, foi percebido o treino dos alunos como estratégia para obter melhores resultados nas avaliações em larga escala. A SME pesquisada realizou, no ano de 2010, um projeto que contava com o apoio de estagiários de graduação para atender aos alunos do 4º ano e que fariam, em 2011, a Prova Brasil, além da adoção de simulados realizados periodicamente para que os estudantes se familiarizassem com o formato da prova. Os relatos evidenciaram o foco do trabalho – não só de professoras, mas também

de gestores – no conteúdo que é cobrado pelas avaliações. As professoras consideraram que a preocupação com a Prova Brasil toma muito tempo do trabalho pedagógico, e um dos gestores escolares relatou preocupar-se o tempo todo se as crianças estão tendo contato com o conteúdo que “vai cair na prova”, e que “[...] tudo está voltado para melhorar os resultados na prova Brasil e no Ideb” (AMARO, 2013, p. 35).

Ao buscar compreender os impactos da avaliação em larga escala sobre o trabalho de professores na rede estadual paulista, o artigo de Novaes (2014) traz dados de entrevistas realizadas com 21 professoras do Ensino Fundamental I de três escolas: uma que havia atingido a meta determinada pela Secretaria da Educação, uma que ficou abaixo e outra que ficou acima da meta estabelecida. Os relatos obtidos apresentaram posições variadas dos docentes em relação à implementação de avaliações externas e índices de desempenho. Entretanto, em sua maioria, destaca-se o argumento de que o problema não estaria nos instrumentos da prova, mas no uso que se faz deles. Mostraram-se, em geral, contrários à divulgação ampla dos resultados e à elaboração de *rankings*, por entenderem que tal classificação se reflete nas escolas e nos professores, que acabam sendo também classificados como “bons” ou “ruins”, a depender dos resultados alcançados. Além disso, alegaram que tal divulgação faz crescer a pressão sobre o trabalho do professor, tanto nas cobranças por parte da Secretaria da Educação do Estado de São Paulo, quanto nos julgamentos negativos da sociedade (NOVAES, 2014).

Uma questão central existente nos relatos dos professores está vinculada à política de bonificação do governo de São Paulo, na qual o desempenho da escola no Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp) está vinculado à concessão de Bônus por Resultados (BR). A maioria deles considera injusta a política de vincular incentivos financeiros aos resultados obtidos nas avaliações e aponta implicações sobre o cotidiano escolar na forma de “[...] conflitos de escolas versus escolas, professores versus professores, precarização do trabalho docente, cobranças por resultados, hierarquização das escolas, tanto pela mídia quanto pela própria comunidade” (NOVAES, 2014, p. 300). Além disso, segundo o autor, os docentes mostraram-se desanimados por precisarem mudar o modo de ensinar, atendendo às determinações dos gestores, tanto externos quanto internos à escola, com o objetivo de atingirem os resultados esperados.

O autor destaca o quadro de precarização do trabalho docente decorrente da política de bonificação por resultados, que acaba por tirar da escola a autonomia de gerir seu próprio currículo, que estimula a realização de treinamentos para as provas, desrespeitando o trabalho dos professores, prejudicando sua autoestima e criando “[...] novas subjetividades e maneiras nem sempre claras de se relacionar com a profissão, o que somente serve para aumentar a angústia e a frustração” (NOVAES, 2014, p. 304).

Martins (2015) apresenta o processo de aplicação de instrumentos de avaliação em larga escala em escolas públicas do Mato Grosso e analisa quantitativamente os resultados tabulados com o objetivo de compreender razões para as dificuldades apresentadas na Prova Brasil por estudantes do 9º ano do Ensino Fundamental na realização de problemas matemáticos. O artigo excerto de dissertação de Mestrado, defendida no âmbito do Programa Observatório da Educação (OBEDUC), foca em Matemática e Iniciação às Ciências, um protocolo de parceria interinstitucional entre três Universidades, com financiamento público, cujo objetivo é diagnosticar as maiores dificuldades em Matemática e Iniciação às Ciências de alunos da Educação Básica de escolas da rede pública de ensino.

O autor aplicou testes para 239 alunos de três escolas participantes do Projeto Observatório da Educação e entrevistou três alunos sobre os seus respectivos testes. Durante a aplicação dos testes, era solicitado que cada aluno utilizasse uma folha para elaborar os cálculos necessários para responder às questões e que, ao final, entregassem também a folha, além da prova. Apenas 49 alunos apresentaram tentativas de resolução dos cálculos de, pelo menos, metade das questões; e somente 12 apresentaram tentativas de resolução em todas as questões da prova. Para o autor, ficou evidenciado que “[...] muitos dos alunos não tentaram fazer o teste e até mesmo preencheram os cartões respostas diretamente, sem assinalar alternativas no teste”, o que o levou a concluir que “[...] foi perceptível, para nós pesquisadores, certo desinteresse por parte dos alunos para com a resolução dos cálculos” (MARTINS, 2015, p. 147).

Em Gonçalves e Brito (2015), encontra-se, também, resultados de pesquisa desenvolvida no âmbito do OBEDUC com foco em Matemática e Iniciação às Ciências (MARTINS, 2015). O artigo descreve a experiência de discussão dos resultados do Ideb com os professores das escolas participantes do projeto como possibilidade de planejamento e de replanejamento das ações em sala de aula e atividades

extraclasse para melhorar o índice. Comentando o contato inicial com os docentes, Gonçalves e Brito (2015, p. 53) destacam que apenas alguns “[...] participantes tinham alguma intimidade ou conhecimento acerca das metodologias empregadas em avaliações de larga escala”, o que exigiu que a equipe procurasse experiências para debate e aprofundamento teórico e prático.

A pesquisa teve como objetivo construir e aplicar “Simulados da Prova Brasil”, a partir da conceitualização do Inep de descritores e com a adoção de uma metodologia adaptada ao objetivo de “[...] identificar dificuldades conceituais de matemática” (GONÇALVES; BRITO, 2015, p. 54) e utilizar os instrumentos elaborados como “[...] parâmetros articuladores das intervenções didáticas dos [...] bolsistas junto às escolas participantes do projeto [...]” (GONÇALVES; BRITO, 2015, p. 54), após a constatação de que os testes do Saeb e Prova Brasil são adequados para avaliar redes ou sistemas de ensino e não alunos individualmente, conforme textos do Inep (GONÇALVES; BRITO, 2015). Concluem os autores que, embora seja importante ampliar as análises sobre os instrumentos elaborados, é possível “[...] confirmar sua utilidade mais direta para identificar os pontos nevrálgicos dos planejamentos iniciais” (GONÇALVES; BRITO, 2015, p. 63). Além disso, eles acrescentam que “[...] várias das ações em parceria com as escolas pautaram-se nos relatórios estatísticos, preferencialmente nos estágios iniciais, servindo como catalisadores dos planos de ação” dos professores (GONÇALVES; BRITO, 2015, p. 63).

Os artigos de Martins (2015) e Gonçalves e Brito (2015) demonstram resultados de pesquisas desenvolvidas no âmbito de projeto de parceria interinstitucional com objetivo de contribuir com uma melhor compreensão e apreensão das avaliações em larga escala pelos profissionais de escolas públicas e estudantes. É relevante notar que os dois estudos foram pautados na elaboração e na aplicação de outros instrumentos, além daqueles que já compõem a avaliação em larga escala. É certo, como asseverou Vianna (2003), que a avaliação não pode existir apartada de outras políticas, pois é fundamental que ela esteja atrelada a programas de formação docente e até mesmo à pesquisa da realidade da escola e dos problemas educacionais. Todavia, há de problematizar-se se as escolas não estão sendo instadas de forma excessiva pelas avaliações em larga escala e pelas pesquisas sobre essas avaliações que introduzem mais instrumentos como forma de análise e de compreensão daqueles já existentes.

O segundo artigo de Amaro (2016, p. 1970) que compõe a amostra deste estudo é fruto da pesquisa “Observatório das Periferias Urbanas” (Observatório da Educação – Inep/Capes) e utilizou-se de observações e entrevistas de duas diretoras e sete professoras que atuavam em duas escolas públicas situadas na região metropolitana periférica do Rio de Janeiro. Nele, o autor aborda a “lógica performativa do trabalho docente” (AMARO, 2016, p. 1970), criada pelas avaliações que exercem controle por meio de metas definidas pelo poder central e que promovem a intensificação do trabalho e alteram o próprio sentido do que é ser professor ao vincularem seu trabalho ao desempenho e à busca de resultados. Mais uma vez foi possível perceber nos relatos a centralidade assumida pelas avaliações em larga escala no trabalho tanto de professores quanto de gestores, a pressão por melhorar os resultados da escola, o treino dos alunos para realização das provas e a frustração de toda a equipe escolar com os resultados insatisfatórios.

Conclui o autor que a ênfase colocada na responsabilização exclusiva das escolas no sucesso ou fracasso dos estudantes nas provas estabelece um cenário de “falsa descentralização” (AMARO, 2016, p. 1972), uma vez que os órgãos centrais se eximem de quaisquer responsabilidades e impõem para os profissionais das escolas o total “[...] compromisso pela ‘aprendizagem’ ou ‘não aprendizagem’ dos alunos, independentemente dos contextos, das diferenças e situações vivenciadas no cotidiano” (AMARO, 2016, p. 1972).

O artigo de Ghisleni e Luce (2016), fruto de tese de doutoramento, aborda os impactos da política nacional de avaliação em larga escala na gestão política e pedagógica de escolas municipais de Porto Alegre e explora as relações e as utilizações elaboradas pelos profissionais da educação. As entrevistas realizadas foram analisadas com o uso do *software* N-Vivo 10 e tiveram o objetivo de detectar as ênfases discursivas dos/as entrevistados/as. O *software* encontrou 12 códigos em um total de 197 referências assim distribuídas: limitações (28); formação de professores (27); caracterização político-pedagógica da mantenedora (26); encaminhamentos dos professores (26); gestão da escola (22); sistema de avaliação municipal (18); conhecimento (16); histórico (9); pais (9); possibilidades (8); evasão e manutenções (5); e encaminhamento considerados possíveis (3).

Para as autoras, a adoção da avaliação em larga escala como mecanismo fulcral da gestão escolar e da prática pedagógica das escolas municipais desconsidera “[...] elementos históricos e de contexto da política e das práticas educativas, aparece no limite entre uma imposição hierárquica” (GHISLENI; LUCE, 2016, p. 17-18), o que gera tensão entre a Secretaria e as escolas que demandam “[...] por mais clareza política, pedagógica e curricular” (GHISLENI; LUCE, 2016, p. 18). Ao examinar o conteúdo das entrevistas, as pesquisadoras ressaltam os destaques observados dos códigos “limitações” que aparecem na forma de “[...] angústia, inconformidade, crítica e busca de alternativas para superar a situação atual da política de avaliação” (GHISLENI; LUCE, 2016, p. 18) e, também, “formação de professores” (GHISLENI; LUCE, 2016, p. 18), visto como caminho para equacionar as limitações impostas pela política avaliativa do órgão central.

Ao finalizarem, as autoras concluem que o desafio colocado é o de “[...] construir a possibilidade de uma política de avaliação que enseje o desvelamento dos conflitos, a crítica e a superação das distâncias entre a gestão política e a gestão pedagógica” (GHISLENI; LUCE, 2016, p. 19) e que a avaliação deve ser “[...] capaz de conciliar fatores vinculados à aprendizagem em condições democráticas e aos elementos contextuais em que a escola está envolvida” (GHISLENI; LUCE, 2016, p. 19).

As sínteses conclusivas de Ghisleni e Luce (2016) em relação à concepção avaliativa dos/as profissionais entrevistados/as aproxima-se daquela sustentada por Vianna (2003, p. 26), na qual a avaliação “não é um valor em si”, devendo, portanto, ser integrada “[...] ao processo de transformação do ensino/aprendizagem e contribuir, desse modo, ativamente, para o processo de transformação dos educandos” (VIANNA, 2003, p. 26). Ao mesmo tempo, distancia-se da perpetrada pela política da gestão municipal de Porto Alegre que, ao que parece, investe em um método que impõe mais dificuldade do que colaboração no desenvolvimento do trabalho docente.

O artigo de Santiago, Santos e Melo (2017) resulta de pesquisa que utilizou dados do Observatório Nacional da Educação Especial (Oneesp) no Estado do Rio de Janeiro (Oerj) a respeito dos processos de avaliação e inclusão em educação. Os dados utilizados foram coletados durante o “Ciclo de Formação Continuada para Professores das Salas de Recursos Multifuncionais” (SANTIAGO; SANTOS; MELO, 2017, p. 633), por meio da realização de seis grupos focais dos quais participaram 63 professores. O texto discute três tipos de avaliação que foram previamente definidas para o segundo encontro do ciclo, no qual foi trabalhado o eixo temático da avaliação do estudante com Necessidades Educacionais Especiais (NEE): a realizada para identificar alunos que necessitam de atendimento na sala de recursos multifuncionais; a avaliação realizada no cotidiano escolar; e as avaliações em larga escala, sendo esta última perspectiva a que será abordada aqui.

Os relatos apresentados reafirmam o fato de grande parte dos alunos com NEE não realizarem as avaliações externas. Em relação às implicações no contexto escolar, as autoras apontam que o processo de inclusão pressupõe que os alunos possam participar de todas as atividades escolares e que o formato padronizado das avaliações em larga escala se revelaria como uma prática excludente. Uma das professoras chega a mencionar a não realização da prova pelos alunos com NEE para evitar prejudicar o rendimento da escola no resultado da avaliação. Nesse sentido, Santiago, Santos e Melo (2017, p. 641) concluem que a “[...] concepção de exame (cultura), configurada a partir de padrões (política) que estimulam a uniformização (prática) dos processos e dos resultados [...] molda as práticas pedagógicas, reduzindo as possibilidades de estímulo à diferença”.

A pesquisa de Ivo e Hypolito (2017) utilizou-se de entrevistas realizadas na rede municipal de ensino de um município de médio porte da região central do Rio Grande do Sul, com gestoras da Secretaria Municipal de Educação (SMED) e com gestoras e professoras de duas escolas, selecionadas a partir de seu desempenho no Ideb, uma delas acima da média nacional (nomeada de Escola Amarela) e a outra abaixo (nomeada Escola Verde). O artigo apresenta algumas similaridades entre as duas escolas: ambas possuem infraestruturas semelhantes, já que a Prefeitura Municipal havia realizado reformas em todas as escolas da rede nos anos anteriores; e a formação dos professores apresentava a grande maioria dos docentes com especialização em ambas as escolas, e três professoras com Mestrado na Escola Amarela e duas na Escola Verde. Ambos os fatos são apontados por uma das gestoras da SMED para afirmar não encontrar justificativa para as diferenças no Ideb entre escolas na rede, supondo que parte das diferenças

viria da falta de compreensão da política pública de avaliação, por parte dos trabalhadores das escolas com baixo desempenho, que geraria um movimento de negação e de resistência.

Entretanto, a pesquisa apontou diferenças importantes entre as duas escolas que são camufladas pelo Ideb e desconsideradas nas falas das gestoras da Secretaria de Educação, a começar pelo contexto educacional no qual estão inseridas: a Escola Verde situa-se em “[...] um contexto de grande vulnerabilidade social, com graves problemas sociais e econômicos” (IVO; HYPOLITO, 2017, p. 796). Os alunos, de modo geral, vivem em condições precárias, uma das professoras relata que crianças vão para a escola com fome, sem material escolar, e alguns sem agasalho, mesmo no inverno. Já a Escola Amarela, conforme informado pela gestora escolar, possui, em sua maioria, alunos de classe média, cujos pais trabalham na Universidade ou na base aérea próxima à escola.

As diferenças relacionadas à origem socioeconômica das crianças têm reflexos inegáveis no trabalho desenvolvido, que, na Escola Verde, inclui “[...] resolver ou contornar problemas sociais, por meio de atividades assistencialistas e sociais que reduzem o tempo para o trabalho pedagógico e escolar de organização e sistematização dos conteúdos que são avaliados pelos exames padronizados” (IVO; HYPOLITO, 2017, p. 798). Além disso, apesar de não serem considerados no Ideb, representam conquistas importantes para alunos, professoras e gestoras. Reflete-se, também, nas expectativas que esses alunos nutrem em relação à educação, quando os alunos da Escola Amarela já se preparam para o Ensino Médio, a grande maioria presta exame para cursos técnicos, e os da Escola Verde mal têm aspiração de terminar o Ensino Médio e estão mais preocupados, conforme apontado pela gestora escolar, com a sobrevivência diária (IVO; HYPOLITO, 2017).

Outra diferença, de acordo com os autores, em parte fomentada pelas avaliações em larga escala, mostra-se nas relações estabelecidas entre a Secretaria de Educação e as escolas. No caso da Escola Amarela, existe uma proximidade, e a Secretaria assume uma posição significativa no contexto escolar. Contudo, os relatos das gestoras da Escola Verde, apesar de pontuarem ações positivas, demonstram um distanciamento e um desconforto em participar das atividades promovidas pela SMED (IVO; HYPOLITO, 2017). Essas relações são ainda mais tensionadas pela postura do representante central do governo municipal, já que as duas gestoras da Escola Verde relataram com insatisfação, em suas entrevistas, o mesmo acontecimento no qual o prefeito, em reunião com os diretores, teria afirmado que as escolas com baixo Ideb envergonham o município (IVO; HYPOLITO, 2017).

Finalizam Ivo e Hypolito (2017) pontuando que a padronização quantitativa e descontextualizada proporcionada pelas avaliações em larga escala acaba sendo mais um fator de ocultamento e de reprodução das desigualdades educacionais de nosso país. Destacam que propor “[...] políticas ‘iguais’ para atingir metas ‘desiguais’ [...] faz com que as metas proposta[s] não sejam atingidas para todas as escolas e todos os estudantes, já que alguns convivem com contextos desfavoráveis” (IVO; HYPOLITO, 2017, p. 806).

Com pesquisa também realizada no estado do Rio de Janeiro, mas tendo como foco a rede municipal da capital, o artigo de Fernandes e Nazareth (2018) resultou de observações em duas escolas da zona sul do município. Os pesquisadores elegeram uma turma do Ciclo de Alfabetização em cada escola para acompanhar a aplicação das provas bimestrais, conselho de classe e entrevistar as professoras.

Segundo os autores, no município do Rio de Janeiro, são realizadas duas avaliações em larga escala: a Prova Rio, instituída em 2009 para avaliação da rede de ensino; e as Provas Bimestrais da rede, que, desde 2010, pretendem acompanhar a aprendizagem dos alunos, além do Índice de Desenvolvimento da Educação do município do Rio de Janeiro (IDE-Rio). As observações e as entrevistas realizadas demonstram as contradições nas percepções das professoras em relação à realização de avaliações em larga escala. Uma das professoras chega a se referir a elas como “massacre”; contudo, há outro relato em que é demonstrado pela professora como positivo o aspecto de habituar as crianças à avaliação e de fazê-las “cair na regra” (FERNANDES; NAZARETH, 2018, p. 901). No entanto, a política de bonificação adotada pelo município, de caráter meritocrático, vinculada ao resultado das avaliações, gera implicações que vão além do momento de aplicação dos exames e se fazem notar em momentos nos quais os alunos simulam a realização das provas, em nítidas situações de treino, mesmo se tratando ainda do Ciclo de Alfabetização. Além disso, há evidências da adoção de prática semelhante à relatada por Santiago, Santos e Melo (2017) na pesquisa apresentada anteriormente, que consiste em comunicar “[...] aos responsáveis pelas crianças com necessidades educativas especiais que elas podem permanecer em casa nos dias dos

exames externos” (FERNANDES; NAZARETH, 2018, p. 893), tendo como intuito aumentar seus índices para garantir a bonificação.

O artigo de Lockmann e Machado (2018) teve o objetivo de questionar o que os números são capazes de medir e o que escaparia às avaliações em larga escala. A opção por priorizar a análise sobre a Avaliação Nacional de Alfabetização (ANA) deu-se por ser ela a primeira forma de avaliação em larga escala “[...] que cogita a possibilidade não só de elaborar instrumentos adaptados, mas também de discutir a construção de matrizes específicas considerando os alunos incluídos” (LOCKMANN; MACHADO, 2018, p. 882). Com base nisso, a pesquisa investigou leis e regulamentos vinculados à ANA e às narrativas obtidas a partir da realização de entrevistas com professores e diretores de sete municípios do Rio Grande do Sul, tendo sido selecionadas as duas escolas de cada município que atendem ao maior número de alunos de inclusão.

Apesar de a documentação vinculada à ANA evidenciar a tentativa de aproximação entre as políticas de avaliação em larga escala e de inclusão, de acordo com os relatos obtidos, o que se mostra no contexto escolar é que, apesar de garantir a realização da prova pelos alunos deficientes e de propor alguns materiais adaptados e protocolos de inclusão para realização das avaliações, esses sujeitos não são contemplados em suas singularidades, o que prejudica o resultado do aluno e, por consequência, diminui o índice alcançado pela escola, sobretudo aquelas que atendem a um grande número de alunos com deficiência (LOCKMANN; MACHADO, 2018).

Como implicações nas escolas, segundo Lockmann e Machado (2018), pode-se perceber que o trabalho com currículos adaptados e o atendimento diferenciado realizado pela escola e pelos professores para atender aos alunos com deficiência é invisibilizado – e, portanto, desvalorizado, desestimulado – quando avaliações em larga escala desconsideram tais especificidades, mas não só isso. A relação “in/exclusão” da situação, na qual os alunos com deficiência são convocados a realizar a ANA, mas esta não é adaptada àqueles, é agravada ao tornar esses sujeitos “[...] indesejados dentro do jogo, pois diminuem os índices escolares” (LOCKMANN; MACHADO, 2018, p. 885).

Ao abordarem a percepção dos professores de ciências sobre as avaliações em larga escala, Garcia, Bizzo e Rosa (2019) utilizaram-se de entrevistas com 77 professores de ciências que atuavam nos anos finais do Ensino Fundamental, em escolas da região do ABC Paulista, composta pelos municípios de Santo André, São Bernardo do Campo, São Caetano do Sul, Diadema, Mauá, Ribeirão Pires e Rio Grande da Serra, entre os anos de 2016 e 2017. Consoante os autores, os docentes destacaram que o processo de desvalorização de seu trabalho que já existia (exemplificado por 90% deles nas precárias condições dos laboratórios de ciências das escolas onde atuavam), foi agravado após o início das avaliações externas, restritas às disciplinas de língua portuguesa e matemática.

O artigo demonstra o cerceamento do desenvolvimento profissional dos professores de ciências provocado pela priorização das disciplinas consideradas nas avaliações em larga escala. Segundo Garcia, Bizzo e Rosa (2019), nos aspectos financeiros, os professores de português e matemática dispunham de mais recursos materiais como papel, tinta e impressão, além de serem favorecidos na compra de *softwares*, livros, materiais didáticos e jogos.

No acompanhamento pedagógico, também foi registrado um tempo maior para reuniões com diretores e coordenadores pedagógicos dado aos professores de português e matemática. Nas reuniões gerais ou nos encontros pedagógicos coletivos, as discussões giravam em torno da Prova Brasil ou do Saesp e de seus respectivos índices, abordando a proficiência dos alunos ou o fluxo escolar, em busca de soluções para aumentar os índices alcançados pelas escolas. Nesses casos, “[...] as discussões eram, quase que exclusivas, sobre as disciplinas de português e matemática” (GARCIA; BIZZO; ROSA, 2019, p. 16). Além disso, os professores municipais de São Bernardo do Campo e de São Caetano do Sul expuseram uma atuação mais próxima dos coordenadores pedagógicos em relação aos professores de português e matemática, dos quais assistiam às aulas, analisavam conteúdos e provas, verificavam resultados, entre outros.

Outro ponto evidenciado pelos relatos relaciona-se às oportunidades profissionais ofertadas no ambiente escolar. Mais de 80% dos professores alegaram que os professores de português e matemática tinham mais tempo para formação continuada, “[...] seja nas secretarias ou diretorias de educação, ou

nas escolas, por meio de participação em palestras, cursos, workshops e reuniões formativas” (GARCIA; BIZZO; ROSA, 2019, p. 16.). Além disso, a maioria das escolas possuía projetos de reforço escolar que se restringiam às disciplinas de português e matemática, inviabilizando a participação dos professores de ciências e negando a oportunidade de aumentar sua remuneração com as aulas de reforço.

Assim sendo, é possível perceber, a partir do estudo de Garcia, Bizzo e Rosa (2019), um processo de intensificação das desigualdades, já existentes anteriormente, entre os professores de língua portuguesa e matemática e os demais, favorecido e potencializado pela restrição da maioria das avaliações em larga escala a essas disciplinas. Os autores argumentam que esse contexto de desmotivação dos professores de ciências pode também reduzir o interesse dos alunos pela disciplina, conclusão que pode ser extrapolada também para o ensino das demais disciplinas preteridas pelas avaliações externas, reforçando o argumento de alguns autores analisados por Bauer, Alavarse e Oliveira (2015). De acordo com os autores, a definição de padrões de proficiência impulsionaria as redes de ensino a direcionarem seus esforços aos conteúdos cobrados nas avaliações, “[...] deixando de lado outros conteúdos fundamentais para a educação básica das crianças e adolescentes em idade escolar” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 1374), em um movimento de empobrecimento/afunilamento curricular.

Por fim, também abordando a inclusão/exclusão no contexto das avaliações em larga escala, o artigo de Witeze e Silva (2019) sistematiza resultados de pesquisa realizada em 2015 em escola estadual de Goiás, que reunia os critérios previamente definidos de oferta de Atendimento Educacional Especializado (AEE) e Ensino Fundamental I e II, e que estivesse entre as escolas com maior número de matrículas de alunos com deficiências, TGD e/ou altas habilidades/superdotação e possuísse as maiores médias no IDEB. Como procedimentos para levantamento de dados, utilizou-se observação participante das rotinas escolares, entrevistas semiestruturadas com gestores, coordenadores pedagógicos, profissionais de apoio à inclusão e professores/as de recursos.

Os depoimentos colhidos na escola, que obteve um aumento do Ideb entre 2005 e 2013 de 5,0 para 7,3, apontam, entre as explicações possíveis, o empenho em trabalhar as habilidades/competências exigidas pela Prova Brasil em língua portuguesa e matemática, que “[...] foi adquirindo centralidade como estratégia de controle dos resultados, mobilizando grandes esforços da equipe a cada nova edição” (WITEZE; SILVA, 2019, p. 9). Pôde ser observado, durante o período de campo da pesquisa, a preparação dos alunos para a edição de 2015 da Prova Brasil, por meio do oferecimento de aulas de reforço aos sábados, nos turnos matutino e vespertino, por professores e estagiários. De acordo com a diretora, a pressão sofrida pela escola no período antecedente à Prova Brasil torna necessária “[...] uma boa estratégia de ação e de uma boa preparação para o teste. Segundo a servidora, em tempos de avaliação externa, é preciso organizar uma ‘força tarefa’ que integre todos os estudantes e funcionários da escola” (WITEZE; SILVA, 2019, p. 9).

Assim como na pesquisa de Lockmann e Machado (2018), foi apontada por Witeze e Silva (2019) a incompatibilidade existente entre o que é proposto nas políticas de inclusão e o formato adotado pelas avaliações em larga escala – nesse caso, a Prova Brasil. De acordo com os pesquisadores, a diretora destaca o trabalho realizado na escola para adaptar conteúdos e avaliações às particularidades dos alunos com deficiências, transtornos globais de desenvolvimento (TGD) e outros distúrbios. Contudo, segundo ela, a Prova Brasil não contempla a perspectiva de inclusão adotada pela escola. Dessa forma, é novamente percebida a não participação de alunos na realização das avaliações para evitar que prejudiquem os índices da escola, tornando o ambiente escolar “[...] espaço de tensão entre diferentes projetos de formação humana, no qual o respeito à diversidade é suplantado pela perspectiva pragmática, mercadológica e classificatória difundida pelos testes padronizados” (WITEZE; SILVA, 2019, p. 1).

As políticas avaliativas produzem reverberações diversas e adversas para os sistemas educativos, os profissionais das escolas e a comunidade escolar em geral. Os artigos examinados trazem informações imprescindíveis para, em um contexto de expansão quantitativa e qualitativa de testes padronizados e índices de qualidade da educação nos âmbitos nacional, estadual e municipal, compreender as eventuais implicações das avaliações em larga escala no cotidiano das escolas da Educação Básica. Mais do que isso, as pesquisas representadas nos artigos permitem conhecer e interpretar como as escolas têm respondido às demandas provocadas por essas políticas que, muitas vezes, se afastam da garantia da educação como direito básico de todos e todas, conforme preconiza a Constituição Federal de 1988 (BRASIL,

1988) e a Lei de Diretrizes e Bases da Educação Nacional - Lei No 9.394, de 20 de dezembro de 1996 (BRASIL, 1996).

REFLEXÕES FINAIS

Este estudo teve o objetivo de, por meio de revisão bibliográfica de artigos do Portal de periódicos da Capes, responder quais têm sido as implicações das avaliações em larga escala no cotidiano das escolas de Educação Básica evidenciadas por investigações que, além do debate teórico, enveredaram pelo interior das instituições educacionais por meio de pesquisas de campo. Foram selecionados 12 artigos resultantes de estudos empíricos em um levantamento que encontrou 118 títulos que abordavam a temática da avaliação em larga escala. Ao finalizar a empreitada, há de destacarem-se aspectos significativos observados nesse processo.

Salienta-se, inicialmente, que a diferença entre a quantidade de artigos teóricos e ensaios publicados é substancial em relação aos artigos resultantes de pesquisas empíricas: 44 no primeiro caso, incluindo-se dois textos sobre o Ensino Superior, e um em relação ao contexto internacional; e 12 no segundo, evidenciando que o debate no campo da avaliação educacional carece de mais estudos de campo.

Outro destaque relevante diz respeito à classificação Qualis Capes (2017-2018) dos periódicos nos quais os artigos foram publicados. Dos 12 textos analisados, nove foram publicados em revistas do estrato A, sendo um em A3, seis artigos em A2 e dois em revista A1. Os demais, três manuscritos, foram publicados em revistas classificadas como B2. Tal fato pode denotar que as revistas com classificação mais alta estão priorizando a seleção de textos que realcem a relação das avaliações em larga escala com o cotidiano escolar.

Ainda cabe apontar a disparidade em relação aos períodos de publicação dos artigos. Nota-se que o primeiro texto teórico registrado na Plataforma da Capes data de 2002, enquanto o primeiro artigo selecionado nesta amostra somente foi publicado em 2013. Ao que tudo indica, as pesquisas de campo foram impulsionadas pela inflexão provocada, especialmente, pela criação da Prova Brasil e do Ideb, em 2005 e 2007 respectivamente, que estreitou a “[...] distância entre o avaliador (governo) e o avaliado (escola)”, conforme aponta Machado (2012, p. 72).

Tal inflexão é, também, reafirmada pelo sistema de avaliação em larga escala que mais aparece como objeto de análise nas investigações veiculadas nos artigos. Dentre as 12 publicações examinadas, seis artigos, metade da amostra, tiveram como foco de sua investigação a Prova Brasil e o Ideb (AMARO, 2013, 2016; GHISLENI; LUCE, 2016; GONÇALVES; BRITO, 2015; IVO; HYPOLITO, 2017; WITEZE; SILVA, 2019).

A centralidade assumida pela avaliação em larga escala no cotidiano das escolas de Educação Básica, identificada nas pesquisas refletidas nos artigos analisados, foi evidenciada por meio de implicações profundas. Sinteticamente, destaca-se: intensificação do trabalho docente (AMARO, 2016; NOVAES, 2014), provocando a adesão, muitas vezes de forma involuntária, à “corrida performativa” (AMARO, 2016, p. 1971); incentivo e fortalecimento de práticas tecnicistas de treinos para os testes (AMARO, 2013, 2016; FERNANDES; NAZARETH, 2018; NOVAES, 2014; WITEZE; SILVA, 2019); estreitamento de currículos, recursos e oportunidades formativas (AMARO, 2013; GARCIA; BIZZO; ROSA, 2019; NOVAES, 2014); incremento das desigualdades entre alunos/as dentro das escolas e entre escolas (GARCIA; BIZZO; ROSA, 2019; IVO; HYPOLITO, 2017); responsabilização pública dos resultados, notadamente, dos fracos ou inferiores resultados (IVO; HYPOLITO, 2017); e ocultação das desigualdades educacionais produzidas pelos contrastes econômicos e sociais próprios da histórica ocupação territorial desordenada dos municípios brasileiros (IVO; HYPOLITO, 2017).

As políticas de bonificação, investigadas tanto por Fernandes e Nazareth (2018), na cidade do Rio de Janeiro, quanto por Novaes (2014), no estado de São Paulo, também apontam implicações importantes no trabalho escolar e exemplificam o que Bonamino e Sousa (2012) chamam de avaliação de terceira geração ou avaliação de responsabilização forte. Discutem as autoras o risco existente, no caso desse tipo de avaliação, de estreitamento do currículo e da situação conhecida como “ensinar para o teste”

(BONAMINO; SOUSA, 2012, p. 383), na qual os professores acabam priorizando os tópicos avaliados em detrimento de outros aspectos importantes do currículo, fato confirmado tanto por Novaes (2014) quanto por Fernandes e Nazareth (2018), bem como processos de precarização do trabalho docente e controle curricular (NOVAES, 2014).

Em três publicações (LOCKMANN; MACHADO, 2018; SANTIAGO; SANTOS; MELO, 2017; WITEZE; SILVA, 2019), foi possível perceber a contradição existente entre as políticas de inclusão e as avaliações em larga escala. Enquanto aquelas buscam promover um atendimento inclusivo, com conteúdos, atividades e avaliações adaptados aos alunos com NEE, estas padronizam sua forma de avaliação e, mesmo no caso da ANA – avaliação pesquisada por Lockmann e Machado (2018), justamente por possuir materiais adaptados e protocolos de inclusão –, desconsideram as especificidades desses alunos. Os processos de inclusão/exclusão dos alunos com NEE mostram-se com forte limitação das avaliações em larga escala, com consequências que podem ir muito além do aumento ou da diminuição do índice das escolas, em direção à exclusão desses alunos dos momentos de avaliação e ao desestímulo do trabalho de inclusão realizado pelos professores.

O exame dos artigos permitiu notar que as equipes escolares e até mesmo as Secretarias de Educação têm dificuldades de compreensão e apropriação dos resultados produzidos pelas avaliações em larga escala (AMARO, 2013; GONÇALVES; BRITO, 2015), reforçando o argumento de Vianna (2003, p. 26, grifo do autor) de que a avaliação “não subsiste isoladamente”, requer estar “[...] associada a outros programas, destacando-se, inicialmente, o de *capacitação docente*; em que a área da avaliação deve integrar, necessariamente, o conjunto das atividades que levam à formação de professores em quaisquer dos níveis de ensino”.

Chamou atenção a constatação de uma lacuna existente nas pesquisas refletidas nos artigos: a ausência de investigações que tenham como objeto captar, compreender e interpretar as percepções de alunos e pais sobre as avaliações em larga escala e suas implicações no trabalho desenvolvido nas escolas e na formação dos estudantes. Além disso, o desenvolvimento deste estudo, ainda que sem o objetivo de esgotar o tema, lançou luzes sobre alguns reflexos das avaliações em larga escala nas escolas de Educação Básica e possibilitou perceber que são muitas as experiências pouco colaborativas que as escolas têm tido com a implementação e as consequências das políticas avaliativas.

Ao retomar a reflexão proposta no início deste artigo, que indagou quais as implicações das avaliações em larga escala no cotidiano das escolas de Educação Básica, cabe realçar que, embora seja possível perceber “influências diferenciadas”, conforme Bauer, Alvarado e Oliveira (2015, p. 1377) apontam, das avaliações no cotidiano escolar, torna-se incontornável reconhecer que as repercussões dessas políticas, como propostas analisadas nos textos da amostra, pouco contribuem com avanços na gestão e na organização do trabalho escolar e do ensino oferecido.

Em que pese a premência de outras pesquisas que invistam em compreender, por meio de investigações já realizadas, como as escolas têm convivido com as avaliações em larga escala no seu cotidiano, vale ressaltar que este estudo revelou que o uso que tem sido feito da avaliação tem servido muito mais a provocar tensionamento no cotidiano escolar do que à construção e alternativas educacionais “com base em evidências” (INEP, 2020, n.p.).

Nesse sentido, também é premente que os gestores centrais, proponentes, em geral, dessas políticas, se disponham e proponham uma avaliação da avaliação, uma meta-avaliação, com a perspectiva de analisar criticamente se os contornos dessas avaliações realmente têm contribuído com a melhoria da qualidade da educação.

REFERÊNCIAS

AMARO, I. A (in) visibilidade da escola: implicações das avaliações externas no contexto escolar. **Educação: Teoria e Prática**, Rio Claro, v. 23, n. 43, p. 24-43, maio/ago. 2013. DOI: <https://doi.org/10.18675/1981-8106.vol23.n43.p24-43>

AMARO, I. As políticas de avaliação em larga escala e trabalho docente: dos discursos eficientistas aos caminhos contrarregulatórios. **RIAAE – Revista Ibero-Americana de Estudos em Educação**, São Carlos, v. 11, n. 4, p. 1960-1978, 2016.

ANDRÉ, M. E. D. A. Pesquisas sobre a escola e pesquisas no cotidiano da escola. **EccoS Revista Científica**, São Paulo, v. 10, p. 133-145, jul. 2008.

BAUER, A. *et al.* Avaliação em larga escala em municípios brasileiros: o que dizem os números? **Estudos em Avaliação Educacional**, São Paulo, v. 26, n. 62, p. 326-352, maio/ago. 2015.

BAUER, A.; ALAVARSE, O. M.; OLIVEIRA, R. P. de. Avaliações em larga escala: uma sistematização do debate. **Educação e Pesquisa**, São Paulo, v. 41, n. especial, p. 1367-1382, dez. 2015. DOI: <http://dx.doi.org/10.1590/S1517-9702201508144607>

BONAMINO, A.; SOUSA, S. Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. **Educação e Pesquisa**, São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012. DOI: <https://doi.org/10.1590/S1517-97022012005000006>

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil. Brasília, DF: Senado, 1988.**

BRASIL. **Lei Nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. Diário Oficial da União: seção 1, Brasília, DF, n. 248, p. 27833-27841, 23 dez. 1996.**

FERNANDES, C. de O.; NAZARETH, H. D. G. de. Resultados de pesquisas sobre as políticas de avaliação em larga escala em educação e seus impactos na escola. **Revista Educação Especial**, Santa Maria, v. 31, n. 63, p. 893-906, out./dez. 2018. DOI: <https://doi.org/10.5902/1984686X32906>

FONTANIVE, N. S. A divulgação dos resultados das avaliações dos sistemas escolares: limitações e perspectivas. **Ensaio: Avaliação e Políticas Públicas Educacionais**, Rio de Janeiro, v. 21, n. 78, p. 83-100, 2013.

GARCIA, P. S.; BIZZO, N.; ROSA, S. S. da. Políticas de avaliação e o desenvolvimento profissional do professor de ciências. **Educação**, Santa Maria, v. 44, p. 1-24, 2019.

GHISLENI, A. C.; LUCE, M. B. Efeitos da política nacional de avaliação em larga escala na gestão política e pedagógica de escolas municipais de Porto Alegre: explorando relações e utilizações. **Arquivos Analíticos de Políticas Educativas [online]**, v. 24, n. 71, p. 1-21, jun. 2016. DOI: <https://doi.org/10.14507/epaa.24.2328>

GONÇALVES, W. V.; BRITO, A. J. de. Das avaliações em larga escala às intervenções didáticas: experiências do Observatório da Educação (OBEDUC) - polo UNEMAT. **Revista Eventos Pedagógicos, Cáceres**, v. 6, n. 1, edição especial temática, p. 52-65, jan./maio 2015.

INEP. Instituto Nacional de Estudos e Pesquisas Anísio Teixeira. **Saeb – 30 anos. 2020.** Disponível em: <http://portal.inep.gov.br/web/guest/educacao-basica/saeb>. Acesso em: 7 jun. 2020.

IVO, A. A.; HYPOLITO, A. M. Sistemas de avaliação em larga escala e repercussões em diferentes contextos escolares: limites da padronização gerencialista. **RBPAAE [online]**, v. 33, n. 3, p. 791-809, set./dez. 2017.

LOCKMANN, K.; MACHADO, R. B. Invisibilidades na Avaliação Nacional da Alfabetização (ANA): os alunos com deficiência e o trabalho docente. **Revista Educação Especial**, Santa Maria, v. 31, n. 63, p. 879-892, out./dez. 2018. DOI: <https://doi.org/10.5902/1984686X33090>

MACHADO, C. Avaliação externa e gestão escolar: reflexões sobre o uso dos resultados. **Revista @ mbienteeducação**, São Paulo, v. 5, n. 1, p. 70-82, jan./jun. 2012.

MACHADO, C. Políticas estaduais de avaliação externa: tendências e implicações. **Quaestio**, Sorocaba, v. 22, n. 1, p. 205-223, jan./abr. 2020. DOI: <http://dx.doi.org/10.22483/2177-5796.2020v22n1p205-223>

MACHADO, C.; ALAVARSE, O. M.; ARCAS, P. H. Sistemas estaduais de avaliação: interfaces com qualidade e gestão da educação. **Revista Brasileira de Política e Administração da Educação** [online], v. 31, n. 3, p. 667-680, set./dez. 2015. DOI: <https://doi.org/10.21573/vol31n32015.63800>

MARTINS, E. A. As relações entre a aprendizagem matemática e a aprendizagem de ciências naturais: um estudo em três escolas acompanhadas pelo Projeto Observatório da Educação com foco em matemática e iniciação às ciências. **Revista Eventos Pedagógicos**, Cáceres, v. 6, n. 1, edição especial temática, p. 130-149, jan./maio 2015.

NOVAES, L. C. Os impactos da avaliação externa sobre o trabalho de professores na rede estadual paulista. **Roteiro**, Joaçaba, v. 39, n. 2, p. 283-310, jul./dez. 2014.

SANTIAGO, M. C.; SANTOS, M. P. dos; MELO, S. C. de. Inclusão em educação: processos de avaliação em questão. **Ensaio: Avaliação e Políticas Públicas Educacionais**, Rio de Janeiro, v. 25, n. 96, p. 632-651, jul./set. 2017. DOI: <https://doi.org/10.1590/s0104-40362017002500652>

SILVA, M. C. V.; MELETTI, S. M. F. Estudantes com Necessidades Educacionais Especiais nas Avaliações em Larga Escala: Prova Brasil e ENEM. **Revista Brasileira de Educação Especial**, Marília, v. 20, n. 1, p. 53-68, jan./mar. 2014.

VIANNA, H. M. Fundamentos de um programa de avaliação educacional. **Estudos em Avaliação Educacional**, São Paulo, n. 28, p. 23-38, jul./dez. 2003. DOI: <http://dx.doi.org/10.18222/eae02820032168>

VIEIRA S. L.; VIDAL, E. M.; NOGUEIRA, J. F. F. Gestão da aprendizagem em tempos de Ideb: percepções dos docentes. **RBPAE** [online], v. 31, n. 1, p. 85-106, jan./abr. 2015. DOI: <https://doi.org/10.21573/vol31n12015.58916>

WITEZE, E. M.; SILVA, R. H. dos R. Avaliação em larga escala e educação especial inclusiva: o embate entre duas lógicas. **Revista Educação Especial**, Santa Maria, v. 32, p. 111, 2019. DOI: <https://doi.org/10.5902/1984686X39687>