ALÉM DA PROFICIÊNCIA: um estudo sobre trajetórias individuais e métricas sistêmicas

CAIO GUILHERME VASCONCELOS ALVES [1]
https://orcid.org/0009-0005-7852-1986
aluno.caiogva@ensin-e.edu.br

SILAS PEREIRA DE OLIVEIRA [2]
https://orcid.org/0009-0006-2416-4255
aluno.silaspo@ensin-e.edu.br

SOPHIA CRISTINE SOARES MELLO [3]
https://orcid.org/0009-0002-1793-8751
aluna.sophiacsm@ensin-e.edu.br

TATIANE GONÇALVES MORAES [4]
https://orcid.org/0000-0002-5910-2381
profa.tatiane.moraes@ensin-e.edu.br

RESUMO
Este artigo analisa como registros qualitativos de aprendizagem produzidos por uma professora de Matemática do 6º ano no cotidiano escolar dialogam — ou não — com os indicadores numéricos das avaliações externas em larga escala no Brasil. A pesquisa, desenvolvida como estudo de caso, acompanhou ao longo de um ano letivo uma estudante do 6º ano e comparou sua trajetória individual aos dados gerados pelas escalas de proficiência do Saeb. Os registros docentes revelaram avanços cognitivos, afetivos e socioemocionais, fortemente influenciados por intervenções pedagógicas, feedbacks contínuos e elogios específicos, elementos que impulsionaram o engajamento e a autonomia da estudante. Em contraste, as avaliações externas produzem métricas agregadas voltadas ao monitoramento de redes, incapazes de capturar nuances relacionais e processuais do aprender. Os resultados mostram que avaliação interna e externa cumprem funções distintas e complementares: a primeira ilumina processos individuais; a segunda identifica tendências coletivas. O estudo conclui que o sentido pedagógico dos números depende da articulação responsável entre essas duas camadas de evidência, reconhecendo que trajetórias reais de aprendizagem ultrapassam aquilo que a métrica estatística é capaz de mostrar.

Palavras-chave: Avaliação em larga escala. Escala de proficiência. Uso de resultados.

ABSTRACT
His article analyzes how qualitative learning records produced by a teacher in everyday classroom practice do—or do not—dialogue with the numerical indicators generated by large-scale external assessments in Brazil. Developed as a case study, the research followed a 6th-grade student over the course of one school year and compared her individual learning

1 Faculdade Ensine, Juiz de Fora, Minas Gerais (MG), Brasil, Estudante do curso de Matemática e IA.
2 Faculdade Ensine, Juiz de Fora, Minas Gerais (MG), Brasil, Estudante do curso de Matemática e IA.
3 Faculdade Ensine, Juiz de Fora, Minas Gerais (MG), Brasil, Estudante do curso de Matemática e IA.
4 Faculdade Ensine, Juiz de Fora, Minas Gerais (MG), Brasil, Professora Doutoranda em Educação pela UFJF e orientadora do projeto de extensão da Faculdade Ensine.

trajectory with the data produced by Saeb proficiency scales. The teacher's records revealed cognitive, affective, and socioemotional progress, strongly influenced by pedagogical interventions, continuous feedback, and specific praise—elements that fostered the student's engagement and autonomy. In contrast, external assessments generate aggregated metrics designed to monitor educational systems, which cannot capture relational and processual nuances of learning. The results show that internal and external assessment fulfill distinct yet complementary functions: the former illuminates individual processes, while the latter identifies collective trends. The study concludes that the pedagogical meaning of numbers depends on the responsible articulation of these two layers of evidence, acknowledging that real learning trajectories extend beyond what statistical metrics are able to reveal.

Keywords: Large-scale assessment. Proficiency scale. Use of results.

RESUMEN
Este artículo analiza cómo los registros cualitativos de aprendizaje producidos por la profesora en el cotidiano escolar dialogan —o no— con los indicadores numéricos generados por las evaluaciones externas a gran escala en Brasil. Desarrollada como un estudio de caso, la investigación acompañó durante un año lectivo a una estudiante de 6.º grado y comparó su trayectoria individual con los datos producidos por las escalas de proficiencia del Saeb. Los registros docentes revelaron avances cognitivos, afectivos y socioemocionales, fuertemente influenciados por intervenciones pedagógicas, retroalimentación continua y elogios específicos, elementos que impulsaron el compromiso y la autonomía de la alumna. En contraste, las evaluaciones externas generan métricas agregadas destinadas a monitorear las redes educativas, incapaces de capturar los matices relacionales y procesuales del aprender. Los resultados muestran que la evaluación interna y la externa cumplen funciones distintas y complementarias: la primera ilumina procesos individuales; la segunda identifica tendencias colectivas. El estudio concluye que el sentido pedagógico de los números depende de la articulación responsable entre estas dos capas de evidencia, reconociendo que las trayectorias reales de aprendizaje van más allá de lo que las métricas estadísticas pueden mostrar.

Palabras clave: Evaluación a gran escala. Escala de proficiencia. Uso de resultados.

1. INTRODUÇÃO

Em tempos em que resultados de avaliações em larga escala se transformam rapidamente em manchetes, rankings e metas, é fundamental lembrar que toda métrica carrega escolhas metodológicas, intenções políticas e significados que vão muito além do número divulgado. No Brasil, o Sistema de Avaliação da Educação Básica (Saeb) e o Índice de Desenvolvimento da Educação Básica (Ideb) tornaram-se os principais referenciais numéricos da qualidade da educação básica, orientando políticas públicas e moldando discursos sobre eficiência e desempenho das redes de ensino (INEP, 2016; BRASIL, 2005, 2013, 2020). O Ideb sintetiza duas dimensões distintas — fluxo escolar e aprendizagem — esta última estimada pelas escalas de proficiência do Saeb, construídas a partir de modelos

da Teoria de Resposta ao Item (TRI).

Nesse modelo, cada item funciona como um instrumento de medida com parâmetros próprios, permitindo estimar a proficiência do estudante em uma escala contínua de habilidade latente. Assim, a proficiência não reflete apenas quantos itens foram acertados, mas o nível de competências demonstrado diante de tarefas de complexidade graduada (Klein, 2011). Esse autor compara a escala de proficiência a uma escala termométrica: sua origem e unidade são arbitradas, mas é a interpretação pedagógica de cada ponto que lhe confere sentido. Da mesma forma, interpretar que um grupo de estudantes se encontra, por exemplo, em determinado patamar de proficiência significa inferir probabilisticamente quais habilidades já dominam e quais ainda não foram consolidadas.

A escala, portanto, é uma linguagem técnica do desenvolvimento de competências — linguagem que, nem sempre, é apropriada de maneira adequada pelas escolas e redes ao transformá-la apenas em metas a cumprir. Quando médias de proficiência ou índices sintéticos são lidos exclusivamente como marcadores de sucesso ou fracasso, corre-se o risco de reduzir a avaliação a um instrumento de responsabilização, esvaziando sua potência de diagnóstico e compreensão dos processos de aprendizagem. Em contraste, registros cotidianos produzidos pelo professor — como descrever que uma estudante tímida, organizada e retraída passou a interagir mais, realizar todas as atividades e demonstrar avanços graduais, ainda que com dificuldades de compreensão — revelam dimensões subjetivas, interpretativas e relacionais que nenhuma média de proficiência isolada é capaz de capturar.

É justamente nessa tensão entre o estudante que “aparece” como um ponto em uma escala estatística e o estudante concreto, que vive processos, hesitações, avanços e vínculos na sala de aula, que este artigo se insere. Logo, este estudo busca compreender em que condições os números das avaliações externas podem dialogar com os registros qualitativos da prática docente, e como esse diálogo pode produzir interpretações mais ricas e fundamentadas sobre o aprender.

Portanto, o objetivo central desta pesquisa é discutir a natureza e o uso das métricas produzidas pelas avaliações em larga escala, analisando seus limites e possibilidades quando comparadas às narrativas qualitativas de aprendizagem construídas pela professora de Matemática no acompanhamento longitudinal de uma estudante do 6º ano. Pretende-se, assim, explorar como a métrica ganha sentido quando articulada à observação pedagógica, e como essa articulação pode contribuir para uma compreensão mais ampla das

desigualdades, dos avanços e dos desafios da aprendizagem no contexto escolar.

2. METODOLOGIA DO TRABALHO

Este estudo caracteriza-se como uma pesquisa qualitativa, de natureza descritivo-analítica e abordagem exploratória (Lüdke; André, 1986; Bogdan; Biklen, 1994). O foco central é analisar como registros qualitativos de aprendizagem — produzidos no contexto da avaliação interna — dialogam com os resultados padronizados das avaliações externas, cujo objetivo é avaliar redes e sistemas educacionais e não o desempenho individual de cada estudante (INEP, 2016; Brasil, 2005; Bonamino; Sousa, 2012). Assim, a pesquisa busca compreender o distanciamento e as possíveis aproximações entre essas duas formas de produzir evidências sobre o processo de aprender.

A investigação assumiu o formato de estudo de caso (Stake, 1995; Yin, 2015), com acompanhamento longitudinal de uma estudante do 6º ano do ensino fundamental de uma escola municipal ao longo de um ano letivo. A seleção da aluna foi intencional, fundamentada em critérios teóricos e metodológicos. Três características orientaram essa escolha: ser uma estudante laudada, apresentar comportamento timidamente retraído e demonstrar forte organização e regularidade no uso de materiais e na execução de tarefas. Esses critérios não se configuram como limitações, mas como elementos que qualificam a escolha e ampliam a riqueza analítica do caso. Do ponto de vista do desenvolvimento cognitivo, estudantes laudados frequentemente apresentam perfis heterogêneos de aprendizagem, com potenciais específicos e dificuldades localizadas, tornando visíveis processos que, em estudantes sem laudo, podem permanecer menos evidentes (APA, 2014).

À luz da perspectiva sociocultural de Vygotski (1995), tais perfis tornam mais clara a dinâmica da zona de desenvolvimento proximal, especialmente quando mediada por intervenções pedagógicas planejadas. A timidez e o retraimento, por sua vez, permitem observar como dimensões socioemocionais (aquelas que afetam a forma como o estudante interage, participa, se arrisca cognitivamente e constrói significado) são invisibilizadas em avaliações externas, cuja padronização não captura aspectos afetivo-motivacionais (Eisenberg et al., 2004; Buck, 2019). A presença desses elementos confere ao caso uma profundidade analítica importante para o problema investigado: entender o que os números mostram e o que deixam de mostrar.

A organização e o cumprimento sistemático das tarefas, por fim, funcionam como um critério metodológico de controle. Estudantes com esses perfis reduzem ruídos típicos de

estudos longitudinais (como irregularidades no acompanhamento, tarefas não realizadas ou grandes variações comportamentais), permitindo analisar com maior precisão o efeito das intervenções pedagógicas e o progresso efetivamente observado (Snow; Corno, 1998). O acompanhamento foi realizado pelos estudantes do projeto de extensão curricular de matemática da faculdade Ensine. Ao longo do ano, foram planejadas intervenções pontuais em conteúdos nos quais a estudante demonstrava maior dificuldade. Após cada sequência de atividades, a professora da turma registrava aspectos qualitativos da aprendizagem, tais como estratégias de resolução, compreensão conceitual, autonomia, participação, persistência e mudanças observadas em relação a momentos anteriores. Esses registros constituíram o corpus qualitativo principal da análise.

A pesquisa desenvolveu-se em duas etapas. Primeiro, os registros foram examinados por meio da análise de conteúdo (Bardin, 2011), originando categorias que representaram dimensões cognitivas, afetivo-relacionais e organizacionais. Em seguida, essas evidências foram comparadas ao tipo de informação produzido pelas avaliações externas (escalas de proficiência, níveis de desempenho e médias da rede) não com o propósito de estabelecer equivalências diretas, mas de analisar como a trajetória singular da aluna pode (ou não) ser representada pelas métricas agregadas que caracterizam o Saeb e, posteriormente, se evidenciam no Ideb (INEP, 2016; Klein, 2011; Pontes, 2017).

O estudo integra um projeto de extensão universitária vinculado ao curso de Matemática e Inteligência Artificial da Faculdade Ensine, conduzido por três estudantes de graduação e pela professora orientadora. Os estudantes atuaram na análise dos registros, na sistematização dos dados, na articulação teórica e na discussão dos achados, reforçando a aproximação entre universidade e escola e promovendo uma formação que integra análise de dados educacionais e compreensão qualitativa dos processos de aprendizagem.

3. FUNDAMENTAÇÃO TEÓRICA

As avaliações educacionais em larga escala tornaram-se pilares das políticas públicas nas últimas décadas, oferecendo indicadores capazes de monitorar sistemas e orientar decisões governamentais. No Brasil, o Saeb e o Ideb desempenham esse papel ao produzir resultados padronizados que traduzem, em números, tendências coletivas de desempenho escolar (INEP, 2016; Brasil, 2005). É importante salientar que as Avaliações Externas têm finalidade sistêmica, e não individual, o que implica limites importantes para sua interpretação pedagógica (Bonamino; Sousa, 2012).

A adoção da Teoria de Resposta ao Item (TRI) no Saeb, a partir de 1995, permitiu estimar a proficiência dos estudantes em uma escala contínua de habilidade latente. Cada item funciona como um instrumento de medida, caracterizado por parâmetros que descrevem sua dificuldade, discriminação e possibilidade de acerto ao acaso (Baker, 2001). Assim, a proficiência não corresponde à contagem de acertos, mas a uma medida probabilística construída segundo critérios psicométricos rigorosos. Klein (2011) compara sua lógica à de uma escala termométrica: valores arbitrados ganham significado apenas quando interpretados à luz das habilidades que representam.

Essa natureza técnica evidencia que o uso das métricas requer cautela. Como ressaltam Messick (1989) e os Standards (AERA; APA; NCME, 2014), a validade de uma medida está relacionada não só ao cálculo da pontuação, mas também às interpretações e consequências produzidas por ela. Resultados agregados, portanto, não capturam trajetórias individuais, nem aspectos socioemocionais, motivacionais ou relacionais do aprender — dimensões centrais na prática pedagógica cotidiana.

Em contraste, a avaliação interna, especialmente quando orientada por princípios formativos, produz evidências contínuas e individualizadas sobre o processo de aprendizagem. Black e Wiliam (1998) demonstram que o feedback de qualidade tem um dos maiores impactos conhecidos sobre o desempenho dos estudantes. Brookhart (2011) argumenta que registros qualitativos (observações, descrições de estratégias, momentos de dúvida, modos de participação) revelam elementos que não emergem em métricas padronizadas, mas que são fundamentais para orientar o ensino.

Para interpretar esse tipo de registro, teorias do desenvolvimento e da aprendizagem oferecem suporte essencial. Na perspectiva sociocultural de Vygotski (1995), aprender é um processo mediado, que se realiza na interação entre o estudante e o professor dentro da Zona de Desenvolvimento Proximal. Perrenoud (1999) destaca que as competências se constroem gradualmente, à medida que os estudantes enfrentam desafios e reorganizam esquemas de ação. A Taxonomia de Bloom revisada (Anderson; Krathwohl, 2001) permite situar evidências qualitativas em níveis cognitivos distintos, orientando a análise pedagógica.

Esses referenciais ajudam a compreender trajetórias como a da estudante acompanhada neste estudo, cuja timidez, organização e necessidade de apoio docente revelam nuances que não aparecem em indicadores sistêmicos. Tais características influenciam a forma como ela participa, arrisca, persiste e constrói significado — dimensões essenciais para compreender seu desenvolvimento na Matemática.

Dessa forma, a literatura contemporânea aponta para a necessidade de articular evidências de diferentes naturezas. Avaliações externas oferecem um panorama amplo e comparável das redes; avaliações internas revelam processos individuais, afetivos e cognitivos que sustentam a aprendizagem (Brookhart, 2011; Shepard, 2019). O desafio atual não é escolher entre uma ou outra, mas compreender como cada tipo de dado cumpre uma função distinta e complementar, permitindo interpretações mais completas e sensíveis sobre o que significa aprender.

4. DESENVOLVIMENTO

Esta pesquisa foi realizada ao longo de um ano letivo com uma estudante do 6º ano do ensino fundamental de uma escola municipal, bem como na análise comparativa entre os registros qualitativos produzidos pela professora de matemática da estudante e as informações geradas pelas avaliações externas em larga escala. A escola investigada está localizada em um território urbano marcado por desafios sociais e econômicos, atendendo predominantemente estudantes de famílias de baixa renda.

Esse contexto produz uma dinâmica escolar heterogênea, na qual diferentes experiências de vida, repertórios culturais e ritmos de aprendizagem coexistem e se expressam de maneiras diversas. A estudante escolhida para o estudo apresentava três características centrais: é uma aluna laudada, demonstrava comportamento timidamente retraído e mantinha elevada organização e regularidade na execução de atividades e no uso dos materiais. Esses critérios possibilitaram observar uma trajetória de aprendizagem com nuances cognitivas e socioemocionais particularmente visíveis, o que favoreceu a análise aprofundada da relação entre o olhar individualizado da avaliação interna e o caráter agregado das avaliações externas.

Durante todo o ano letivo, a professora de Matemática produziu registros qualitativos contínuos sobre o desempenho, a participação e as estratégias utilizadas pela aluna. Esses registros incluíam observações sobre sua compreensão conceitual, sua autonomia, suas dificuldades específicas, sua reação a desafios e erros, e a forma como interagia com colegas e com a própria docente. A professora também conduziu intervenções pedagógicas pontuais voltadas ao fortalecimento de conteúdos nos quais a estudante demonstrava fragilidade, como operações com números racionais, leitura e interpretação de problemas e justificativa de procedimentos matemáticos.

Outro aspecto de destaque ao longo do acompanhamento foi a dimensão relacional

estabelecida entre a professora e a estudante. A cada avanço, por menor que fosse, a docente oferecia elogios específicos e contextualizados, reforçando o esforço, a estratégia utilizada ou a persistência demonstrada pela estudante. Ao longo do acompanhamento, a relação entre professora e estudante tornou-se um elemento central. Elogios específicos diante de pequenos avanços fortaleceram gradualmente a autoestima acadêmica da aluna, antes tímida e insegura. Com isso, ela passou a participar mais, expressar dúvidas e confiar na mediação docente. O reconhecimento verbal criou um ambiente emocional seguro, a ponto de a estudante compartilhar espontaneamente questões pessoais, evidenciando o vínculo construído. Esse fenômeno evidencia a interdependência entre dimensões afetivas e cognitivas no processo de aprendizagem, como apontam Vygotski (1995) e Eisenberg et al. (2004).

Os registros da professora revelam um percurso marcado por avanços graduais. Identificou-se aumento da participação em sala, melhora na compreensão conceitual de conteúdos matemáticos, ampliação da autonomia e persistência diante de tarefas mais complexas. Apesar disso, algumas dificuldades permaneceram ao longo do ano, como a justificativa de procedimentos, a leitura cuidadosa de enunciados e a verbalização do raciocínio utilizado na resolução de problemas. Esses elementos, embora sutis, configuram indicadores relevantes da aprendizagem, pois evidenciam não apenas o que a estudante domina, mas como ela pensa matematicamente.

Quando esses registros são contrastados com as informações produzidas pelas avaliações externas — no caso, as escalas de proficiência do Saeb e os indicadores derivados do Ideb —, percebe-se uma diferença substantiva na natureza dos dados. As avaliações externas oferecem, sobretudo, resultados numéricos agregados, derivados de modelos estatísticos que situam estudantes em uma escala contínua de habilidade. Esses números, embora tecnicamente precisos, representam tendências coletivas e não capturam processos emocionais, interacionais ou estratégicos, como a importância do elogio no desenvolvimento da estudante. Na avaliação externa, essa aluna seria representada apenas por sua proficiência final, sem que seu percurso, suas inseguranças, seus avanços socioemocionais e sua relação com a professora fossem considerados.

Observam-se, contudo, alguns pontos de aproximação. Certas habilidades que se materializam nas escalas de proficiência dialogam com objetos do conhecimentos trabalhados pela professora, como resolver problemas que envolvem operações com números racionais, localizar pontos e deslocamentos no plano cartesiano, resolver

problemas envolvendo perímetro, área e volume, entre outros. Também é possível que melhorias observadas qualitativamente — como maior persistência e compreensão conceitual — repercutissem em melhor desempenho em uma avaliação externa. Porém, essas conexões são indiretas, pois o escopo da avaliação externa não contempla manifestações individuais de aprendizagem, mas padrões de desempenho da rede ou do sistema.

O caso analisado evidencia o contraste entre duas formas de compreender a aprendizagem. A avaliação interna revela o processo — as oscilações entre insegurança e progresso, o papel do afeto, as estratégias adotadas e a construção da autonomia. Já a avaliação externa revela o produto, situando a estudante em um ponto da escala de proficiência. Cada perspectiva cumpre uma função distinta: a interna ilumina nuances individuais invisíveis em métricas padronizadas, enquanto a externa oferece um panorama coletivo útil para identificar tendências e desigualdades. Assim, ambas são complementares, mas não intercambiáveis.

O desafio, como aponta a literatura (Shepard, 2019; Brookhart, 2011), é articular esses dois mundos (o da singularidade e o da generalidade) de forma ética e pedagogicamente significativa.

5. ANÁLISE E DISCUSSÃO DOS DADOS OBTIDOS

A análise dos dados coletados no acompanhamento da estudante permitiu compreender, de forma aprofundada, como os registros qualitativos produzidos pela professora revelam nuances do processo de aprendizagem que não se expressam nos indicadores numéricos das avaliações externas. Essa discussão, ancorada nos referenciais teóricos da avaliação educacional, da psicologia do desenvolvimento e da psicometria, permite responder ao problema central da pesquisa: em que medida descrições individuais de aprendizagem dialogam, ou não, com as métricas agregadas produzidas pelas avaliações em larga escala?

Os registros produzidos ao longo do ano revelaram uma trajetória marcada por avanços cognitivos, evoluções socioemocionais e transformações na postura da estudante em sala de aula. A combinação entre intervenções pedagógicas planejadas, feedbacks personalizados e elogios constantes desencadeou um processo de fortalecimento da autoestima acadêmica, aspecto apontado pela literatura como determinante para o engajamento cognitivo (Eisenberg et al., 2004; Bandura, 1997). Esse movimento confirma a

tese vygotskiana de que a aprendizagem é, simultaneamente, social e individual, construída por meio de interações que expandem a Zona de Desenvolvimento Proximal (Vigotski, 1995).

Ao ser encorajada pela professora, a estudante passou a arriscar respostas, formular perguntas e compartilhar dificuldades, demonstrando que os afetos são motores da evolução cognitiva. Do ponto de vista da aprendizagem matemática, observou-se um crescimento gradual na compreensão de conceitos e no uso de estratégias, ainda que persistissem dificuldades em justificar procedimentos e verbalizar raciocínios. Esse tipo de progresso (incremental, contextual e dependente da mediação docente) alinha-se ao que Perrenoud (1999) denomina desenvolvimento de competências, entendido como a ampliação de esquemas de ação capazes de mobilizar saberes de maneira flexível.

Ao relacionar esses elementos com os pressupostos da avaliação formativa (Black; Wiliam, 1998; Brookhart, 2011), percebe-se que os registros da professora funcionaram como instrumentos de monitoramento contínuo, permitindo ajustar intervenções e promover avanços específicos. A narrativa qualitativa capturou dimensões essenciais que expressam como a estudante pensa, sente e aprende — dimensões que, como demonstra Shepard (2019), são fundamentais para decisões pedagógicas, mas não se traduzem em métricas padronizadas.

A segunda etapa da análise consistiu em confrontar esses registros com os elementos fornecidos pelas avaliações externas. Ao fazê-lo, emerge um contraste evidente entre o caráter processual da avaliação interna e o caráter pontual da avaliação externa. As escalas de proficiência expressam tendências agregadas em termos de níveis de habilidade, derivadas de modelos estatísticos de Teoria de Resposta ao Item (Klein, 2011; Baker, 2001). Esses números cumprem a função de monitorar sistemas educacionais, identificar desigualdades estruturais e orientar políticas públicas (Brasil, 2005; INEP, 2016). Entretanto, como adverte Messick (1989), a validade de uma medida depende da adequação de seus usos; utilizá-la para inferir características individuais é um desvio de finalidade.

A análise indica que os avanços da estudante — maior participação, fortalecimento da autoestima, desenvolvimento de estratégias de resolução e ampliação da autonomia — não seriam capturados por uma avaliação de larga escala. O exame externo revelaria apenas seu nível de proficiência em um dado momento, que é uma informação valiosa para o sistema, mas insuficiente para compreender a riqueza de seu percurso. Por outro lado, algumas habilidades específicas observadas qualitativamente, como resolver problemas envolvendo números racionais, poderiam refletir-se em um aumento de proficiência na escala do Saeb.

Assim, há pontos de convergência entre o micro e o macro, mas esses pontos são parciais e limitados pelo desenho de cada sistema avaliativo.

Esse descompasso confirma o que defende Brookhart (2011): avaliações externas e avaliações internas têm naturezas, funções e alcances distintos; não são substituíveis entre si, mas complementares. A interpretação pedagógica dos dados depende da articulação dessas duas camadas (a camada do sistema e a camada do sujeito), reconhecendo que métricas agregadas orientam políticas, enquanto registros individualizados orientam práticas de ensino.

Os dados analisados também revelam como a avaliação interna permite observar fenômenos que, embora invisíveis nas métricas externas, têm impacto direto na aprendizagem. O fato de os elogios terem aumentado a confiança da estudante corrobora ideias da Psicologia Socioemocional de que o reconhecimento positivo constitui, para muitos alunos, um gatilho para engajamento e autorregulação (Zimmerman, 2002). Esse resultado reforça a necessidade de políticas que valorizem condições pedagógicas de cuidado, vínculo e acompanhamento próximo, pois tais dimensões não aparecem nos indicadores, mas influenciam diretamente os resultados que neles se mostram.

Em síntese, a análise dos dados evidenciaram que:

1. A avaliação interna captura processos; a externa captura resultados pontuais. Cada uma tem finalidades distintas e não pode ser interpretada fora de seu escopo teórico.
2. A trajetória individual da estudante mostra nuances que se perdem quando traduzidas em proficiência, revelando limites epistemológicos das métricas agregadas.
3. Há pontos de contato entre avanço qualitativo e desempenho quantitativo, mas essa relação é probabilística, não determinística.
4. A dimensão afetiva tem papel central na aprendizagem, e seu impacto, embora real, permanece ausente das avaliações externas.
5. A articulação entre avaliações interna e externa é essencial, mas exige responsabilidade interpretativa, como afirmam os Standards (AERA; APA; NCME, 2014).

Assim, a discussão revela que o desafio contemporâneo da avaliação educacional não consiste em substituir um modelo pelo outro, mas em construir interpretações que integrem, de modo crítico, o sujeito real das aprendizagens e o sistema que busca monitorá-las.

6. CONSIDERAÇÕES FINAIS

O estudo investigou como os registros qualitativos de aprendizagem dialogam com os indicadores numéricos das avaliações externas, a partir do acompanhamento longitudinal de uma estudante do 6º ano. Os registros da professora revelaram avanços cognitivos, afetivos e socioemocionais, fortemente influenciados pela relação pedagógica e pelo uso consistente de feedbacks e elogios, o que fortaleceu a autoestima e ampliou o engajamento da aluna — elementos coerentes com os princípios da avaliação formativa.

Ao comparar esse percurso com os dados das avaliações externas, verificou-se que cada modelo cumpre funções distintas: a avaliação interna ilumina processos e singularidades; a externa oferece informações agregadas que orientam políticas e diagnósticos sistêmicos. Integradas, essas perspectivas permitem interpretações mais completas sobre a aprendizagem.

O caso evidencia que dimensões afetivas como timidez, segurança, vínculo e organização, essenciais para o aprender, não aparecem nos indicadores sistêmicos, reforçando a necessidade de práticas avaliativas múltiplas que articulem diferentes tipos de evidência. Também destaca o valor formativo do diálogo entre universidade e escola no âmbito do projeto de extensão.

Sugere-se, para estudos futuros, a ampliação da investigação para outros perfis de estudantes, análises comparativas entre casos e pesquisas que integrem dados qualitativos e quantitativos em redes de ensino, de modo a fortalecer políticas e práticas sensíveis às complexidades das trajetórias individuais.

REFERÊNCIAS

AERA; APA; NCME. **Standards for educational and psychological testing.** Washington, DC: American Educational Research Association, 2014.

AMERICAN PSYCHOLOGICAL ASSOCIATION. **Diagnostic and statistical manual of mental disorders**. 5. ed. Washington, DC: APA, 2014.

ANDERSON, Lorin W.; KRATHWOHL, David R. (org.). **A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.** New York: Longman, 2001.

BAKER, Frank B. T**he basics of item response theory.** 2. ed. Washington, DC: ERIC Clearinghouse, 2001.

BANDURA, Albert. Self-efficacy: **The exercise of control**. New York: Freeman, 1997.

BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 2011.

BLACK, Paul; WILIAM, Dylan. **Assessment and classroom learning**. Assessment in Education, London, v. 5, n. 1, p. 7-74, 1998.

BOGDAN, Robert; BIKLEN, Sari. **Investigação qualitativa em Educação**. Porto: Porto Editora, 1994.

BONAMINO, Alicia; SOUSA, Sandra Zákia. **Avaliação Educacional: uma discussão do Saeb, Enem e Prova Brasil.** Educação & Realidade, Porto Alegre, v. 37, n. 2, p. 629-645, 2012.

BRASIL. Ministério da Educação. **Portaria nº 435**, de 3 de julho de 2005. Institui o Sistema de Avaliação da Educação Básica – SAEB. Brasília, DF: MEC, 2005. Disponível em: https://abmes.org.br/arquivos/legislacoes/PORTARIA435-21deoutubrode2025.pdf Acesso em: 20 ago. 2025.

BRASIL. Ministério da Educação. **Portaria nº 482**, de 7 de junho de 2013. Dispõe sobre a tradução metodológica do SAEB e integra a Avaliação Nacional da Alfabetização – ANA. Brasília, DF: MEC, 2013.

BRASIL. Ministério da Educação. **Portaria nº 458**, de 05 de maio de 2020. Institui normas complementares da Política Nacional de Avaliação da Educação Básica. Brasília, DF: MEC, 2020.

BROOKHART, Susan. **Formative assessment strategies for every classroom.** 2. ed. Alexandria: ASCD, 2011.

BUCK, Ross. **The affective neuroscience of social behavior.** Cambridge: Cambridge University Press, 2019.

EISENBERG, Nancy et al. The relations of emotionality and regulation to children's externalizing and internalizing problem behavior. **Child Development**, v. 65, n. 2, p. 486-498, 2004.

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Relatório SAEB/ANEB e ANRESC 2005–2015: **panorama da década.** Brasília, DF: Inep, 2016.

KLEIN, Ruben. **Escalas de proficiência: fundamentos, construção e interpretação**. Rio de Janeiro: Fundação Cesgranrio, 2011.

LÜDKE, Menga; ANDRÉ, Marli. **Pesquisa em educação: abordagens qualitativas**. São Paulo: EPU, 1986.

MESSICK, Samuel. Validity. In: LINN, Robert (org.). **Educational measurement**. 3rd. ed. New York: Macmillan, 1989. p. 13-103.

PERRENOUD, Philippe. **Construir competências desde a escola.** Porto Alegre: Artmed, 1999.

PONTES, Luís António Fajardo. A construção do Índice de Desenvolvimento da Educação Básica: fundamentos, limites e possibilidades. 2017. **Tese (Doutorado em Educação)** — UFJF, Juiz de Fora, 2017.

SHEPARD, Lorrie A. Classroom assessment to support teaching and learning. In: **The future of assessment**. New York: Routledge, 2019.

SNOW, Richard; CORNO, Lynn. Conative and affective processes in educational situations. In: BERLINER, David; CALFEE, Richard (org.). **Handbook of educational psychology**. New York: Macmillan, 1998.

STAKE, Robert. **The art of case study research.** Thousand Oaks: Sage, 1995.

VYGOTSKI, Lev S. **A formação social da mente.** São Paulo: Martins Fontes, 1995.

YIN, Robert. **Case study research:** design and methods. 5. ed. **Thousand Oaks: Sage**, 2015.

ZIMMERMAN, Barry. Becoming a self-regulated learner. **Theory into Practice**, v. 41, n. 2, p. 64-70, 2002