ISSN: 2358-8829

AVALIAÇÕES ESTADUAIS DE ALFABETIZAÇÃO E SAEB: CONSTRUÇÃO E LIMITES DA COMPARABILIDADE DOS RESULTADOS

Rosilene Cândida de Araújo[1]

Mestre em Educação pela Universidade de Brasília - UnB, rosi.2712@gmail.com.

1 INTRODUÇÃO

A avaliação em larga escala ocupa lugar relevante na política educacional brasileira porque permite produzir informações sobre redes de ensino que dificilmente poderiam ser obtidas apenas por avaliações realizadas no cotidiano das escolas. Em um país marcado por profundas desigualdades sociais, territoriais e educacionais, conhecer diferenças de desempenho pode contribuir para orientar políticas e acompanhar a garantia do direito à educação. Reconhecer essa função, contudo, exige delimitar aquilo que os resultados avaliativos efetivamente permitem afirmar. Desempenho e aprendizagem não são categorias equivalentes. O desempenho corresponde ao que o estudante demonstra diante de determinadas tarefas e condições de avaliação; a aprendizagem constitui um processo mais amplo, desenvolvido ao longo do tempo e mediado por experiências pedagógicas, linguísticas, culturais e sociais. A avaliação pode produzir evidências relevantes sobre dimensões desse processo, mas não sua representação integral.

Essa distinção torna-se especialmente importante no campo da alfabetização. Em 2023, o Decreto nº 11.556 instituiu o Compromisso Nacional Criança Alfabetizada (CNCA), articulando União, estados, Distrito Federal e municípios em torno da garantia do direito à alfabetização ao final do 2º ano do ensino fundamental (BRASIL, 2023). A política encontrou uma realidade já constituída: diferentes unidades da Federação possuíam sistemas próprios de avaliação, com instrumentos e procedimentos distintos. Para que seus resultados participassem do acompanhamento nacional, tornou-se necessário construir condições de comparação com uma referência comum.

A solução adotada não consistiu em substituir as avaliações estaduais por uma única prova nacional. A Portaria Inep nº 351/2023 estabeleceu que os sistemas estaduais deveriam organizar-se de forma complementar ao Sistema de Avaliação da Educação Básica (Saeb), com compatibilidade metodológica quanto às medidas de desempenho, padronização de procedimentos e produção de subsídios para diagnóstico e políticas educacionais (INEP, 2023a). No mês seguinte, a Portaria nº 441 instituiu o Grupo de Trabalho em Comparabilidade

[1]Mestre em Educação Pela Universidade de Brasília (UnB), rosi.2712@gmail.com

entre Avaliações da Educação Básica, voltado à elaboração de diretrizes psicométricas para essa aproximação (INEP, 2023b). A comparabilidade tornou-se, assim, parte da arquitetura institucional do CNCA.

A escolha cumpria uma função de política pública: permitir uma leitura nacional da alfabetização sem extinguir os sistemas estaduais já existentes. Em fevereiro de 2024, os estados autorizaram o compartilhamento de resultados; os dados passaram a alimentar o Indicador Criança Alfabetizada e metas anuais até 2030 (INEP, 2026a). A operação técnica de relacionar instrumentos distintos passou, portanto, a sustentar a forma pela qual a política acompanha publicamente a alfabetização.

A análise documental realizada neste estudo mostra, entretanto, que a comparabilidade não surgiu pronta nem decorreu automaticamente da adoção da escala do Saeb. Em 2023, os estados permaneceram responsáveis por seus instrumentos e utilizaram itens previamente calibrados pelo Inep como elementos de ligação com a escala nacional. Nos anos seguintes, a documentação registrou heterogeneidades no uso da escrita, na composição dos testes, nas condições de aplicação e nos procedimentos de equalização. Essas diferenças motivaram análises, recomendações e maior padronização até 2025 (INEP, 2025a; INEP, 2025b; INEP, 2025c).

Esse percurso sustenta o primeiro achado do artigo: utilizar uma referência comum não basta, por si só, para produzir comparabilidade. Ela depende de decisões sobre construto, itens, desenho dos testes, aplicação, equalização, estimação e interpretação. O segundo achado decorre da natureza do resultado produzido: aquilo que se torna diretamente comparável é o desempenho, posteriormente convertido em estimativas de proficiência; inferências sobre aprendizagem exigem um passo interpretativo adicional. O terceiro é político: ao alimentar indicadores e metas, a comparabilidade passa a participar da forma pela qual o Estado torna a alfabetização visível e governável.

Diante desse percurso, pergunta-se: como foi construída, entre 2023 e 2025, a comparabilidade entre as avaliações estaduais de alfabetização e o Saeb e quais são os limites dessa medida comum para representar a aprendizagem das crianças? O objetivo é analisar a construção técnico-institucional da comparabilidade, identificar as operações utilizadas para relacionar avaliações distintas à referência nacional e discutir suas implicações pedagógicas e políticas. A posição que orienta o estudo pode ser sintetizada em uma premissa: avaliar é necessário; confundir desempenho mensurado com aprendizagem não é.

ISSN: 2358-8829

2 METODOLOGIA

A pesquisa adota abordagem qualitativa, documental, longitudinal e analítico-interpretativa, com recorte entre 2023 e 2025. O período compreende a instituição do CNCA, a criação das instâncias de articulação entre Saeb e sistemas estaduais e a progressiva normatização dos procedimentos de comparabilidade. A análise documental foi utilizada não como simples extração de informações, mas como leitura das fontes em relação a seu contexto, finalidade e função no objeto investigado, conforme orientação de Sá-Silva, Almeida e Guindani (2009).

O corpus primário foi composto por atos normativos, orientações e relatórios oficiais diretamente relacionados à institucionalização, operacionalização ou revisão da comparabilidade. A seleção obedeceu a três critérios: participação direta na regulamentação da articulação entre avaliações estaduais e Saeb; descrição de procedimentos de elaboração, aplicação, equalização ou interpretação das medidas; e registro de divergências, modificações ou recomendações de aperfeiçoamento no período. Documentos de divulgação foram utilizados apenas para contextualização.

Integram o corpus: o Decreto nº 11.556/2023; as Portarias Inep nº 351 e nº 441/2023; as Orientações Técnicas de 2023; o relatório de equalização dos itens de leitura da edição de 2023; o relatório das curvas de informação das edições de 2023 e 2024; o Parecer do Comitê Consultivo de Especialistas de 2025; a Instrução Normativa Inep nº 2/2025; e os relatórios técnico e pedagógico da Avaliação da Alfabetização 2025.

Para sistematização da evidência, cada documento foi registrado em matriz analítica com onze dimensões: função no processo; concepção de alfabetização; forma de produção do desempenho; procedimento de construção da comparabilidade; mudanças e limites reconhecidos; atores e vozes presentes; lugar atribuído ao professor alfabetizador; lugar atribuído à criança; uso político dos resultados; evidência documental; e interpretação crítica. A separação entre evidência e interpretação foi mantida para evitar que justificativas institucionais fossem tomadas como comprovação de eficácia ou que hipóteses da autora fossem apresentadas como conteúdo literal das fontes.

A análise foi conduzida em duas dimensões articuladas. A dimensão técnico-institucional reconstruiu como resultados de instrumentos diferentes foram relacionados à referência nacional, considerando construto, matriz, itens comuns, resposta construída, desenho dos testes, aplicação, equalização e estimação das proficiências. A dimensão político-pedagógica examinou os usos atribuídos aos resultados, os atores formalmente envolvidos e o lugar ocupado por professores e crianças na representação institucional da alfabetização.

ISSN: 2358-8829

Um cuidado específico foi dedicado à distinção entre desempenho e aprendizagem. Preservou-se, sempre que possível, a terminologia dos documentos. A referência oficial a medidas de desempenho não foi reescrita como medida direta de aprendizagem. Considerou-se o percurso entre respostas aos itens, modelagem, proficiência e interpretação, tratando qualquer afirmação sobre aprendizagem como inferência que precisa ser proporcional às evidências produzidas pelo instrumento.

O desenho não realiza reanálise psicométrica dos bancos de itens nem permite afirmar efeitos causais das avaliações sobre currículos ou práticas docentes, pois não houve observação ou coleta direta nas escolas. Também não se interpreta a ausência de representação formal de professores alfabetizadores em determinados colegiados como ausência desses profissionais em toda a política. A inferência limita-se às instâncias e formas de participação registradas no corpus. Referências ao Saeb de 2027 têm caráter prospectivo e não integram os resultados empíricos.

3 RESULTADOS E DISCUSSÃO

3.1 A comparabilidade como construção técnico-institucional

A articulação entre avaliações estaduais e Saeb foi institucionalizada em 2023 como componente do acompanhamento do CNCA. A Portaria nº 351 definiu a compatibilidade metodológica das medidas de desempenho e a Portaria nº 441 criou um grupo técnico especificamente voltado à produção de diretrizes para comparabilidade (INEP, 2023a; 2023b). Não se tratou, portanto, de aproximar informalmente avaliações já existentes, mas de criar uma arquitetura destinada a produzir uma referência comum a partir de instrumentos administrados por sistemas distintos.

O arranjo preservou parte da autonomia estadual. Na edição de 2023, cada estado permaneceu responsável pela elaboração de seus itens, aplicação e análise, enquanto o Inep disponibilizou 20 itens calibrados na escala de alfabetização do Saeb. Esses itens funcionaram como elementos de ligação, e os estados tiveram independência para compor suas provas dentro de orientações comuns (BORGATTO; ANDRADE, 2024). A solução não foi aplicar uma única prova, mas tornar interpretáveis em uma referência compartilhada desempenhos produzidos por testes parcialmente diferentes.

Essa opção precisa ser lida também como política. Ball (1993) propõe compreender a política como texto e discurso, enquanto Mainardes (2006) chama atenção para os atores e contextos que participam de sua produção e reinterpretação. No caso analisado, a manutenção das avaliações estaduais combinada à referência nacional materializa uma negociação entre diversidade federativa e necessidade de monitoramento comum. Quando os resultados passam a

alimentar o Indicador Criança Alfabetizada e metas anuais, a operação estatística adquire consequência institucional: aquilo que é tornado comparável passa a ser utilizado para acompanhar redes e orientar a ação pública (INEP, 2026a).

3.2 Heterogeneidades e ampliação da padronização

A documentação posterior mostra que compartilhar uma escala não encerrou o problema. O relatório de 2025 registra que, em 2023, 10 dos 23 estados participantes não utilizaram itens de resposta construída; em 2024, eram 7 entre 26. Para comparar as três edições, a análise longitudinal precisou considerar inicialmente apenas os itens de leitura de múltipla escolha (INEP, 2025c). A decisão revela um limite concreto: instrumentos podem ser colocados na mesma escala e, ainda assim, diferir na forma como representam o construto e distribuem informação ao longo dessa escala.

A escrita é o exemplo mais claro. As orientações de 2023 já trabalhavam com Apropriação do Sistema de Escrita Alfabética, Leitura e Produção Textual, mas esses componentes não apareceram de forma uniforme nos testes. Somente em 2025 todos os estados e o Distrito Federal incorporaram itens de escrita. O relatório técnico considera essa mudança fundamental para melhorar a comparabilidade porque a interpretação da escala envolve leitura e escrita (INEP, 2025c). A trajetória mostra que comparabilidade psicométrica e alinhamento do construto são exigências relacionadas, mas não equivalentes.

Em 2025, a equalização foi reforçada com 21 itens calibrados pelo Inep, além de regras mais detalhadas de montagem, aplicação, cálculo e envio dos dados, consolidadas pela Instrução Normativa nº 2/2025 (INEP, 2025d). O movimento não autoriza afirmar que a comparabilidade inexistia em 2023; tampouco permite tratá-la como resolvida desde o início. A evidência aponta para uma comparabilidade instituída, colocada à prova e progressivamente aperfeiçoada.

Essa leitura dialoga com Kane (2013), para quem não se valida um escore de maneira abstrata, mas as interpretações e usos sustentados por uma cadeia de inferências. Quanto mais ampla a conclusão pretendida, maior a exigência de evidências para seus pressupostos. Aplicada a este estudo, a formulação impede que a presença de uma métrica comum seja tomada como garantia suficiente: a interpretação comparativa depende também do construto, do desenho, da aplicação e dos procedimentos de estimação.

O próprio arranjo institucional reconheceu essa necessidade. O Comitê Consultivo criado para examinar divergências entre Saeb e avaliações estaduais recomendou maior padronização da equalização, pesos amostrais, matrizes, escrita, correção e condições de aplicação, além de maior transparência sobre rotinas de cálculo e amostragem do Saeb (INEP,

2025b). A crítica à comparabilidade automática, portanto, não é externa aos documentos; emerge das dificuldades reconhecidas pelo próprio sistema.

3.3 O que se torna comparável: desempenho não é aprendizagem

A expressão medidas de desempenho, utilizada pela Portaria nº 351/2023, é conceitualmente precisa e deve ser preservada. A criança responde a tarefas; as respostas são submetidas a procedimentos de calibração e modelagem; estima-se uma proficiência; e o resultado é interpretado em uma escala. Há uma cadeia entre aquilo que foi observado e aquilo que se pretende afirmar. A aprendizagem pode ser inferida a partir dessas evidências, mas não está diretamente depositada no escore.

Kane (2013) ajuda a situar o problema: afirmar que um estudante demonstrou determinado desempenho é menos abrangente do que concluir que a medida representa seu processo de aprendizagem. No campo da alfabetização, a distância é ainda mais importante. Soares (2004) preserva a especificidade da apropriação do sistema alfabético sem separá-la das práticas sociais de leitura e escrita. Sua contribuição amplia o objeto pedagógico que a avaliação tenta representar: alfabetizar não é apenas acertar tarefas, embora determinadas tarefas possam produzir evidências relevantes sobre aspectos da alfabetização.

O relatório pedagógico de 2025 reconhece essa amplitude ao organizar o construto nos eixos Apropriação do Sistema de Escrita Alfabética, Leitura e Produção de Textos (INEP, 2025e). A inclusão da escrita melhora a representação avaliativa, mas não altera a natureza do instrumento: toda prova seleciona habilidades, tarefas e condições de resposta. Soares e Kane permitem, assim, ler o mesmo problema por ângulos complementares: a primeira dimensiona a complexidade do objeto; o segundo obriga a limitar a inferência ao que a medida consegue sustentar.

Essa precaução também se aproxima de Oliveira, Pereira Junior e Horta Neto (2025), que questionam a redução da qualidade educacional aos resultados de testes cognitivos. Qualidade e aprendizagem não são conceitos equivalentes, mas a advertência é pertinente: um indicador pode ser tecnicamente relevante sem constituir representação suficiente do fenômeno mais amplo ao qual é associado. Resultados baixos podem revelar desigualdades e orientar ação pública; isso não exige transformá-los em definição total da aprendizagem. A medida informa, e a inferência precisa permanecer proporcional à evidência.

3.4 Do desempenho ao indicador: usos, autoridade e saber pedagógico

A transformação das proficiências em percentuais e metas modifica a função social dos resultados. Bauer, Alavarse e Oliveira (2015) mostram que o debate sobre avaliações em larga escala não comporta a oposição simples entre aceitá-las e recusá-las: elas podem orientar

ISSN: 2358-8829

políticas e, simultaneamente, exigir crítica quanto aos usos. Bonamino e Sousa (2012) também demonstram que as consequências associadas aos resultados alteram o lugar ocupado pelas avaliações e podem produzir repercussões sobre redes e currículos. O CNCA não deve ser automaticamente enquadrado nos modelos de responsabilização discutidos por essas autoras; interessa aqui a proposição mais geral de que resultados não têm efeitos independentes dos usos institucionais que recebem.

Quando a porcentagem de estudantes que alcança determinado padrão passa a representar o avanço de uma rede em direção a uma meta nacional, o desempenho avaliado adquire circulação política. Ozga (2009) mostra como a produção de dados pode deixar de funcionar apenas como descrição e participar dos próprios modos de governar sistemas educacionais. No caso analisado, resultados comparáveis ampliam a capacidade de acompanhar desigualdades, mas também conferem maior visibilidade institucional às dimensões da alfabetização que conseguem ser traduzidas em escala, percentual e meta.

A composição das instâncias técnicas ajuda a observar quais formas de conhecimento recebem autoridade nessa etapa. A Comissão de Articulação reúne Inep, Consed, Undime e Consec; o GT de Comparabilidade trabalhou com equipes técnicas e estatísticos indicados pelas secretarias estaduais (INEP, 2026a). Essa composição é coerente com uma tarefa que exige conhecimento especializado. A questão não é contestar a psicometria, mas reconhecer que ela responde a perguntas diferentes daquelas enfrentadas por quem acompanha diariamente as crianças em processo de alfabetização.

A Pesquisa Alfabetiza Brasil mostra que professoras alfabetizadoras participaram da definição pedagógica do padrão nacional: 251 docentes de 206 municípios analisaram tarefas associadas ao que se esperava de uma criança alfabetizada (INEP, 2023d). Entretanto, não se identificou representação formal específica desse grupo nas instâncias técnicas examinadas de comparabilidade. Surge, então, uma distinção relevante entre participar da construção do padrão pedagógico e participar das decisões que definem como instrumentos diferentes serão relacionados, equalizados e convertidos em resultados comparáveis.

A pergunta que decorre desse achado é menos sobre presença e mais sobre autoridade interpretativa: em que momento o conhecimento de quem alfabetiza retorna ao processo de leitura dos resultados? O professor observa hipóteses de escrita, estratégias de leitura, respostas à mediação e mudanças ao longo do tempo; esse conhecimento não substitui a medida psicométrica, mas tampouco pode ser substituído por ela. A mesma cautela vale para a criança: para existir como informação de sistema, suas respostas são agregadas em escore, proficiência e

percentual. O dado é uma representação necessária, não o próprio sujeito nem seu processo de aprendizagem.

Ball (1993) e Mainardes (2006) permitem formular essa questão sem supor interesses ocultos. Políticas produzem formas autorizadas de nomear e interpretar problemas. Quando escalas, percentuais e metas se tornam linguagem central da política de alfabetização, a pergunta crítica é quais conhecimentos ganham maior capacidade de definir publicamente o que conta como evidência e até onde essa autoridade deve chegar.

3.5 Comparar também expõe a referência: aprendizagem institucional e Saeb 2027

A análise do corpus revelou um movimento que amplia a interpretação inicial. A aproximação entre sistemas estaduais e Saeb não produziu apenas exigências de adequação dos estados. O Parecer do Comitê Consultivo recomenda divulgação das rotinas de cálculo e das informações amostrais do Saeb, um modelo conceitual articulado entre MEC e Inep, maior convergência entre matrizes, padronização da aplicação e oferta regular de itens calibrados (INEP, 2025b). A referência nacional permanece sendo o Saeb, mas o documento não reduz toda divergência a um defeito do instrumento estadual.

Esse resultado permite propor uma leitura autoral da comparabilidade como processo de aprendizagem institucional recíproca. O Saeb fornece a métrica de referência; os sistemas estaduais, ao operacionalizarem avaliações em contextos diversos, também produzem evidências sobre dificuldades e escolhas que precisam ser revistas. Se duas avaliações apresentam resultados divergentes, a questão científica não é decidir previamente qual expressa a verdadeira aprendizagem, mas examinar construto, itens, aplicação, amostragem, equalização, precisão e interpretação. A exigência de Kane (2013) vale para ambos: as afirmações precisam ser sustentadas pelas evidências disponíveis.

A edição do Saeb de 2027 poderá constituir um ponto de observação dos desdobramentos desse ciclo, mas não integra o corpus e não é tratada como resultado. A questão deixada pelo estudo é se a maior padronização produzirá apenas números mais próximos ou também interpretações mais defensáveis. Convergência estatística pode indicar maior alinhamento técnico; não demonstra, sozinha, que a avaliação passou a representar com maior abrangência a alfabetização.

A contribuição central pode ser formulada nesse limite: aperfeiçoar a comparabilidade não deve significar apenas aproximar números, mas tornar mais defensável aquilo que esses números nos autorizam a dizer. Se a finalidade da política é garantir o direito à alfabetização, a precisão estatística é necessária, mas não suficiente. A pergunta pedagógica permanece: o que esse resultado permite conhecer sobre a criança e o que continua fora de seu alcance?

ISSN: 2358-8829

4 CONSIDERAÇÕES FINAIS

Este estudo analisou a construção da comparabilidade entre avaliações estaduais de alfabetização e Saeb entre 2023 e 2025. A documentação mostra que a comparabilidade foi instituída desde 2023 por meio de itens calibrados, equalização e orientações comuns, mas não se estabilizou naquele primeiro desenho. Diferenças no construto efetivamente representado, na presença da escrita, na aplicação e nos procedimentos técnicos tornaram-se objeto de revisão e levaram a maior padronização em 2025.

O principal limite conceitual identificado diz respeito ao que efetivamente se compara. Os instrumentos produzem evidências de desempenho diante de tarefas determinadas e as transformam em estimativas de proficiência. A aprendizagem é mais ampla e não pode ser tratada como sinônimo do resultado. Essa distinção não enfraquece a avaliação; ao contrário, delimita a extensão legítima das inferências e protege a medida de usos que extrapolem aquilo que ela consegue sustentar.

Quando esses resultados alimentam o Indicador Criança Alfabetizada e metas nacionais, a comparabilidade deixa de ser apenas questão de mensuração. Ela participa da maneira pela qual o Estado representa e acompanha a alfabetização. Essa capacidade pode revelar desigualdades e orientar políticas, mas exige cautela para que aquilo que é mais facilmente mensurável não adquira autoridade para definir sozinho aquilo que pedagogicamente importa.

O estudo também mostrou que a relação com o Saeb não é inteiramente unilateral. As divergências produziram recomendações que alcançam o próprio arranjo nacional, permitindo compreender a comparabilidade como espaço de aprendizagem institucional. Essa leitura poderá ser retomada em 2027, sem pressupor mudanças ainda não regulamentadas, para observar se o aperfeiçoamento técnico resultará apenas em maior convergência estatística ou também em interpretações mais defensáveis.

A contribuição do trabalho está, portanto, em deslocar a pergunta restrita sobre se os resultados são comparáveis para três questões articuladas: o que precisa acontecer para que sejam comparáveis; o que exatamente se torna comparável; e até onde essa comparação autoriza falar sobre alfabetização. A posição defendida é favorável à avaliação em larga escala justamente porque reconhece sua utilidade pública. Avaliar é necessário; transformar desempenho mensurado em sinônimo de aprendizagem não é. Os números devem servir à garantia do direito à alfabetização, sem reduzir a criança nem o trabalho pedagógico ao que o instrumento consegue contar.

ISSN: 2358-8829

REFERÊNCIAS

BALL, Stephen J. What is policy? Texts, trajectories and toolboxes. Discourse: Studies in the Cultural Politics of Education, v. 13, n. 2, p. 10-17, 1993. DOI: 10.1080/0159630930130203.

BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela de. Avaliações em larga escala: uma sistematização do debate. Educação e Pesquisa, São Paulo, v. 41, n. esp., p. 1367-1384, 2015. DOI: 10.1590/S1517-9702201508144607.

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, 2012. DOI: 10.1590/S1517-97022012005000006.

BORGATTO, Adriano Ferreti; ANDRADE, Dalton Francisco de. Relatório técnico detalhado: análise das equalizações dos itens de leitura realizadas pelas redes estaduais com o Saeb para os testes de alfabetização. Brasília, DF: Inep, 2024.

BRASIL. Decreto nº 11.556, de 12 de junho de 2023. Institui o Compromisso Nacional Criança Alfabetizada. Brasília, DF: Presidência da República, 2023.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 351, de 4 de agosto de 2023. Estabelece diretrizes e orientações para a organização complementar entre os sistemas estaduais de avaliação e o Saeb. Brasília, DF: Inep, 2023a.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 441, de 25 de setembro de 2023. Institui o Grupo de Trabalho em Comparabilidade entre Avaliações da Educação Básica. Brasília, DF: Inep, 2023b.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Orientações técnicas 2023: avaliação da alfabetização. Brasília, DF: Inep, 2023c.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Pesquisa Alfabetiza Brasil: diretrizes para uma política nacional de avaliação da alfabetização das crianças. Brasília, DF: Inep, 2023d.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Relatório técnico: análise das curvas de informação das avaliações do sistema estadual - edições 2023 e 2024. Brasília, DF: Inep, 2025a.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Parecer do Comitê Consultivo de Especialistas para Aperfeiçoamento dos Processos de Avaliação da Educação Básica. Brasília, DF: Inep/MEC, 2025b.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Relatório da Avaliação da Alfabetização 2025: desenho do teste e procedimentos de análise dos itens. Brasília, DF: Inep, 2025c.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Instrução Normativa nº 2, de 26 de novembro de 2025. Dispõe sobre diretrizes e orientações para a organização dos sistemas estaduais e distrital de avaliação de forma complementar ao Saeb. Diário Oficial da União, Brasília, DF, 28 nov. 2025. 2025d.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Relatório Pedagógico da Avaliação da Alfabetização 2025. Brasília, DF: Inep, 2025e.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Avaliação da Alfabetização: articulação entre os sistemas estaduais e o Saeb; Indicador Criança Alfabetizada e metas até 2030. Brasília, DF: Inep, 2026a.

KANE, Michael T. Validating the interpretations and uses of test scores. Journal of Educational Measurement, v. 50, n. 1, p. 1-73, 2013. DOI: 10.1111/jedm.12000.

MAINARDES, Jefferson. Abordagem do ciclo de políticas: uma contribuição para a análise de políticas educacionais. Educação & Sociedade, Campinas, v. 27, n. 94, p. 47-69, 2006. DOI: 10.1590/S0101-73302006000100003.

OLIVEIRA, Dalila Andrade; PEREIRA JUNIOR, Edmilson Antonio; HORTA NETO, João Luiz. Qualidade da educação básica: para além dos testes cognitivos em larga escala. Cadernos de Pesquisa, v. 55, e11295, 2025. DOI: 10.1590/1980531411295.

OZGA, Jenny. Governing education through data in England: from regulation to self-evaluation. Journal of Education Policy, v. 24, n. 2, p. 149-162, 2009. DOI: 10.1080/02680930902733121.

SÁ-SILVA, Jackson Ronie; ALMEIDA, Cristóvão Domingos de; GUINDANI, Joel Felipe. Pesquisa documental: pistas teóricas e metodológicas. Revista Brasileira de História & Ciências Sociais, v. 1, n. 1, 2009. DOI: 10.63595/rbhcs.v1i1.10351.

SOARES, Magda. Letramento e alfabetização: as muitas facetas. Revista Brasileira de Educação, n. 25, p. 5-17, 2004. DOI: 10.1590/S1413-24782004000100002.