

AVALIAÇÃO PADRÃO NO SISTEMA EDUCACIONAL: ANÁLISE TEÓRICA SOBRE A AVALIAÇÃO EM LARGA ESCALA

Maria Eunice Ribeiro Machado

Veni Creator Christian University.

<https://orcid.org/0009-0000-7750-8951>

E-mail: eunice.prelindinha@gmail.com

Sandra Karina Mendes do Vale

<https://orcid.org/0009-0009-5684-8303>

E-mail: karinamendes2232@gmail.com

DOI-Geral: <http://dx.doi.org/10.47538/RA-2025.V4N3>

DOI-Individual: <http://dx.doi.org/10.47538/RA-2025.V4N3-16>

RESUMO: Busca apresentar os estudos teóricos realizados sobre a avaliação padrão no sistema educacional por entendermos que este tipo de avaliação tem se tornado um tópico central nas discussões sobre a qualidade e a equidade da educação em diversas partes do mundo. Neste contexto, esta revisão de literatura visa analisar os primeiros estudos, os locais mais pesquisados, os subtemas associados, as principais teorias, as metodologias utilizadas e as lacunas existentes nesse campo. Os primeiros estudos sobre avaliação em larga escala datam de décadas passadas, destacando-se por abordagens que buscavam entender o impacto das avaliações no desempenho escolar e nas políticas educacionais. Esses trabalhos pioneiros lançaram as bases para discussões mais contemporâneas sobre a efetividade e a relevância das avaliações no contexto educacional brasileiro. O método utilizado para o desenvolvimento deste estudo foi a revisão da por entender-se que, é um tipo de abordagem utilizada em estudos que se baseia na análise e interpretação de fontes já existentes, como livros, artigos científicos, teses, dissertações e outros documentos. Os resultados da pesquisa evidenciaram que as avaliações em larga escala são frequentemente utilizadas como indicadores de desempenho que podem impactar negativamente as comunidades escolares, levando a estigmatizações e desigualdades sociais, uma vez que a responsabilidade pelo sucesso ou fracasso educacional é muitas vezes atribuída às escolas e aos educadores, enquanto fatores externos, como a infraestrutura e o financiamento, são negligenciados. Portanto, é fundamental que as avaliações em larga escala sejam repensadas e reformuladas, buscando uma abordagem mais justa e inclusiva, que considere as diversas realidades dos estudantes e que foque não apenas em resultados quantitativos, mas também na qualidade e na equidade da educação.

PALAVRAS-CHAVE: Avaliação padrão. Sistema educacional. Avaliação em larga escala.

STANDARDIZED ASSESSMENT IN THE EDUCATIONAL SYSTEM: THEORETICAL ANALYSIS OF LARGE-SCALE ASSESSMENT

ABSTRACT: It aims to present the theoretical studies carried out on standardized assessment in the educational system, understanding that this type of evaluation has become a central topic in discussions about the quality and equity of education worldwide. In this context, this literature review seeks to analyze the initial studies, the

most researched locations, associated subtopics, main theories, methodologies used, and existing gaps in the field. The first studies on large-scale assessment date back decades, focusing on understanding the impact of assessments on school performance and educational policies. These pioneering works laid the foundation for more contemporary discussions on the effectiveness and relevance of assessments in the Brazilian educational context. The method used was a literature review, understood as an approach based on the analysis and interpretation of existing sources such as books, scientific articles, theses, dissertations, and other documents. The research results show that large-scale assessments are often used as performance indicators that can negatively impact school communities, leading to stigmatization and social inequalities, as responsibility for educational success or failure is often attributed to schools and educators, while external factors such as infrastructure and funding are overlooked. Therefore, it is essential that large-scale assessments be reconsidered and restructured, seeking a fairer and more inclusive approach that considers students' diverse realities and focuses not only on quantitative results but also on the quality and equity of education.

KEYWORDS: Standardized assessment. Educational system. Large-scale assessment.

INTRODUÇÃO

Busca apresentar os estudos teóricos realizados sobre a avaliação padrão no sistema educacional por entendermos que este tipo de avaliação tem se tornado um tópico central nas discussões sobre a qualidade e a equidade da educação em diversas partes do mundo.

Compreendemos que a avaliação da aprendizagem escolar é um tema abrangente que permite várias compreensões ao longo do tempo. Tradicionalmente, a avaliação era vista principalmente como um instrumento de medida do desempenho acadêmico dos alunos, focando em resultados quantitativos, como provas e notas. No entanto, essa concepção tem se expandido, considerando também aspectos qualitativos do processo de ensino-aprendizagem.

Nesse contexto, abordar a avaliação em larga escala implica considerar a tomada de decisões que visem aprimorar tanto o ensino quanto a aprendizagem dos alunos. Refletir sobre como orientar a avaliação nesse sentido requer a definição de objetivos claros e a análise das funções que essa avaliação deve desempenhar.

Entendemos a escola como um espaço que, sob a pressão da lógica competitiva, incentiva os professores a realizar a avaliação da aprendizagem em sala de aula de modo semelhante à avaliação em larga escala. Isso se traduz em testes e questionários que se

concentram exclusivamente nos resultados, baseando-se nas matrizes das avaliações nacionais ou estaduais, sem se preocupar em analisar o processo pelo qual os alunos desenvolvem suas aprendizagens.

A utilização de avaliações externas, especialmente aquelas realizadas pelas escolas, exige um conjunto de ações que ultrapassa a mera aplicação de provas e questionários, bem como a divulgação de seus resultados. Isso se deve ao fato de que diversos fatores do contexto local ou regional de uma secretaria de educação podem influenciar positivamente ou negativamente a compreensão e a utilização dos resultados em suas diferentes esferas (Silva et al., 2013).

Com os resultados dessa avaliação, é viável elaborar um diagnóstico que permita implementar procedimentos e técnicas voltadas à melhoria da qualidade da educação em nosso país. Cada região, ou melhor, cada unidade federativa ou escola, pode efetivamente analisar os resultados divulgados e, dessa forma, desenvolver estratégias específicas para sua rede ou instituição, visando aprimorar a qualidade do ensino (Nogueira, 2021).

O artigo possui como objetivo geral, realizar um mapeamento da avaliação com ênfase na identificação das lacunas de pesquisa, bem como as principais ênfases temáticas, teóricas e metodológicas. Por meio de uma revisão teórica, a pesquisa pretende atingir os seguintes objetivos específicos, identificar os primeiros estudos e os locais onde o tema é mais pesquisado analisando as principais teorias, autores e metodologias associadas, explorar os subtemas relacionados à avaliação em larga escala evidenciando as ênfases, resultados e lacunas existentes na literatura e, por fim, analisar a evolução dos estudos sobre a temática destacando tendências e áreas que requerem mais investigação.

O método utilizado para o desenvolvimento deste estudo foi a revisão da literatura por entendermos que, é um tipo de abordagem utilizada em estudos que se baseia na análise e interpretação de fontes já existentes, como livros, artigos científicos, teses, dissertações e outros documentos. Essa metodologia é fundamental para fundamentar teoricamente o trabalho, contextualizar a problemática de pesquisa e identificar lacunas no conhecimento.

A partir das palavras-chave avaliação padrão, sistema educacional e avaliação de larga escala realizamos buscas em fontes secundárias de plataformas digitais como Scielo,

Google Acadêmico, IBICIT, entre outros. Assim, a partir das pesquisas realizadas por Dias Sobrinho (2002), Bonamino; Sousa (2012), Sousa; Ferreira (2019) e Borges; Almeida e Santana (2021) foi possível compreender e discutir sobre suas teorias aqui apresentadas.

O texto apresenta uma estrutura bem definida, começando com os estudos teóricos da avaliação em larga escala, onde apresentamos os estudos, subtemas, e locais da pesquisa, metodologia e lacunas. Em seguida, fazemos uma análise dos dados coletados a partir revisão da literatura sobre a temática. O estudo prossegue com a apresentação de algumas teorias relacionadas às três gerações da avaliação em larga escala. Por fim, nas considerações finais, são analisados os tópicos discutidos anteriormente, avaliando se os objetivos propostos foram atingidos.

AVALIAÇÃO EM LARGA ESCALA: ESTUDOS TEÓRICOS

Nos últimos vinte anos, as políticas educacionais brasileiras passaram por marcos significativos, entre os quais se destacam as avaliações que incorporam elementos comuns a propostas implementadas em outros países, refletindo uma agenda global. Essas iniciativas de avaliação, além de outros objetivos, estão conectadas à melhoria da qualidade do ensino e, em última instância, contribuem para o estabelecimento de novos parâmetros de gestão nos sistemas educacionais (Bonamino; Sousa, 2012).

Diante disto, a revisão da literatura nos mostra os estudos realizados com relação às teorias, metodologias, resultados e lacunas existentes referentes à avaliação em larga escala. Autores renomados trazem contribuições significativas que nos ajudam a compreender sobre tal temática.

Quadro 1 – Resultados do levantamento – parte 1

Primeiros estudos	Locais da pesquisa	Subtemas
Universidade e Avaliação: entre a ética e o mercado	Universidades brasileiras	• A relação entre Universidade e Avaliação • Ética e Mercado
Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola.	Legislações e diretrizes nacionais que regem a educação básica no Brasil	• Avaliação e currículos • Contexto histórico • Impacto nas práticas pedagógicas • Desigualdades educacionais

Avaliação de larga escala e da aprendizagem na escola: um diálogo necessário	Práticas e contextos no Brasil, especialmente em São Paulo	• Avaliação de Larga Escala vs. Avaliação da Aprendizagem • Perspectivas Históricas da Avaliação Educacional • Impactos das Avaliações na Prática Pedagógica • Lacunas e Necessidades Identificadas
Avaliação Educacional: o Saeb, seus pressupostos, finalidades e repercussões	Documentos oficiais do Inep, MEC e CAEd. Documentos de Referência – Versão 1.0 do Saeb (2018).	• Sistema de Avaliação da Educação Básica (Saeb) e sua influência na formulação de Políticas Públicas Educacionais. • Críticas à lógica neoliberal na avaliação educacional e suas consequências para a escola e o estudante. • Importância da análise documental para compreender o Saeb, suas características e objetivos. • Relação entre avaliação, currículo e formação de competências.

Fonte: Elaborado pelas autoras.

Estudos realizados por Sousa e Ferreira (2019) destacam a importância das avaliações de larga escala no contexto educacional, sugerindo que elas funcionam como ferramentas que oferecem uma visão clara dos pontos fortes e fracos do sistema educacional. Entretanto, é importante levar em consideração alguns aspectos. Em primeiro lugar, a ideia de que as avaliações revelam o que “vai bem” e o que “não vai” na educação implica que os resultados obtidos são interpretados de forma objetiva e que as conclusões tiradas são indiscutíveis. No entanto, isso pode ser problemático.

Entendemos que, as avaliações em larga escala frequentemente se concentram em aspectos quantificáveis do aprendizado, como habilidades em leitura e matemática, e podem não captar a complexidade do processo educativo, incluindo fatores socioeconômicos, culturais e emocionais que influenciam o desempenho dos alunos.

Além disso, essas avaliações tendem a promover uma cultura de comparação e competitividade entre escolas, estados e até países. Essa abordagem pode levar a um enfoque excessivo em resultados numéricos, em detrimento de práticas pedagógicas mais holísticas e menos mensuráveis. Isso pode resultar em um “ensino para o teste”, onde

professores e alunos focam apenas na preparação para essas avaliações, negligenciando aspectos importantes da aprendizagem, como criatividade, pensamento crítico e habilidades socioemocionais.

Vejamos as lacunas encontradas em algumas pesquisas de autores que discutem sobre a avaliação em larga escala em vários contextos.

Quadro 2 – Resultados do levantamento – parte 2

Autores	Metodologias	Lacunas
Dias Sobrinho	• Caráter qualitativo • Abordagem analítica e crítica • Revisão bibliográfica e estudos de caso de instituições de ensino superior.	• Atualização dos dados • Falta de estudos comparativos internacionais sobre o tema
Bonamino e Sousa (2012)	• Abordagem qualitativa • Análise documental • Revisão de literatura.	• Falta de dados empíricos • Desconsideração de contextos regionais • Necessidade de estudos longitudinais
Sousa; Ferreira (2019)	• Revisão teórica	• Necessita de uma integração efetiva entre as avaliações em larga escala e os métodos de avaliação prática na sala de aula • Necessidade de políticas educacionais que considerem as realidades locais.
Borges; Almeida e Santana (2021)	• Pesquisa de caráter qualitativo com enfoque crítico participativo. - Análise documental dos Documentos de Referência do Saeb e revisão da literatura sobre avaliação educacional.	• A necessidade de um entendimento mais profundo sobre as lacunas nas Políticas Públicas Educacionais em relação ao Saeb, além das falhas nas práticas pedagógicas nas escolas. • Falta de integração entre as avaliações externas e o contexto das práticas educacionais nas escolas, evidenciando a necessidade de sistematização e diálogo entre os diferentes níveis de avaliação.

Fonte: elaborado pelas autoras.

Os resultados das avaliações em larga escala devem ser utilizados como ferramenta para aprimorar os objetivos educacionais, incentivando, assim, a aprendizagem dos alunos, que é um dos principais focos. O uso adequado desses

resultados nos permite planejar intervenções que contribuirão significativamente para o processo educativo. Para garantir a eficácia dessas intervenções e promover uma educação de qualidade, é fundamental que as ações sejam elaboradas por grupos educacionais, seja de forma individual, como nas escolas, ou em conjunto, como em municípios ou estados. Esses coletivos devem pensar em estratégias que atendam às necessidades específicas dos seus grupos e dificuldades, sempre com o objetivo de melhorar a aprendizagem dos alunos. Essa abordagem, por sua vez, refletirá positivamente na qualidade da educação e nos resultados do IDEB¹ (Nogueira, 2021).

PRIMEIROS ESTUDOS, LOCAIS DA PESQUISA E OS SUBTEMAS ASSOCIADOS

As avaliações de larga escala, que têm como objetivo analisar o desempenho dos alunos são frequentemente utilizadas como ferramentas de diagnóstico e planejamento educacional. Estas avaliações, que incluem testes padronizados, buscam oferecer uma visão geral sobre a eficácia dos sistemas de ensino, permitindo comparações entre diferentes regiões e contextos.

Em 2001, a promulgação da Lei nº 10.172 instituiu o Plano Nacional de Educação - PNE (Brasil, 2001), fruto da participação ativa da sociedade, associações e entidades educacionais. Apesar das críticas que recebeu, a elaboração do PNE ocorreu por meio de um processo democrático, que assegurou ampla discussão. Desde sua implementação, o plano enfatiza a relevância dos sistemas de avaliação em todos os níveis de ensino. Sua proposta busca desenvolver sistemas de informação e avaliação para as diversas

1 O Índice de Desenvolvimento da Educação Básica (Ideb) foi criado em 2007 e reúne, em um só indicador, os resultados de dois conceitos igualmente importantes para a qualidade da educação: o fluxo escolar e as médias de desempenho nas avaliações. O Ideb é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho no Sistema de Avaliação da Educação Básica (SAEB). O Ideb agrega ao enfoque pedagógico das avaliações em larga escala a possibilidade de resultados sintéticos, facilmente assimiláveis, e que permitem traçar metas de qualidade educacional para os sistemas. O índice varia de 0 a 10. A combinação entre fluxo e aprendizagem tem o mérito de equilibrar as duas dimensões: se um sistema de ensino reter seus alunos para obter resultados de melhor qualidade no Saeb, o fator fluxo será alterado, indicando a necessidade de melhoria do sistema. Se, ao contrário, o sistema apressar a aprovação do aluno sem qualidade, o resultado das avaliações indicará igualmente a necessidade de melhoria do sistema. Disponível: <https://www.gov.br/>

modalidades de ensino, com o intuito de aprimorar a coleta e a disseminação de dados, além de melhorar a gestão e a qualidade da educação.

No que tange ao Ensino Fundamental, o PNE reafirma a importância de consolidar e aprimorar tanto o Censo Escolar quanto o Sistema de Avaliação da Educação Básica (Saeb), além de propor a criação de sistemas complementares nos estados e municípios. As metas e objetivos do PNE também incluem a articulação das funções de supervisão e inspeção no sistema de avaliação, ressaltando a importância de:

Assegurar a elevação progressiva do nível de desempenho dos alunos, mediante a implantação, em todos os sistemas de ensino, de um programa de monitoramento que utilize os indicadores do Sistema Nacional de Avaliação da Educação Básica e dos sistemas de avaliação dos estados e municípios que venham a ser desenvolvidos (Brasil, 2001, S/P).

Nota-se a importância de melhorar continuamente o desempenho dos alunos nas escolas. Para atingir esse objetivo, é essencial implementar um programa de monitoramento em todos os sistemas de ensino, que utilize indicadores estabelecidos pelo Sistema Nacional de Avaliação da Educação Básica, além de sistemas de avaliação específicos desenvolvidos por estados e municípios.

A partir dos dados apresentados no Quadro 1, podemos afirmar que Dias Sobrinho (2002) realiza uma análise com base em exemplos de universidades brasileiras, oferecendo uma perspectiva regional que reflete as particularidades do contexto educacional no Brasil. O autor explora como as instituições de ensino superior são avaliadas e as implicações dessa avaliação no seu funcionamento e na qualidade do ensino.

Dias Sobrinho (2002) discute ainda o dilema ético que surge na avaliação acadêmica, especialmente quando as métricas de avaliação são influenciadas por interesses mercadológicos. Suas teorias sobre avaliação institucional fazem reflexões sobre qualidade e a dualidade entre a missão educativa e a concorrência no mercado educacional. A análise crítica sobre a mercantilização do ensino superior também é um aspecto relevante, levando em consideração teorias da educação e da sociologia.

Já Bonamino e Sousa (2012) analisam a evolução das teorias de avaliação da educação básica no Brasil, destacando três gerações. A relação entre as práticas

avaliativas e os currículos implementados nas escolas. Por meio de sua abordagem qualitativa, utilizando análise documental e revisão de literatura, permite uma compreensão aprofundada das políticas e práticas avaliativas em suas diferentes gerações, destacando a evolução dos padrões e paradigmas ao longo do tempo.

As autoras abordam em suas teorias a evolução das políticas de avaliação ao longo das décadas e como elas se relacionam com as diretrizes educacionais no Brasil (Bonamino; Sousa, 2012). Sua análise se baseia nas legislações e diretrizes nacionais que regem a educação básica no Brasil, contemplando o cenário educacional do país como um todo. Isso inclui dados de políticas federais e estaduais, além de práticas em diferentes contextos escolares.

As pesquisas realizadas por Sousa e Ferreira (2019) discutem as diferenças e interações entre as avaliações em larga escala e as avaliações realizadas em sala de aula. Consideram a avaliação de larga escala uma ferramenta de diagnóstico educacional e formulação de políticas públicas. As autoras fazem uma evolução histórica da avaliação educacional, destacando a transição de enfoques quantitativos para metodologias que consideram contextos sociais e educativos.

Sousa e Ferreira (2019) tratam da influência das avaliações em larga escala na forma como os professores avaliam, muitas vezes limitando as práticas avaliativas a testes padronizados. Elas reforçam que é necessário um diálogo mais consistente entre diferentes formas de avaliação, como a larga escala e a da aprendizagem em sala de aula, para promover melhorias significativas na educação.

PRINCIPAIS TEORIAS, METODOLOGIAS, RESULTADOS E LACUNAS SOBRE AS TRÊS GERAÇÕES DA AVALIAÇÃO EM LARGA ESCALA

De acordo com Dias Sobrinho (2002), a importância da avaliação está diretamente relacionada ao poder que ela exerce, o que explica seu destaque em praticamente todas as áreas da intervenção social atualmente. É fundamental reiterar essa ideia para evidenciar que a avaliação não se limita apenas ao contexto escolar ou educacional. Na verdade, ela transcende esses setores e, gradualmente, tem ampliado suas aplicações e impactos nos campos econômico e político. O autor reforça essa argumentação ao assinalar que, nos

países modernos, a avaliação é uma ferramenta essencial utilizada para implementar reformas do Estado.

Existem críticas em relação à adequação das avaliações externas às exigências das agências internacionais que as definem, uma vez que elas determinam o currículo e valorizam a educação formal em detrimento da educação não formal. Embora reconheçam a importância de certos conhecimentos sistematizados ao longo da história que devem ser ensinados, essas críticas também ressaltam que o indivíduo adquire saberes valiosos por meio de suas experiências sociais e atividades extracurriculares. Assim, enfatizam a necessidade de uma formação que prepare o indivíduo para a vida, indo além do simples ensino voltado para a realização de testes (Borges; Almeida; Santana, 2021).

Ainda de acordo com Borges, Almeida; e Santana (2021) a avaliação está intrinsecamente ligada a outros instrumentos fundamentais nos processos de ensino e aprendizagem. Ela pode ser extremamente produtiva quando encarada como um meio de verificação do aprendizado, possibilitando ajustes no planejamento pedagógico. É de interesse social que todas as escolas ofereçam um ensino de qualidade, promovendo a emancipação dos indivíduos, especialmente das classes populares, que constituem uma parcela significativa da rede pública de ensino.

A avaliação em larga escala se tornou um instrumento significativo de poder, conforme observou Sordi (2002), e passou a receber forte apoio das mídias. Essa centralidade das avaliações possibilitou uma melhor compreensão de diferentes instrumentos, como o Sistema de Avaliação da Educação Básica (SAEB), a Avaliação Nacional de Alfabetização (ANA), a Prova Brasil e o Exame Nacional do Ensino Médio (ENEM). Além disso, inclui também as avaliações na Educação Superior, como o Exame Nacional da Educação Superior (ENADE) e a Avaliação da Pós-Graduação realizada pela CAPES.

Bonamino e Sousa (2012) destacam que, o impacto dessas avaliações, pode ser compreendido ao analisar as três gerações da avaliação educacional de larga escala no Brasil. A primeira geração focou na realização de diagnósticos relacionados à qualidade da educação.

Assim sendo, desconsiderar a relevância da avaliação de larga escala, sobretudo no Brasil, seria leviandade e injustiça, principalmente porque

ela representa não só aqui, mas em muitos países, uma ação eficaz de reestruturação da Escola e do Sistema de Educação, capaz de definir critérios essenciais pelos quais se deve compreender a qualidade do trabalho educacional (Sousa; Ferreira, 2019, p. 16).

Compreende-se que a avaliação de larga escala é um importante componente das políticas educacionais, utilizado para medir e monitorar a qualidade do ensino e a aprendizagem dos alunos em um contexto mais amplo. Historicamente, essa avaliação passou por três gerações distintas, cada uma com características, objetivos e metodologias específica.

No Brasil, avaliações de primeira geração são aquelas cuja finalidade é acompanhar a evolução da qualidade da educação. De um modo geral, essas avaliações divulgam seus resultados na Internet, para consulta pública, ou utilizam-se da mídia ou de outras formas de disseminação, sem que os resultados da avaliação sejam devolvidos para as escolas (Bonamino; Sousa, 2012, p. 375).

A primeira geração de avaliações educacionais no Brasil tem um foco específico na análise diagnóstica da qualidade do sistema educacional. Esse tipo de avaliação busca compreender e identificar as características e o desempenho da educação oferecida nas escolas brasileiras. No entanto, é importante ressaltar que essas avaliações não têm o efeito de gerar consequências diretas para as instituições de ensino ou para os conteúdos do currículo. Isso significa que, apesar de fornecer informações valiosas sobre a situação educacional, elas não resultam em mudanças obrigatórias ou imediatas nas práticas escolares ou nas diretrizes de ensino.

A primeira geração de avaliação de larga escala, focava na coleta de dados sobre o desempenho escolar em larga escala. Avaliações como testes padronizados foram realizar para entender o nível de conhecimento dos alunos em disciplinas específicas, como matemática e língua portuguesa. Uma das intenções era comparar resultados entre diferentes escolas, regiões e até países.

As avaliações de segunda geração, por sua parte, incluem não apenas a divulgação pública, mas também o retorno dos resultados para as escolas, sem que isso acarrete consequências diretas. Neste contexto, as consequências são de natureza simbólica e resultam da divulgação e da apropriação das informações sobre os resultados escolares por parte dos pais e da sociedade. Esse tipo de mecanismo de responsabilização parte do pressuposto de que o conhecimento dos resultados estimula a mobilização das equipes

escolares em busca da melhoria da educação, além de gerar pressão por parte dos pais e da comunidade sobre a escola (Zaponi; Valença, 2009). A Prova Brasil, é considerada avaliação de segunda geração, pois, a mesma:

Que ocorre a cada dois anos, foi idealizada para produzir informações a respeito do ensino oferecido por município e escola, com o objetivo de auxiliar os governantes nas decisões sobre o direcionamento de recursos técnicos e financeiros e no estabelecimento de metas e implantação de ações pedagógicas e administrativas, visando à melhoria da qualidade do ensino. De outra parte, considera-se que essa avaliação pode funcionar como um elemento de pressão, para pais e responsáveis, por melhoria da qualidade da educação de seus filhos, uma vez que, a partir da divulgação dos resultados, eles podem cobrar providências para que a escola melhore (Bonamino; Sousa, 2012, p. 378-379).

Portanto, este tipo de avaliação possui ênfase qualitativa e diagnóstica. Além da medida quantitativa, passou-se a valorizar aspectos qualitativos da educação. A segunda geração enfatizava a importância do feedback para professores e instituições, permitindo uma reflexão sobre as práticas pedagógicas.

Já a terceira geração é uma avaliação Formativa e Inclusiva, como o SAEB. Para, Bonamino e Sousa (2012), há especificidades nas avaliações educacionais e no uso de seus resultados que ilustram as características das relações entre avaliações de terceira geração, políticas de responsabilização e currículo escolar.

As avaliações são integradas ao processo educativo, ajudando no acompanhamento contínuo da aprendizagem e permitindo ajustes pedagógicos em tempo real. Há um maior foco em garantir que todos os estudantes, independentemente de suas condições sociais, econômicas ou culturais, tenham acesso a uma educação de qualidade. Avaliações se propõem a identificar e eliminar desigualdades. Mas, existem autores que discordam desta concepção e argumentam que:

A avaliação a ser desenvolvida pelo professor em sala de aula não pode ficar condicionada a esse formato da avaliação de larga escala que limita e restringe a coleta de informação e a análise de resultados por utilizar, quase que maciçamente, a medição das respostas por itens do tipo objetivo. As medidas realizadas pelo professor devem ir além dos resultados de uma prova objetiva. Devem permitir identificar não só o que o aluno sabe, mas compreender por que não sabe (Sousa; Ferreira, 2019, p.18-19).

Corroboramos com a ideia dos autores, pois, a avaliação realizada pelos professores em sala de aula não deve se basear exclusivamente em testes de larga escala, que geralmente utilizam questões objetivas. Essas avaliações limitam a coleta de informações e a análise dos resultados, pois focam apenas nas respostas corretas. As avaliações dos professores devem ir além desse formato, permitindo não apenas identificar o conhecimento dos alunos, mas também entender suas dificuldades e lacunas no aprendizado. Isso sugere uma abordagem mais profunda e qualitativa na avaliação, que considera o processo de aprendizagem de forma mais ampla.

A avaliação das aprendizagens é um esforço conjunto entre professor e aluno, em que é fundamental que o aluno compreenda claramente as expectativas relacionadas ao seu aprendizado, as metas a serem alcançadas e, a cada momento, tenha a capacidade de se situar, bem como informar o professor sobre seu progresso (Sousa; Ferreira, 2019).

As avaliações de segunda e terceira geração, juntamente com a implementação de políticas de responsabilização que se baseiam em consequências simbólicas e materiais, visam estabelecer incentivos para que os professores se empenhem no aprendizado dos alunos.

Contudo, evidências tanto nacionais quanto internacionais indicam que a utilização dos resultados das avaliações de terceira geração para fundamentar iniciativas de responsabilização rigorosas pode acarretar riscos para o currículo escolar. Um dos principais riscos é a prática conhecida como “ensinar para o teste”, na qual os docentes priorizam seu esforço em tópicos que são objetos de avaliação, negligenciando aspectos significativos do currículo, inclusive aqueles relacionados ao desenvolvimento não cognitivo dos estudantes (Bonamino; Sousa, 2012).

Dentro desse contexto, o currículo delineado pela BNCC serve como a espinha dorsal das avaliações, sejam elas internas ou externas, pois traduz os objetivos da educação formal. As matrizes de referência utilizadas nas avaliações externas definem claramente o que se pretende avaliar através de um conjunto de descritores que elencam habilidades consideradas essenciais para o desenvolvimento cognitivo dos alunos ao longo de sua trajetória escolar, passíveis de serem mensuradas por meio de testes padronizados. Assim, é possível afirmar que, no que diz respeito ao planejamento

pedagógico, o currículo deve ser visto como a principal referência orientadora. Essa relação não apenas aponta a obrigatoriedade de alinhar a prática pedagógica às diretrizes curriculares, mas também propõe uma reflexão crítica sobre a eficácia e a relevância dessas avaliações na construção de um ensino que efetivamente atenda às necessidades dos estudantes e da sociedade (Borges; Almeida; Santana, 2021).

Dias Sobrinho (2002) utilizou em sua pesquisa de caráter qualitativo, utilizando uma abordagem analítica e crítica, com revisão bibliográfica e estudos de caso de instituições de ensino superior.

Observou-se algumas lacunas deixadas na pesquisa de Dias Sobrinho (2002), pois, embora o autor forneça uma análise abrangente, há lacunas relacionadas à atualização dos dados, uma vez que o contexto de avaliação educacional e a dinâmica do mercado mudaram desde a publicação em 2002. Notamos também a falta de estudos comparativos internacionais sobre o tema também é uma limitação, sugerindo a necessidade de uma investigação mais ampla que considere outras realidades globais.

Observamos por meio dos estudos de Sousa; Ferreira (2019) que a presente literatura ainda carece de uma integração efetiva entre as avaliações em larga escala e os métodos de avaliação prática na sala de aula, além da necessidade de políticas educacionais que considerem as realidades locais e ofereçam suporte estrutural e financeiro para efetivar melhorias nas práticas de avaliação e ensino.

Estudos feitos por Borges, Almeida e Santana (2021) em documentos oficiais do Inep, do e do CAEd demonstraram que, O Saeb reflete uma ideologia neoliberal, buscando uniformizar a avaliação de estudantes independentemente de seu contexto social, provocando um ranqueamento das instituições. Os instrumentos de avaliação utilizados tendem a estabelecer padrões de competência, mas desconsideram a complexidade das realidades sociais dos alunos. Observamos que, apesar de o Saeb ser uma ferramenta estruturante nas Políticas Públicas, ele possui limitações em sua capacidade de promover qualidade educativa equitativa. Pesquisa de caráter qualitativo com enfoque crítico participativo.

CONSIDERAÇÕES FINAIS

A partir da revisão da literatura observamos que os autores apresentaram uma ampla discussão ao abordarem a relação entre as avaliações em larga escala e as políticas públicas. O autor analisa como essas avaliações influenciam a formulação de políticas educacionais e as práticas pedagógicas nas escolas.

Os estudos ressaltam a necessidade de considerar as realidades locais ao interpretar os resultados das avaliações. O autor critica uma abordagem homogenizadora que ignora as especificidades de cada contexto educacional. Vimos ainda que os autores sugerem que as avaliações devem ser utilizadas não apenas para comparação, mas também como uma ferramenta para o desenvolvimento de políticas que atendam às necessidades reais das escolas e estudantes.

Há evidências de uma tensão entre a busca por qualidade educacional e a pressão do mercado, enfatizando como isso afeta a autonomia das universidades. O artigo de Dias Sobrinho (2002) ressalta que a avaliação pode tanto promover melhorias nas instituições quanto contribuir para um ambiente onde as metas de mercantilização prevalecem sobre os princípios éticos.

Os principais resultados dos estudos de Bonamino e Sousa (2012) do estudo indicam que as diferentes gerações de avaliação refletem mudanças significativas nas concepções de ensino e aprendizagem. A transição de uma avaliação punitiva para uma perspectiva mais formativa e inclusiva representa avanços no reconhecimento da diversidade dos alunos e na necessidade de uma avaliação que dialogue com o currículo escolar. Os autores concluem que a integração entre avaliação e currículo é vital para promover a qualidade da educação.

Observamos por meio dos estudos de Sousa e Ferreira (2019) uma análise contemporânea das avaliações em larga escala, com um olhar crítico sobre os métodos e práticas adotadas nas últimas décadas. Eles discutem as consequências das avaliações em larga escala sobre o currículo e as práticas pedagógicas. As autoras apresentam questões relacionadas à pressão por resultados que pode distorcer o propósito educativo das avaliações. Eles também enfatizam a importância de um uso ético e reflexivo dos dados gerados por essas avaliações. Vimos que elas defendem que as avaliações devem ser

acompanhadas de um diálogo efetivo com os educadores, promovendo uma reflexão sobre as práticas pedagógicas ao invés de uma simples busca por resultados.

Os estudos evidenciam, portanto, como as avaliações, tanto em larga escala quanto as formativas, podem se complementar e, quando utilizadas de maneira harmônica, podem contribuir para a melhoria da qualidade da educação.

Foi visto que, a avaliação em larga escala passou por três gerações distintas, cada uma refletindo diferentes contextos e objetivos educacionais. A primeira geração focou em medir resultados de aprendizado de forma quantitativa, visando apenas a comparação entre escolas e estudantes. Na segunda geração, houve uma ampliação do escopo, incorporando aspectos qualitativos e buscando entender os processos de ensino-aprendizagem. Por fim, a terceira geração enfatiza a utilização dos dados das avaliações para promover melhorias nos sistemas educacionais, priorizando a análise crítica e a reflexão sobre as práticas pedagógicas. Essa evolução mostra uma crescente valorização do uso das avaliações não apenas como ferramentas de mensuração, mas como meios para promover desenvolvimento e transformação educacional.

A avaliação em larga escala, embora tenha a intenção de medir e garantir a qualidade da educação, apresenta desafios e limitações que precisam ser considerados. Um dos principais aspectos críticos é a padronização dos testes, que tende a desconsiderar a diversidade cultural e socioeconômica dos alunos. Essas avaliações muitas vezes privilegiam um modelo único de ensino, desconsiderando as diferentes formas de aprendizagem e as particularidades de cada contexto.

Além disso, a pressão que essas avaliações exercem sobre escolas e professores pode levar a práticas pedagógicas que priorizam o ensinar para o teste, em detrimento de um aprendizado mais profundo e significativo. Isso pode resultar em um ambiente educacional que valoriza a memorização em vez do desenvolvimento de habilidades críticas e criativas.

Compreendemos que, os resultados das avaliações em larga escala são frequentemente utilizados como indicadores de desempenho que podem impactar negativamente as comunidades escolares, levando a estigmatizações e desigualdades sociais, uma vez que a responsabilidade pelo sucesso ou fracasso educacional é muitas

vezes atribuída às escolas e aos educadores, enquanto fatores externos, como a infraestrutura e o financiamento, são negligenciados.

Portanto, é fundamental que as avaliações em larga escala sejam repensadas e reformuladas, buscando uma abordagem mais justa e inclusiva, que considere as diversas realidades dos estudantes e que foque não apenas em resultados quantitativos, mas também na qualidade e na equidade da educação.

REFERÊNCIAS

BONAMINO, Alicia; SOUSA, Sandra Zákia. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. **Educação e Pesquisa**. São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012. Disponível em: <https://www.scielo.br>. Acesso em: 02 fev. 2025.

BORGES, A. P. N. R., ALMEIDA, S. P. N. de C. e, & SANTANA, K. C. L. (2021). **Avaliação Educacional: o Saeb, seus pressupostos, finalidades e repercussões**. Disponível em: *SciELO Preprints*. <https://doi.org>. Acesso em: 02 fev. 2025.

DIAS SOBRINHO, J. **Universidade e Avaliação: entre a ética e o mercado**. Florianópolis: Insular. 2002. 192 p. Disponível em: <https://pesquisa.bvsalud.org>. Acesso em: 20 nov. 2024.

NOGUEIRA, Arcielli Royer. Turnover de professores na região oeste do Paraná e suas influências nos resultados da avaliação em larga escala (Prova Brasil/SAEB) – 2015-2021. **Dissertação** (Mestrado Acadêmico) – Universidade Estadual do Oeste do Paraná, Centro de Educação, Programa de Pós-Graduação em Educação, Cascavel, 2021, 299 f.

SOUSA, Clarilza Prado de; FERREIRA, Sandra Lúcia. Avaliação de larga escala e da aprendizagem na escola: um diálogo necessário. **Psicologia da educação**, n. 48, p. 13-23, 2019. Disponível em: <http://educa.fcc.org.br/pdf/psie/n48/2175-3520-psie-48-13.pdf>. Acesso em: 22 fev. 2025.

Submissão: março de 2025. Aceite: abril de 2025. Publicação: julho de 2025.