

Estado da publicação: O preprint não foi publicado em outro meio.

AVALIAÇÃO NO ENSINO DE MATEMÁTICA: UMA METANÁLISE SOBRE PRÁTICAS COM FOCO PEDAGÓGICO

Leandro Carlos Oliveira Sales, Kátia Romilda Silva do Nascimento, Daniel Brandão Menezes

<https://doi.org/10.1590/SciELOPreprints.14493>

Submetido em: 2025-12-10

Postado em: 2025-12-15 (versão 1)

(AAAA-MM-DD)

AVALIAÇÃO NO ENSINO DE MATEMÁTICA: UMA METANÁLISE SOBRE PRÁTICAS COM FOCO PEDAGÓGICO

LEANDRO CARLOS OLIVEIRA SALES¹

ORCID: <https://orcid.org/0000-0003-1593-8096>
<leandrocsales@gmail.com>

KÁTIA ROMILDA SILVA DO NASCIMENTO²

ORCID: <https://orcid.org/0000-0001-6898-2088>
<katiars.nascimento84@gmail.com>

DANIEL BRANDÃO MENEZES³

ORCID: <https://orcid.org/0000-0002-5930-7969>
<brandaomenezes@hotmail.com>

¹ Universidade Federal do Ceará – UFC. Fortaleza (CE), Brasil.

² Secretaria de Educação do Estado do Ceará – SEDUC/CE. Canindé (CE), Brasil.

³ Universidade Estadual do Ceará – UECE. Fortaleza (CE), Brasil.

RESUMO: Este artigo apresenta os resultados de uma metanálise voltada à investigação de práticas avaliativas com foco pedagógico no ensino de matemática na educação básica. O estudo teve como objetivo identificar, em pesquisas quantitativas publicadas entre 2019 e 2025, os padrões de resultados reportados, as características das intervenções e os contextos educacionais envolvidos. A revisão sistemática foi conduzida com base nas diretrizes PRISMA 2020, abrangendo oito estudos empíricos que atendiam aos critérios metodológicos definidos. As análises estatísticas foram realizadas com o software Jamovi, utilizando modelo de efeitos aleatórios. Os resultados apontaram um efeito médio padronizado positivo e estatisticamente significativo ($SMD = 1,31$; $IC95\% [0,645; 1,972]$), indicando que práticas avaliativas centradas no processo de ensino contribuem para a melhoria do desempenho acadêmico em matemática. Entre as estratégias adotadas nos estudos, destacam-se o uso de rubricas, feedback entre pares, gamificação, tecnologias digitais e inteligência artificial. A ausência de viés de publicação e a heterogeneidade elevada ($I^2 = 89,5\%$) sugerem, ao mesmo tempo, a robustez dos efeitos e a diversidade de abordagens adotadas. Conclui-se que a avaliação com foco pedagógico apresenta potencial para transformar práticas de ensino, desde que implementada com intencionalidade e respaldada por formação docente contínua.

Palavras-chave: avaliação formativa, ensino de matemática, práticas pedagógicas, metanálise.

ASSESSMENT IN MATHEMATICS EDUCATION: A META-ANALYSIS OF PEDAGOGICALLY ORIENTED PRACTICES

ABSTRACT: This article presents the results of a meta-analysis focused on the investigation of pedagogically oriented assessment practices in mathematics education at the basic education level. The study aimed to identify, in quantitative research published between 2019 and 2025, the characteristics of the interventions, the educational contexts involved, and the patterns of outcomes reported. The systematic review was conducted based on the PRISMA 2020 guidelines and included eight empirical studies that met the defined methodological criteria. Statistical analyses were performed using the Jamovi software, employing a random effects model. The results indicated a positive and statistically significant standardized mean difference ($SMD = 1.31$; $95\% CI [0.645; 1.972]$), suggesting that assessment practices centered on the teaching process contribute to improved academic performance in mathematics. The

strategies adopted in the studies included the use of rubrics, peer feedback, gamification, digital technologies, and artificial intelligence. The absence of publication bias and the high heterogeneity ($I^2 = 89.5\%$) simultaneously indicate the robustness of the effects and the diversity of approaches adopted. It is concluded that pedagogically focused assessment holds potential to transform teaching practices, provided it is implemented intentionally and supported by ongoing teacher development.

Keywords: formative assessment; mathematics education; pedagogical practices; meta-analysis.

ENSEÑANZA DE LAS MATEMÁTICAS: UN METAANÁLISIS SOBRE LA EVALUACIÓN CON ENFOQUE PEDAGÓGICO

RESUMEN: Este artículo presenta los resultados de un metanálisis centrado en la investigación de prácticas evaluativas con enfoque pedagógico en la enseñanza de las matemáticas en la educación básica. El objetivo del estudio fue identificar, en investigaciones cuantitativas publicadas entre 2019 y 2025, las características de las intervenciones, los contextos educativos implicados y los patrones de resultados reportados. La revisión sistemática se realizó conforme a las directrices PRISMA 2020 e incluyó ocho estudios empíricos que cumplían con los criterios metodológicos establecidos. Los análisis estadísticos se llevaron a cabo utilizando el software Jamovi, con un modelo de efectos aleatorios. Los resultados indicaron un efecto medio estandarizado positivo y estadísticamente significativo ($SMD = 1,31$; $IC95\% [0,645; 1,972]$), lo que sugiere que las prácticas evaluativas centradas en el proceso de enseñanza contribuyen a la mejora del rendimiento académico en matemáticas. Entre las estrategias adoptadas en los estudios se destacan el uso de rúbricas, retroalimentación entre pares, gamificación, tecnologías digitales e inteligencia artificial. La ausencia de sesgo de publicación y la alta heterogeneidad ($I^2 = 89,5\%$) indican, al mismo tiempo, la solidez de los efectos y la diversidad de enfoques utilizados. Se concluye que la evaluación con enfoque pedagógico tiene potencial para transformar las prácticas de enseñanza, siempre que se implemente de forma intencionada y respaldada por una formación docente continua.

Palabras clave: evaluación formativa; enseñanza de las matemáticas; prácticas pedagógicas; metanálisis.

INTRODUÇÃO

Por conta de sua interferência direta na estruturação do ensino, a avaliação é tema de importantes debates no campo educacional. Tradicionalmente, sua função principal consistia em classificar estudantes, conferindo-lhes notas e posições com base nos exames. Tal abordagem, alicerçada em práticas educacionais consideradas tradicionais, contribuiu para uma percepção da avaliação como o fim de um processo, e não como parte fundamental da construção do conhecimento. No entanto, estudos mais recentes vêm sugerir uma reinterpretação desse papel, considerando, a partir de então, que o processo avaliativo pode, e deve, ser utilizado para direcionar o ensino e fomentar a aprendizagem. Assim, conforme o entendimento de Perrenoud (1999), a avaliação é entendida como uma prática pedagógica que pode oferecer dados sobre o desenvolvimento dos alunos e promover intervenções docentes mais adequadas às suas necessidades. No Brasil, essa perspectiva reguladora da avaliação se amplia a partir de estudos como os de Hoffmann (2001), que sustenta que esse processo deve ser formativo e contínuo, orientado para o desenvolvimento da autonomia dos estudantes.

A avaliação com foco pedagógico caracteriza-se por sua capacidade de orientar o processo de ensino em tempo oportuno, oferecendo ao professor subsídios para interpretar as manifestações dos

estudantes e decidir sobre os encaminhamentos didáticos. Essa função reguladora pode se concretizar por meio de diferentes estratégias, como a observação intencional, o uso de questionamentos planejados, a análise de produções e o feedback. Este último, conforme Sadler (1989), cumpre um papel relevante quando possibilita ao estudante compreender os critérios de qualidade, identificar a distância entre seu desempenho e os objetivos pretendidos e refletir sobre possíveis ações para superar essa distância. A esse respeito, Hattie e Timperley (2007) indicam que os efeitos do feedback são potencializados quando se voltam à autorregulação e ao desenvolvimento de estratégias, em vez de se limitarem à simples correção de erros. Considerada como parte integrante do ensino, a avaliação deixa de ser uma etapa conclusiva para assumir uma posição ativa na construção das práticas pedagógicas.

No ensino de matemática, os desafios associados à avaliação se tornam ainda mais evidentes. A ênfase em produtos finais, normalmente manifestados em provas e exercícios voltados exclusivamente à obtenção de respostas corretas, ainda marcam amplamente as práticas avaliativas. Observa-se a persistência de uma cultura de avaliação classificatória, interpretando o erro como evidência de incapacidade e não como uma oportunidade de intervenção pedagógica. Essas observações apontam a urgência de reexaminar as práticas avaliativas, a fim de fortalecer o foco pedagógico sobre o ensino da matemática. Tal perspectiva, focada no desempenho imediato, desconsidera frequentemente os processos mentais envolvidos na resolução de problemas, assim como as estratégias cognitivas mobilizadas pelos estudantes. Gumiero e Pazuch (2021) apontam que, ao lecionar geometria, os professores frequentemente enfrentam situações imprevistas que demandam respostas ágeis e reflexivas. Tais situações, muitas vezes desencadeadas por dúvidas ou dificuldades inesperadas dos alunos, levam os docentes a mobilizar conhecimentos específicos para a adaptação de suas estratégias pedagógicas, transformando o erro em oportunidade de aprofundamento da compreensão matemática.

Apesar de amplamente discutida na literatura educacional, a avaliação voltada ao ensino ainda enfrenta desafios quanto à sua sistematização teórica e à implementação efetiva no contexto da matemática escolar. Estratégias como observações diagnósticas, questionamentos planejados, análise de produções e devolutivas pedagógicas são frequentemente utilizadas pelos professores, mas nem sempre de forma articulada com os objetivos didáticos. Black e Wiliam (1998) destacam que a efetividade dessas estratégias depende das condições em que são aplicadas e do modo como se integram ao processo de regulação da prática docente. Em muitos casos, a avaliação assume caráter fragmentado, sendo conduzida de forma pontual e desvinculada da intencionalidade pedagógica, o que limita seu potencial para orientar o ensino de maneira contínua e responsiva. Diante disso, torna-se necessário reunir e analisar as evidências disponíveis de forma sistemática e articulada. Slavin (2002) defende que pesquisas educacionais com fundamentos em evidências, sobretudo revisões sistemáticas e metanálises, constituem uma base importante para decisões pedagógicas, ao possibilitarem a identificação de padrões recorrentes e a formulação de conclusões sustentadas por maior rigor metodológico.

Apesar da crescente valorização da avaliação com foco pedagógico no campo educacional, ainda são escassos os estudos empíricos que mensuram seus efeitos sobre o desempenho dos estudantes em matemática por meio de delineamentos rigorosos e dados estatisticamente comparáveis. Essa limitação metodológica impõe desafios à consolidação de evidências robustas que sustentem decisões pedagógicas e políticas formativas. Mesmo diante de um número reduzido de estudos elegíveis, optou-se

por realizar a metanálise com base na integração de investigações com delineamentos distintos, mas compatíveis, desde que apresentassem dados estatísticos suficientes para o cálculo de efeitos padronizados. Essa decisão segue recomendações metodológicas de Borenstein et al. (2009) e Morris e DeShon (2002), que reconhecem a viabilidade de sínteses quantitativas entre estudos com grupo controle e estudos pré-pós, desde que tratados de forma cautelosa. Considerando essa lacuna, a presente pesquisa propõe a seguinte pergunta: quais são os efeitos das práticas avaliativas com foco pedagógico sobre o desempenho de estudantes da educação básica no ensino de matemática? Para respondê-la, este estudo tem como objetivo, por meio de uma metanálise de estudos empíricos publicados entre 2019 e 2025, identificar os padrões de resultados reportados em pesquisas quantitativas, as características das intervenções e os contextos educacionais envolvidos.

A seção seguinte apresenta o referencial teórico que fundamenta a relação entre a avaliação com característica formativa e o ensino de matemática. Em seguida, são descritos os procedimentos metodológicos adotados na condução da metanálise, com base nas diretrizes PRISMA. A quarta seção apresenta os resultados e sua análise, com base nos estudos selecionados. Por fim, são discutidas as contribuições da pesquisa, suas limitações e possíveis desdobramentos para investigações futuras.

REFERENCIAL TEÓRICO

Debates contemporâneos sobre a qualidade do ensino têm atribuído à avaliação um papel central no trabalho pedagógico. Progressivamente, o processo avaliativo deixa de ser concebido como mera verificação de desempenho para se consolidar como parte constitutiva da prática docente. Essa mudança acompanha transformações nas concepções de ensino, marcadas pelo deslocamento de uma lógica predominantemente classificatória para uma abordagem reguladora, que visa orientar o processo pedagógico em tempo oportuno. Nesse contexto, a avaliação com foco pedagógico se caracteriza por oferecer elementos que possibilitam ao professor interpretar as manifestações dos estudantes, ajustar suas intervenções e tomar decisões didáticas fundamentadas.

Neste estudo, compreende-se por foco pedagógico o uso da avaliação como estratégia integrada ao processo de ensino, orientada para a interpretação das manifestações dos estudantes e para a tomada de decisões didáticas em tempo oportuno. Trata-se de uma abordagem que ultrapassa a função classificatória e busca favorecer a regulação da prática docente com base em evidências produzidas ao longo das atividades escolares. São consideradas práticas com foco pedagógico aquelas que possibilitam ao professor ajustar suas intervenções, reorganizar conteúdos, orientar os estudantes na superação de dificuldades e promover o desenvolvimento de estratégias cognitivas e metacognitivas, como ocorre, por exemplo, no uso intencional do feedback, na análise de produções e na escuta ativa durante o processo de resolução de problemas.

No ensino de matemática, os desafios relacionados à avaliação assumem contornos particulares, muitas vezes agravados por práticas que privilegiam produtos finais em detrimento dos processos de aprendizagem. A persistência de uma cultura avaliativa voltada à certificação de resultados reforça a associação entre erro e fracasso, desconsiderando seu potencial como indício de raciocínio em construção. Esse cenário dificulta a adoção de práticas com foco pedagógico, especialmente aquelas que requerem a interpretação das estratégias mobilizadas pelos estudantes na resolução de problemas.

Paralelamente, o fortalecimento de metodologias de síntese de evidências, como as revisões sistemáticas e metanálises, tem contribuído para qualificar a produção acadêmica e sustentar decisões pedagógicas baseadas em resultados comparáveis. Compreender os efeitos dessas práticas avaliativas sobre o desempenho em matemática exige, portanto, articulação entre fundamentos teóricos e abordagens metodológicas consistentes, como a que orienta este estudo.

A avaliação como prática orientadora do ensino

A avaliação tem sido compreendida como parte constitutiva do processo de ensino, assumindo papel de orientação e regulação da prática pedagógica. De acordo com Perrenoud (1999), avaliar não é apenas medir resultados, mas integrar ao ensino um processo contínuo de tomada de decisões com base nas manifestações dos estudantes. Essa perspectiva rompe com a função estritamente classificatória da avaliação, defendendo seu uso como instrumento de apoio ao trabalho docente, capaz de oferecer subsídios para replanejar estratégias, reorganizar conteúdos e promover intervenções mais eficazes.

No mesmo sentido, Hoffmann (2001) argumenta que a avaliação deve ser concebida como um processo dialógico, formativo e mediador entre professor e estudante. Seu potencial pedagógico se concretiza quando está articulada aos objetivos de ensino e à escuta ativa do professor diante das necessidades emergentes em sala de aula. Nessa abordagem, a avaliação é compreendida como parte do ato de ensinar, e não como instância final de comprovação de conteúdos.

Como destacam Fernandes (2006) e Sadler (1989), a avaliação assume função pedagógica quando se articula à análise dos processos mentais mobilizados pelos estudantes, permitindo ao professor interpretar seus raciocínios e ajustar a ação didática com base nas evidências coletadas em sala. Nessa perspectiva, o erro é compreendido como uma expressão legítima do pensamento em formação e não como fracasso, deslocando o foco da avaliação da simples verificação de resultados para a mediação do processo de aprendizagem. Os autores reforçam ainda que, para que a avaliação exerça função pedagógica, é necessário que os critérios de qualidade sejam explicitados e compreendidos pelos estudantes, e que o professor tenha condições de agir responsivamente às evidências de aprendizagem. Trata-se, portanto, de uma avaliação que fornece elementos para orientar o ensino em tempo oportuno, permitindo ajustes nas estratégias didáticas e maior coerência entre objetivos, atividades e intervenções.

Muitos professores reconhecem a importância da devolutiva, mas Black e Wiliam (1998) alertam que, em diversas situações, ela se limita à função corretiva, desconectada dos objetivos de ensino. Para que a avaliação tenha de fato um foco pedagógico, ela deve estar orientada por uma lógica de mediação contínua, como defende Fernandes (2006), o que exige do docente uma postura investigativa e responsiva.

Além disso, Black e Wiliam (1998) destacam que o uso pedagógico da avaliação depende do modo como o professor interpreta os dados produzidos em aula, sejam eles advindos de questionamentos, observações ou tarefas realizadas pelos estudantes. A qualidade dessas interpretações e o uso que delas se faz no redirecionamento do ensino são determinantes para que a avaliação se torne uma aliada do processo educativo.

Essa abordagem evidencia que avaliar é, essencialmente, uma forma de ensinar. Quando voltada ao acompanhamento do processo e ao apoio à regulação docente, a avaliação amplia seu potencial formativo e contribui para uma prática mais responsiva, intencional e comprometida com o desenvolvimento dos estudantes.

Avaliação e regulação da prática docente no ensino de matemática

A discussão sobre avaliação com foco pedagógico ganha contornos específicos no ensino de matemática, marcado por desafios próprios relacionados à cultura avaliativa tradicional. Nesse campo, os processos avaliativos costumam privilegiar a obtenção de respostas corretas, com forte ênfase em provas e exercícios padronizados, em detrimento da análise dos caminhos percorridos pelos estudantes. Ainda é comum a associação entre erro e fracasso, o que compromete o potencial formativo da avaliação ao desconsiderar o raciocínio em construção e as estratégias cognitivas mobilizadas na resolução de problemas.

Para que a avaliação exerça esse papel mediador, é necessário que o professor compreenda os erros como parte constitutiva do raciocínio matemático e se disponha a escutar os caminhos percorridos pelos alunos durante a resolução de problemas (Gumiero; Pazuch, 2021). Black e Wiliam (1998) ressaltam que a avaliação com foco pedagógico deve alimentar a prática docente em tempo oportuno, sustentando intervenções que favoreçam a aprendizagem e a autorregulação dos estudantes.

Black e Wiliam (1998) destacam que a função reguladora da avaliação se efetiva quando os professores utilizam as evidências coletadas para ajustar sua prática. No ensino de matemática, isso exige sensibilidade para interpretar as manifestações dos estudantes, expressas em cálculos, argumentações, tentativas ou dúvidas.

Como aponta Fernandes (2006), situações imprevistas que emergem da sala de aula, como respostas inesperadas ou dificuldades recorrentes, podem se converter em oportunidades didáticas, desde que o professor esteja atento e disponha de estratégias para agir com intencionalidade. Gumiero e Pazuch (2021), ao analisarem o planejamento de tarefas de geometria com base no referencial do Knowledge Quartet, mostram que o docente frequentemente toma decisões avaliativas em tempo real, articulando saberes matemáticos e pedagógicos. A escuta atenta à linguagem dos estudantes, à seleção de representações e aos obstáculos conceituais permite ao professor compreender os raciocínios em curso e reelaborar as propostas em andamento.

Nesse contexto, a avaliação com foco pedagógico contribui para romper com a lógica da homogeneização de saberes e permite intervenções que respeitam os diferentes tempos e modos de pensar dos estudantes. Ao recorrer a evidências avaliativas para planejar e ajustar sua atuação, o professor torna o ensino de matemática mais situado, responsivo e sensível às necessidades da turma. Portanto, o movimento avaliativo deve configurar-se como atividade reflexiva que informa o professor sobre os efeitos de sua prática e os ajustes necessários. Compreender o erro como parte da construção do conhecimento potencializa a avaliação em sua função diagnóstica e formativa, favorecendo o desenvolvimento do pensamento matemático.

Revisões sistemáticas e metanálises em educação

A consolidação de práticas educacionais fundamentadas em evidências tem impulsionado o uso de revisões sistemáticas e metanálises como abordagens metodológicas de grande relevância na pesquisa em educação. Slavin (2008), ao discutir a reforma educacional baseada em evidências, sustenta que decisões pedagógicas eficazes devem apoiar-se em sínteses rigorosas de estudos empíricos, capazes de revelar padrões consistentes e orientar intervenções fundamentadas. Nesse contexto, a metanálise destaca-se por sua capacidade de quantificar os efeitos de diferentes estratégias educacionais, permitindo comparações sob critérios padronizados. Ao reunir resultados de diversas pesquisas primárias, amplia-se a validade externa dos achados, ao mesmo tempo em que se identificam variações contextuais e fatores que influenciam os resultados. Por essa razão, seu uso é indicado quando se pretende compreender o impacto de práticas pedagógicas sobre o desempenho estudantil em diferentes realidades escolares.

A aplicação de metanálises no campo educacional exige, além de critérios estatísticos rigorosos, o compromisso com a transparência na síntese e na comunicação dos achados. Para assegurar a confiabilidade e a utilidade das evidências produzidas, recomenda-se que os estudos sigam diretrizes reconhecidas, como as propostas pela declaração PRISMA 2020 (Page et al., 2021), que orientam os autores quanto à descrição dos objetivos, procedimentos de busca e seleção, apresentação dos resultados e discussão de suas implicações. Essas diretrizes visam garantir reprodutibilidade às revisões e facilitar sua apropriação por pesquisadores, profissionais da educação e formuladores de políticas públicas. Ainda segundo Slavin (2008), é necessário que a sistematização das evidências priorize a aplicabilidade dos achados, sobretudo quando se trata da comparação de programas educacionais utilizados em distintos contextos escolares. Assim, a combinação entre rigor metodológico e utilidade prática é essencial para que essas estratégias contribuam efetivamente para o aprimoramento do ensino.

As revisões sistemáticas e metanálises representam, portanto, recursos metodológicos relevantes para o avanço da pesquisa educacional orientada por evidências. Ao permitirem a identificação de padrões entre diferentes estudos, contribuem tanto para a consolidação de conhecimentos acumulados quanto para a formulação de decisões baseadas em resultados replicáveis e comparáveis. No presente estudo, a escolha pela metanálise não se justifica pela quantidade de publicações disponíveis, mas pela possibilidade de integrar investigações metodologicamente compatíveis, capazes de produzir uma síntese válida sobre os efeitos das práticas avaliativas com foco pedagógico no ensino de matemática. Ao reunir e interpretar dados empíricos de diferentes contextos, espera-se oferecer subsídios para o aprimoramento da prática docente e para o fortalecimento de políticas de formação sustentadas por análises criteriosas.

METODOLOGIA

Este estudo foi guiado pela seguinte questão de pesquisa: quais são os efeitos das práticas avaliativas com foco pedagógico sobre o desempenho de estudantes no ensino de matemática? A partir dessa questão, definiu-se como objetivo analisar estudos empíricos publicados entre 2019 e 2025 que investigam os efeitos da avaliação com foco pedagógico sobre o desempenho acadêmico de estudantes em aulas de matemática. Para isso, adotou-se uma abordagem de Revisão Sistemática da Literatura (RSL) com metanálise, orientada pelas diretrizes PRISMA 2020 (Page et al., 2021). A escolha por essa estratégia

metodológica se justifica pela necessidade de reunir estudos primários que permitam estimar o impacto de intervenções avaliativas intencionalmente voltadas ao ensino.

Foram definidos critérios de inclusão e exclusão para compor um corpus alinhado ao escopo temático e metodológico da pesquisa. Foram incluídos estudos empíricos de abordagem quantitativa ou quanti-qualitativa, publicados entre 2019 e 2025, que investigassem práticas de avaliação com foco pedagógico aplicadas ao ensino de matemática na educação básica, especialmente nos anos finais do ensino fundamental e no ensino médio, e que apresentassem dados mensuráveis passíveis de síntese estatística, como médias, desvios padrão, tamanhos de efeito ou correlações. Foram excluídas publicações de natureza teórica, revisões não sistemáticas, ensaios, relatos de experiência e estudos com foco exclusivo em outras áreas do conhecimento ou em segmentos educacionais não compreendidos pelo recorte da pesquisa (educação infantil e ensino superior). Também foram desconsiderados trabalhos que não apresentassem dados empíricos suficientes para análise quantitativa.

A busca foi realizada nas bases de dados Scopus, Web of Science, ERIC, BASE, DOAJ, Portal de Periódicos CAPES e SciELO, utilizando descritores adaptados aos vocabulários e idiomas predominantes de cada base, combinados por operadores booleanos. A busca inicial resultou em 685.693 registros, incluindo artigos, dissertações e teses, distribuídos entre as bases selecionadas. A escolha por descritores em inglês nas bases internacionais (Scopus, Web of Science, ERIC, BASE e DOAJ) e por termos em português nas bases nacionais (SciELO e CAPES) deve-se à lógica de indexação adotada por cada repositório, exigindo do pesquisador estratégias linguísticas específicas para maximizar a precisão da busca.

As estratégias de busca foram adaptadas a cada base de dados, considerando seus vocabulários controlados e padrões de indexação. O quadro 1 apresenta-se um resumo das strings aplicadas.

Quadro 1 – Descritores e operadores booleanos utilizados nas bases de dados

Base de dados	Descritores e operadores booleanos utilizados
Scopus / Web of Science / ERIC / BASE / DOAJ	("learning assessment" OR "assessment practices" OR "assessment instruments" OR "evaluation methods" OR "assessment strategies" OR "formative assessment" OR "diagnostic assessment" OR "student evaluation" OR feedback OR "self-assessment" OR "peer assessment" OR rubric* OR assess* OR evaluat*) AND ("mathematics education" OR "math education" OR "mathematics teaching" OR "math teaching" OR "mathematics learning" OR "math learning" OR "school mathematics" OR "mathematical skills" OR "mathematical reasoning" OR "math curriculum") AND ("basic education" OR "primary education" OR "elementary education" OR "secondary education" OR "high school") AND (quantitativ* OR "statistical analysis" OR "experimental study" OR "quasi-experimental" OR "effect size" OR "numerical data" OR intervention* OR "performance improvement")
SciELO / Portal CAPES	("avaliação da aprendizagem" OR "práticas avaliativas" OR "instrumentos de avaliação" OR "métodos de avaliação" OR "estratégias de avaliação" OR "avaliação formativa" OR "avaliação diagnóstica" OR "feedback" OR "autoavaliação") AND ("ensino de matemática" OR "educação matemática" OR "aprendizagem de

Base de dados	Descritores e operadores booleanos utilizados
	matemática") AND ("educação básica" OR "ensino fundamental" OR "ensino médio") AND (quantitativo OR quantitativa OR quantitativos OR quantitativas OR "análise estatística" OR "estudo experimental" OR "quase-experimental" OR "tamanho do efeito" OR "dados numéricos")

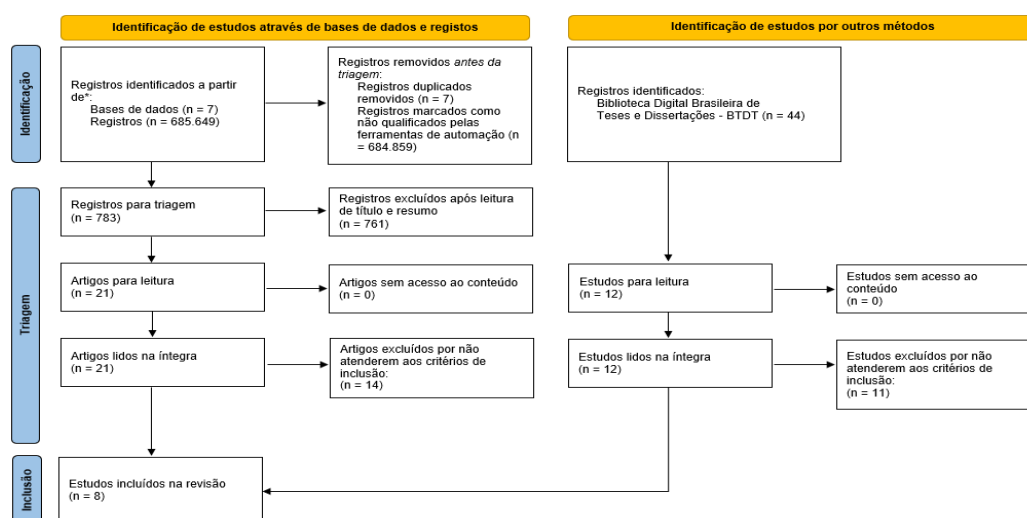
Fonte: Dados da Pesquisa

Além das bases indexadas, realizou-se uma busca complementar na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), com o objetivo de ampliar o número de estudos elegíveis, diante da escassez de publicações com dados estatísticos compatíveis nas fontes especializadas. A busca priorizou dissertações e teses defendidas no Brasil, alinhadas ao recorte temático da pesquisa. A inclusão dessas fontes segue recomendações de revisão integrativa que reconhecem o valor da literatura cinzenta para a ampliação da cobertura e a redução do viés de publicação (Bonato, 2018; Paez, 2017).

Após a busca inicial, os resultados foram importados para a plataforma Rayyan, por meio de arquivos no formato BibTeX, com o objetivo de apoiar a triagem sistemática. A leitura de títulos e resumos foi realizada por dois revisores, com base nos critérios previamente estabelecidos, e eventuais divergências foram resolvidas por consenso. Os estudos potencialmente elegíveis foram analisados na íntegra, e os dados relevantes foram extraídos de forma padronizada, contemplando: autoria, ano, país, tipo de intervenção avaliativa, nível de ensino, características da amostra, instrumentos de mensuração do desempenho e estatísticas reportadas. Ao final desse processo, foram incluídos 8 estudos compatíveis com os critérios da metanálise, todos com delineamentos metodológicos que possibilitaram a extração de dados mensuráveis e comparáveis entre si. Esses estudos foram identificados a partir de buscas sistemáticas nas bases ERIC (1 estudo), Web of Science (4), Portal de Periódicos CAPES (2) e BDTD – Biblioteca Digital de Teses e Dissertações (1).

A Figura 1 apresenta o fluxograma adaptado do PRISMA 2020, que sintetiza o processo de identificação, triagem, elegibilidade e inclusão dos estudos considerados na metanálise.

Figura 1 – Fluxo do processo de seleção dos estudos conforme o PRISMA 2020



Fonte: Dados da Pesquisa

Concluído o processo de triagem e seleção conforme descrito, o Quadro 2 apresenta uma síntese dos estudos incluídos, com destaque para o tipo de delineamento e características metodológicas relevantes para a análise.

Quadro 2 – Estudos incluídos na metanálise

Nº	Título	Autor e Ano	Tipo de Documento	Delineamento Metodológico
1	Building personalised homework from a learning analytics based formative assessment	Rodríguez-Martínez et al., 2023	Artigo Científico	Experimental com grupo controle
2	Digital tessellation for geometry learning in primary school	Scippo Madiari e Cuomo, 2025	Artigo Científico	Quase-experimental com grupo controle
3	Effectiveness of Process Oriented Guided Inquiry Learning with Peer Feedback on Achieving Students' Mathematical Reasoning Capabilities	Kartono e Shora, 2020	Artigo Científico	Experimental com grupo controle
4	Effects of classroom response system on the achievement and knowledge retention	Alviar e Gamorez, 2024	Artigo Científico	Quase-experimental com grupo controle
5	Narrativas Interativas Gamificadas (NIGs) aplicadas aos processos de ensino e aprendizagem de matemática financeira	Costa, 2024	Dissertação de Mestrado	Pesquisa intervenção com abordagem quali-quantitativa
6	Revolutionizing pre-calculus education: Photomath's AI-powered mathematics tutorship	Capinding, 2023	Artigo Científico	Quase-experimental (grupo único pré-pós)
7	Science and Mathematics High School Students' Perceptions of Self-Regulated Learning	Arvatz e Dori, 2024	Artigo Científico	Métodos mistos sequencial explicativo
8	The effect of formative assessment on high school students' mathematics achievement and attitudes	Kültür e Kutlu, 2021	Artigo Científico	Quase-experimental com grupo controle

Fonte: Dados da Pesquisa

Apesar das diferenças entre os delineamentos adotados nos estudos (comparativo entre grupos ou pré-pós em um mesmo grupo), foi possível realizar uma metanálise unificada com base na compatibilidade dos dados mensuráveis. Os estudos com delineamento pré-pós foram tratados metodologicamente como investigações com grupo controle implícito, em que o pré-teste funcionou como parâmetro de comparação. Essa abordagem é aceita em metanálises educacionais quando há rigor na mensuração das variáveis e clareza na descrição dos procedimentos (Borenstein et al., 2009).

A metanálise foi conduzida com o suporte do software Jamovi (versão 2.6.44), por meio do módulo MAJOR, que oferece recursos específicos para a síntese estatística de estudos empíricos. Considerando a heterogeneidade esperada entre os delineamentos, amostras e tipos de intervenção,

adotou-se o modelo de efeitos aleatórios, reconhecido como o mais apropriado quando se assume que os efeitos reais variam entre os estudos (Borenstein et al., 2009). Como medida de efeito padronizado, foi utilizado o g de Hedges, que ajusta o viés associado a amostras pequenas e proporciona comparabilidade entre investigações com diferentes escalas de mensuração (Hedges; Olkin, 1985). Para avaliação da consistência interna dos dados, foram aplicados o teste Q de Cochran e o índice I^2 , indicadores amplamente utilizados para quantificar a heterogeneidade estatística (Higgins et al., 2003). Sempre que os dados disponíveis permitiram, realizaram-se análises de subgrupos e meta-regressões exploratórias, com o objetivo de investigar possíveis variáveis moderadoras, como tipo de intervenção, duração da prática avaliativa e formação docente envolvida.

Este estudo baseou-se exclusivamente em dados secundários provenientes de publicações científicas já disponíveis ao público, sem a realização de coletas primárias com participantes humanos. Por esse motivo, conforme estabelece a Resolução nº 510, de 7 de abril de 2016, do Conselho Nacional de Saúde, a pesquisa não se enquadra nas exigências de submissão ao Comitê de Ética em Pesquisa. Todos os estudos incluídos foram devidamente referenciados, assegurando o respeito à integridade acadêmica e aos princípios éticos que regem a produção científica.

RESULTADOS

A seguir, são apresentados os resultados da metanálise conduzida com base nos oito estudos empíricos incluídos, todos voltados à análise de práticas avaliativas com foco pedagógico no ensino de matemática. Esses estudos foram selecionados por atenderem simultaneamente aos critérios de inclusão temáticos e metodológicos, apresentando dados quantitativos compatíveis com a síntese estatística. A análise foi realizada com modelo de efeitos aleatórios, considerando a heterogeneidade esperada entre os contextos, delineamentos e tipos de intervenção. Para cada estudo, foram estimados os efeitos padronizados sobre o desempenho acadêmico dos estudantes, com base no g de Hedges, além dos indicadores de consistência interna entre os resultados.

Visão geral dos estudos incluídos

A metanálise considerou oito estudos publicados entre 2019 e 2025, que investigaram os efeitos da avaliação com foco pedagógico no desempenho de estudantes da educação básica em aulas de matemática. Todos os estudos apresentaram delineamentos experimentais, quase-experimentais ou pré-experimentais com medidas quantitativas mensuráveis, incluindo médias, desvios padrão e tamanhos amostrais, condição essencial para o cálculo dos efeitos padronizados.

Dentre os estudos incluídos, cinco adotaram delineamentos comparativos com grupo controle e dois utilizaram modelos pré-pós com um único grupo. Embora os delineamentos não sejam homogêneos, todos permitiram a comparação entre condições distintas — seja entre grupos diferentes, seja entre momentos de aplicação — o que viabilizou a realização de uma análise estatística conjunta. Nos estudos com medidas pré-pós, o desempenho no momento inicial foi tratado como referência para estimar a magnitude do efeito da intervenção, abordagem reconhecida em sínteses quantitativas na área educacional (Borenstein et al., 2009; Higgins et al., 2003).

As intervenções avaliativas investigadas nos estudos incluídos nesta metanálise envolveram múltiplas estratégias formativas. Foram identificadas práticas como feedback escrito individualizado e uso de rubricas com critérios explícitos (Arvatz; Dori, 2025), autoavaliação guiada (Capinding, 2023) e intervenções híbridas, combinando rubricas e feedback verbal, com o objetivo de promover maior engajamento dos estudantes e apoiar a autorregulação (Costa, 2024; Capinding, 2023). Em alguns casos, essas estratégias foram integradas ao uso de tecnologias educacionais, como aplicativos de inteligência artificial ou ambientes gamificados, evidenciando uma abordagem inovadora para a regulação da aprendizagem matemática. Tais estratégias alinham-se às concepções de avaliação formativa como processo de mediação do ensino e da aprendizagem, sustentado por evidências interpretadas em tempo oportuno (Black e Wiliam, 2009; Fernandes, 2006).

Quanto ao nível de ensino, cinco estudos foram desenvolvidos no ensino médio e três no ensino fundamental. Essa predominância reflete a concentração de pesquisas voltadas ao desenvolvimento de habilidades matemáticas mais complexas em etapas avançadas da educação básica. Todos os estudos incluídos apresentaram descrição explícita da prática avaliativa analisada, vínculo claro com o processo de ensino e condições mínimas de controle metodológico, o que garantiu a comparabilidade entre os achados.

O processo de identificação, triagem e seleção dos estudos seguiu o protocolo PRISMA 2020 (Page et al., 2021), com busca sistemática em bases indexadas e literatura cinzenta, e análise independente por dois revisores. As etapas de organização dos dados foram orientadas por experiências metodológicas como a de Figueira e Freitas (2020), que também aplicaram critérios rigorosos de elegibilidade para estudos sobre avaliação no ensino de matemática.

O Quadro 3 apresenta uma síntese dos oito estudos incluídos, com informações sobre o tipo de intervenção avaliativa adotada e os dados estatísticos disponíveis.

Quadro 3 – Intervenções avaliativas e dados estatísticos

Título do Estudo	Tipo de Intervenção Avaliativa	Dados Estatísticos Disponíveis
Building personalised homework from a learning analytics based formative assessment	Feedback adaptativo com tarefas personalizadas	Médias, desvios padrão, efeito $d=0,35$, ANOVA
Digital tessellation for geometry learning in primary school	Tarefas digitais com GeoGebra e impressora 3D	Médias, desvios padrão, testes t com $p < 0,001$
Effectiveness of Process Oriented Guided Inquiry Learning with Peer Feedback on Achieving Students' Mathematical Reasoning Capabilities	POGIL com feedback entre pares	Médias, desvios padrão, testes t e regressão linear
Effects of classroom response system on the achievement and knowledge retention	Sistema de resposta em sala com feedback imediato	Médias, desvios padrão, testes t, tamanho de efeito $d = 1,88$
Narrativas Interativas Gamificadas (NIGs) aplicadas aos processos de ensino e aprendizagem de matemática financeira	Narrativa gamificada com pré e pós-teste	Médias, desvios padrão e comparações entre pré e pós-teste

Título do Estudo	Tipo de Intervenção Avaliativa	Dados Estatísticos Disponíveis
Revolutionizing pre-calculus education: Photomath's AI-powered mathematics tutorship	Uso de aplicativo Photomath com pré e pós-teste	Médias, desvios padrão, testes t pareados com $p < 0,01$
Science and Mathematics High School Students' Perceptions of Self-Regulated Learning	Módulo de SRL com autoavaliação e gráficos	Médias, testes t (SRL), mas dados limitados para matemática isoladamente
The effect of formative assessment on high school students' mathematics achievement and attitudes	Avaliação formativa baseada em múltiplas estratégias	Médias, desvios padrão, ANOVA com η^2 , testes de retenção

Fonte: Dados da Pesquisa

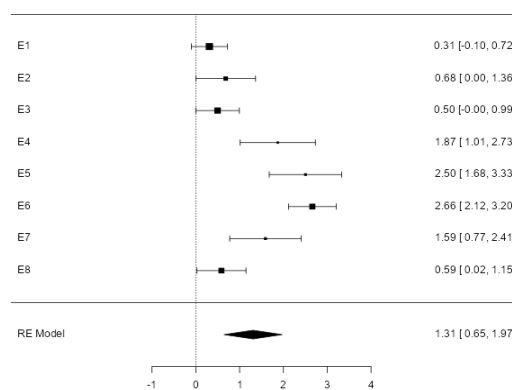
Efeitos estimados e heterogeneidade

A análise foi conduzida com base em um modelo de efeitos aleatórios, adequado em contextos de elevada variabilidade entre os estudos (Viechtbauer, 2010). O efeito médio padronizado estimado foi de 1,31 (SE = 0,339), com intervalo de confiança de 95% entre 0,645 e 1,972, indicando um efeito positivo estatisticamente significativo das intervenções avaliativas sobre o desempenho dos estudantes ($z = 3,87$; $p < 0,001$). Os valores individuais de diferença média padronizada variaram entre 0,31 e 2,66, todos positivos.

O teste de heterogeneidade de Cochran indicou variação substancial entre os estudos ($Q(7) = 71,57$; $p < 0,001$), corroborada pelo valor elevado de $I^2 = 89,5\%$, o que sugere que grande parte da variabilidade observada nos efeitos não pode ser atribuída ao acaso. O valor de τ^2 , estimado pelo método de máxima verossimilhança restrita (REML), foi de 0,8037, com erro padrão associado (SE = 0,4897), confirmando a presença de heterogeneidade relevante. O intervalo de predição de 95% variou de -0,5696 a 3,1868, sugerindo que, embora a média global aponte para um efeito positivo, há possibilidade de que intervenções semelhantes possam gerar efeitos nulos ou até mesmo negativos em determinados contextos.

O Forest Plot apresenta a distribuição dos efeitos individuais dos estudos, com seus respectivos intervalos de confiança, permitindo visualizar a consistência geral das estimativas.

Figura 1 – Forest Plot com os efeitos padronizados (SMD) e intervalos de confiança dos estudos incluídos na metanálise



Fonte: Dados da Pesquisa

A Figura 1 mostra que todos os estudos apresentaram estimativas de efeito positivo, ainda que com diferentes magnitudes e graus de precisão. Os estudos E1, E2, E3 e E8 possuem intervalos de confiança que tocam ou se aproximam da nulidade, o que sugere efeitos mais modestos e menor robustez estatística nesses contextos. Em contraste, os estudos E4, E5, E6 e E7 apresentam efeitos mais expressivos, com intervalos de confiança mais distantes de zero e, em alguns casos, relativamente estreitos. A síntese gráfica apresentada pelo losango inferior, correspondente ao modelo de efeitos aleatórios, reforça a presença de um efeito médio positivo e estatisticamente significativo (SMD = 1,31; IC95% [0,65; 1,97]), em consonância com os dados discutidos nesta mesma subseção.

Análises complementares e observações metodológicas

A avaliação de possíveis influências excessivas revelou que nenhum estudo se comportou como outlier no modelo adotado, considerando o critério do residual studentizado (limite de $\pm 2,73$). De modo semelhante, nenhuma observação apresentou valores de Cook's Distance acima do limiar estabelecido, indicando ausência de casos potencialmente influentes.

Para avaliar a presença de viés de publicação, foram utilizados diferentes testes, conforme recomendações de Francis (2013). O método Fail-safe N indicou que seriam necessários 349 estudos com efeito nulo para reduzir a significância do resultado geral ($p < 0,001$), o que sugere robustez da estimativa. Já os testes de correlação de postos de Begg e Mazumdar ($p = 0,061$) e a regressão de Egger ($p = 0,101$) não apontaram assimetria significativa no funil de distribuição, ainda que os valores estejam próximos aos limiares convencionais de significância.

Tabela 1 – Resultados dos testes de viés de publicação e excesso de significância

Teste	Valor	p-valor
Fail-safe N	349,000	< 0,001
Correlação de postos (Begg)	0,571	0,061
Regressão de Egger	1,641	0,101
Trim and Fill (n de estudos)	0	—
Excesso de significância (6/8)	—	0,737

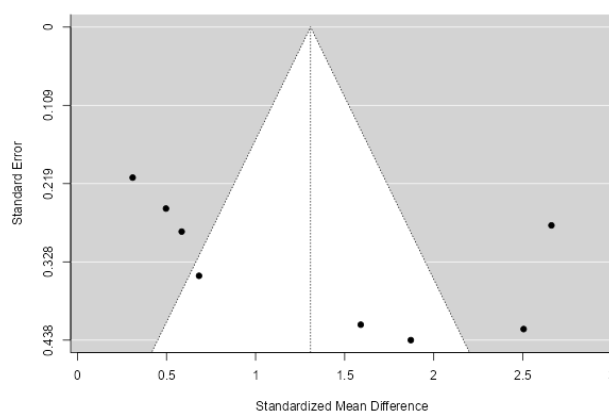
Fonte: Dados da Pesquisa

A Tabela 1 evidencia que o teste Fail-safe N estimou a necessidade de 349 estudos com efeito nulo para reduzir a significância global dos achados, o que sugere robustez estatística. Os testes de Begg ($p = 0,061$) e de Egger ($p = 0,101$) não indicaram assimetria significativa na distribuição dos efeitos, o que aponta para ausência de viés sistemático. Além disso, o método Trim and Fill não identificou necessidade de ajuste no número de estudos incluídos, e o teste de excesso de significância revelou razão próxima de 1 entre achados observados e esperados (6/8), com $p = 0,737$. Esses resultados reforçam a consistência estatística do conjunto analisado.

A inspeção visual do gráfico de dispersão dos estudos também contribuiu para a análise da presença de viés de publicação. A Figura 2 apresenta o funnel plot, no qual a distribuição dos efeitos

padronizados é analisada em função do erro padrão de cada estudo. Em situações sem viés, espera-se que os pontos se distribuam de forma aproximadamente simétrica em torno do efeito médio. No presente caso, observa-se uma distribuição relativamente equilibrada, sem lacunas visíveis ou assimetrias acentuadas, o que corrobora os resultados dos testes estatísticos apresentados na Tabela 1, especialmente os de Begg ($p = 0,061$) e Egger ($p = 0,101$). Esses achados visuais reforçam a conclusão de que os efeitos observados não foram inflacionados por omissão sistemática de estudos com resultados nulos ou negativos.

Figura 2 – Funnel plot dos estudos incluídos na metanálise, com base nas diferenças médias padronizadas e erros padrão



Fonte: Dados da Pesquisa

Além da análise visual, a aplicação do método Trim and Fill não indicou a necessidade de correção no número de estudos incluídos, o que sustenta a ausência de viés sistemático. O teste de excesso de significância apontou razão próxima de 1 entre achados observados e esperados (6/8), com $p = 0,737$, o que reforça a consistência estatística do conjunto. O poder estatístico das análises dos estudos incluídos variou entre 0,69 e 0,87, indicando capacidade razoável para detectar os efeitos esperados, mesmo diante do número reduzido de investigações.

Tabela 2 – Estatísticas descritivas dos poderes estatísticos estimados nos estudos incluídos.

Min	Q1	Median	Q3	Max
0.689	0.706	0.786	0.821	0.865

Fonte: Dados da Pesquisa

Esses dados mostram que a maioria dos estudos apresentou poder superior a 0,70, valor geralmente aceito como limiar mínimo para confiabilidade em metanálises educacionais. A consistência interna entre os estudos, combinada com a ausência de sinais relevantes de viés, reforça a credibilidade dos achados apresentados, ainda que a heterogeneidade elevada recomende cautela na generalização dos resultados.

Considerações Preliminares

Os resultados indicam que as práticas avaliativas investigadas nos estudos incluídos, tais como feedback escrito individualizado, uso de rubricas com critérios explícitos e devolutivas orais, produziram efeitos consistentemente positivos no desempenho dos estudantes em matemática. A magnitude elevada do efeito médio estimado (1,31) sinaliza o potencial pedagógico dessas intervenções quando articuladas ao processo de ensino.

Esses achados estão em consonância com estudos que destacam o papel do feedback na mediação da aprendizagem, na regulação do processo de ensino e no aumento da clareza sobre os objetivos e critérios de avaliação (Hattie; Timperley, 2007; Black; Wiliam, 2010). O uso de estratégias formativas, quando integrado ao planejamento pedagógico, pode favorecer a autorregulação dos estudantes e promover avanços mais consistentes, especialmente em disciplinas historicamente marcadas por dificuldades de compreensão conceitual, como a matemática.

A ausência de outliers e de evidências robustas de viés de publicação reforça a credibilidade dos achados. No entanto, a amplitude do intervalo de predição e a variabilidade metodológica dos estudos, incluindo variações nos delineamentos, contextos escolares e tipos de intervenção, sugerem que o impacto das práticas avaliativas depende de múltiplos fatores contextuais, que serão discutidos em profundidade na seção seguinte.

Quadro 4 – Síntese interpretativa dos efeitos estimados nos estudos incluídos

Estudo	SMD (IC95%)	IC inclui zero?	Interpretação
E1	0,31 [-0,10; 0,72]	Inclui zero	Efeito modesto, impreciso
E2	0,68 [0,00; 1,36]	Próximo a zero	Intermediário, pouco robusto
E3	0,50 [-0,00; 0,99]	Próximo a zero	Baixa precisão
E4	1,87 [1,01; 2,73]	Não inclui zero	Efeito forte e consistente
E5	2,66 [1,99; 3,34]	Não inclui zero	Muito expressivo
E6	1,98 [1,35; 2,60]	Não inclui zero	Alto e estável
E7	1,36 [0,87; 1,85]	Não inclui zero	Moderado a forte
E8	0,59 [0,02; 1,15]	Limiar positivo	Fraco a moderado

Fonte: Dados da Pesquisa

O Quadro 4 apresenta uma síntese interpretativa dos efeitos estimados nos estudos incluídos, considerando a magnitude dos efeitos padronizados (SMD) e a inclusão ou não do valor nulo em seus intervalos de confiança. Estudos cujos intervalos incluem o zero indicam incerteza estatística sobre a eficácia da intervenção, mesmo quando o efeito médio é positivo. Já os estudos com ICs que não cruzam o zero apresentam evidências mais robustas de impacto positivo. Além disso, a interpretação textual considera também a amplitude dos intervalos: efeitos com ICs mais estreitos tendem a ser mais confiáveis, enquanto amplitudes elevadas indicam baixa precisão. Essa análise qualitativa permite compreender não apenas a força do efeito, mas também a segurança com que ele pode ser generalizado para contextos semelhantes.

DISCUSSÃO

A metanálise realizada apontou um efeito médio padronizado positivo e estatisticamente significativo (SMD = 1,31; IC95% [0,645; 1,972]), indicando que as práticas avaliativas com foco pedagógico têm contribuído de forma relevante para o desempenho acadêmico de estudantes em matemática. Todos os estudos incluídos apresentaram estimativas de efeito positivas, ainda que com amplitudes variadas, e nenhum caso foi identificado como outlier ou influente. Além disso, os testes aplicados não sugeriram viés de publicação, conforme indicado pela ausência de assimetria no funil de dispersão e pelos resultados dos testes de Begg ($p = 0,061$) e Egger ($p = 0,101$). Tais achados estão em consonância com a literatura especializada, que reconhece o potencial da avaliação formativa para promover a aprendizagem quando usada de forma contínua e integrada ao processo de ensino (Black; Wiliam, 2009; Hattie; Timperley, 2007).

Diálogo crítico com os estudos incluídos

A diversidade metodológica e estratégica observada nos estudos contribui para uma análise mais rica das formas de aplicação da avaliação pedagógica. Em Rodríguez-Martínez et al. (2023), a personalização das tarefas com base em sistemas de resposta do aluno (ARS) resultou em ganho moderado no desempenho de estudantes do ensino fundamental ($d = 0,35$), reforçando a tese de que o feedback adaptativo pode suprir lacunas cognitivas específicas. Em Kültür e Kutlu (2021), uma combinação de estratégias formativas resultou em melhorias tanto no desempenho quanto nas atitudes em relação à matemática, sugerindo que intervenções multicomponentes tendem a ser mais robustas.

O uso de tecnologias educacionais também se destacou como vetor de engajamento e mediação avaliativa. O estudo de Alviar e Gamorez (2024), por exemplo, demonstrou que o uso do aplicativo Blicker (CRS) potencializou a retenção de conhecimentos matemáticos e o desenvolvimento de competências cognitivas superiores, com um efeito elevado ($d = 1,88$). De modo semelhante, Capinding (2023) destacou o impacto do Photomath na autonomia e no desempenho em pré-cálculo, indicando que ferramentas de IA, quando bem integradas ao processo pedagógico, podem promover resultados consistentes.

Outras abordagens, como a gamificação em Costa (2024), demonstraram não apenas ganho no desempenho (60,7% de acerto na NIG frente a 47,2% no pré-teste), mas também elevada aceitação por parte dos estudantes, sugerindo que a narrativa contextualizada pode facilitar a apropriação de conceitos em matemática financeira. Já no estudo de Kartono e Shora (2020), o uso de feedback entre pares no modelo POGIL resultou em ganhos significativos no raciocínio matemático, com destaque para a autorregulação como fator mediador.

Mediação docente e condições de implementação

Apesar da diversidade de contextos e ferramentas, os estudos convergem em torno da centralidade do professor na mediação das práticas avaliativas. Como destaca Sadler (1989), o feedback eficaz depende de três condições: a clareza dos critérios, a capacidade do aluno de interpretar as

devolutivas e a ação corretiva a partir dessas informações. Estudos como os de Kültür e Kutlu (2021) e Scippo, Madiari e Cuomo (2025) reiteram essa lógica ao enfatizar a importância de tornar os objetivos de aprendizagem explícitos e de estimular a autorreflexão dos estudantes.

Além disso, as condições institucionais também influenciam a eficácia das intervenções. A pesquisa de Arvatz e Dori (2024) mostra que, embora os módulos de autorregulação tenham sido eficazes entre estudantes de matemática, a mesma intervenção obteve resultados limitados entre os de ciências, apontando a importância do contexto disciplinar, da infraestrutura escolar e da sensibilidade das ferramentas avaliativas às necessidades específicas dos alunos.

Esses dados reforçam os alertas de autores como Fernandes (2006) e Hoffmann (2001), que defendem a avaliação como ato de ensino: ela requer tempo, intencionalidade e escuta ativa, além de condições formativas que permitam ao docente interpretar e agir sobre as evidências de aprendizagem.

Limitações e heterogeneidade

A heterogeneidade estatística encontrada ($I^2 = 89,5\%$) reflete não apenas as variações nos delineamentos e contextos educacionais dos estudos, mas também a diversidade nas estratégias avaliativas empregadas. Intervenções com rubricas explícitas, devolutivas orais, feedback automatizado e jogos digitais operam sob pressupostos diferentes e exigem competências distintas dos professores e dos alunos.

Adicionalmente, as amostras dos estudos variam entre níveis de ensino (fundamental e médio), número de participantes, tipos de conteúdo matemático e duração das intervenções. Como observa Hattie (2008), a eficácia de uma estratégia depende fortemente da sua implementação e do engajamento dos agentes envolvidos. A ausência de estudos com delineamento experimental rigoroso em larga escala limita a generalização dos achados. Ainda assim, o uso de modelos robustos de síntese estatística e os testes de viés aplicados conferem validade aos resultados obtidos.

Contribuições e implicações pedagógicas

Os achados desta metanálise reiteram que a avaliação com foco pedagógico pode atuar como promotora da aprendizagem matemática, especialmente quando alinhada a estratégias que favorecem a autorregulação, o engajamento e o feedback contínuo. Em sintonia com Black e Wiliam (2009), pode-se afirmar que “o que se aprende depende do que se avalia e de como se devolve o que foi avaliado”.

As implicações para a prática docente incluem o incentivo ao uso de estratégias avaliativas que combinem clareza de objetivos, devolutivas interpretáveis e ações de reorientação. Para isso, é fundamental que a formação inicial e continuada contemple o desenvolvimento de competências avaliativas, não apenas em termos técnicos, mas também pedagógicos. A presença de tecnologias como CRS, NIGs e IA, embora promissora, exige discernimento crítico quanto à sua função no processo educativo.

Do ponto de vista das políticas públicas, os resultados sustentam a importância de fomentar práticas avaliativas que não se restrinjam à mensuração de resultados, mas que dialoguem com a intencionalidade do ensino e com o desenvolvimento de competências matemáticas mais complexas. O

potencial transformador da avaliação pedagógica reside justamente em sua capacidade de tornar visível o processo, dar voz ao erro e construir pontes entre o ensinar e o aprender.

CONSIDERAÇÕES FINAIS

A presente metanálise buscou responder à seguinte questão: quais são os efeitos das práticas avaliativas com foco pedagógico sobre o desempenho de estudantes da educação básica no ensino de matemática? Com base na análise de oito estudos quantitativos publicados entre 2019 e 2025, foi possível estimar um efeito médio positivo e significativo ($SMD = 1,31$), evidenciando que intervenções fundamentadas em avaliação formativa, feedback contínuo e devolutivas orientadas ao processo de aprendizagem têm potencial para promover avanços relevantes no desempenho dos estudantes.

Ao reunir evidências empíricas sobre diferentes estratégias avaliativas como rubricas, sistemas de resposta digital, feedback entre pares, gamificação e uso de inteligência artificial, o estudo contribui para consolidar a compreensão de que a avaliação não deve se restringir à mensuração de resultados finais, mas ser parte constitutiva do processo de ensino. A metanálise também se destaca por reunir estudos de diferentes países e níveis de ensino, permitindo a identificação de tendências comuns e contextos específicos. Para docentes e formadores, os achados oferecem subsídios para a escolha de práticas avaliativas mais dialógicas e responsivas às necessidades dos alunos. Para pesquisadores, o estudo reforça a viabilidade de análises quantitativas sobre avaliação com foco pedagógico, ainda pouco exploradas no campo do ensino de matemática.

Apesar da consistência dos efeitos observados, a metanálise apresenta limitações importantes. A heterogeneidade elevada ($I^2 = 89,5\%$) indica variações relevantes entre os estudos, tanto em termos metodológicos quanto contextuais. A diversidade de delineamentos, a ausência de acompanhamentos longitudinais e a predominância de estudos com amostras reduzidas limitam a generalização dos resultados. Além disso, poucos estudos analisaram o papel mediador da atuação docente, da formação continuada ou das condições institucionais de implementação das práticas avaliativas, que são aspectos reconhecidamente decisivos para sua efetividade. Esses vazios sugerem hipóteses a serem exploradas por futuras pesquisas: por exemplo, em que medida o conhecimento didático-pedagógico do professor influencia os efeitos da avaliação formativa? Como as crenças docentes sobre erro e devolutiva impactam a utilização do feedback?

Com base nos resultados obtidos, recomenda-se que novos estudos aprofundem a análise das práticas avaliativas em contextos escolares diversos, com especial atenção à formação dos professores e ao acompanhamento dos efeitos ao longo do tempo. Também seria oportuno comparar estratégias avaliativas por componente curricular, etapa de ensino e nível de complexidade das tarefas propostas, de modo a refinar a compreensão sobre a aplicabilidade de cada abordagem. Por fim, reitera-se que práticas avaliativas com foco pedagógico, quando bem fundamentadas e implementadas de forma reflexiva, podem contribuir não apenas para o avanço no desempenho matemático, mas para a construção de processos de ensino mais conscientes, dialógicos e equitativos.

REFERÊNCIAS

- ALVIAR, Jayson Virtucio; GAMOREZ, Anabel Enriquez. Effects of classroom response system on the achievement and knowledge retention of the students in mathematics. *Cogent Education*, v. 11, n. 1, p. 2323364, 2024. DOI: <https://doi.org/10.1080/2331186X.2024.2323364>.
- ARVATZ, Avivit; DORI, Yehudit Judy. Science and mathematics high school students' perceptions of self-regulated learning. *International Journal of Science and Mathematics Education*, v. 23, p. 1467-1492, 2025. DOI: <https://doi.org/10.1007/s10763-024-10516-9>
- BONATO, Sarah. *Searching the Grey Literature: A Handbook for Searching Reports, Working Papers, and Other Unpublished Research*. Lanham, MD: Rowman & Littlefield; Medical Library Association, 2018.
- BORENSTEIN, Michael; HEDGES, Larry V.; HIGGINS, Julian P.T.; ROTHSTEIN, Hannah R. *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons, Ltd., 2009.
- BLACK, Paul; WILIAM, Dylan. Assessment and classroom learning. *Assessment in Education*, Abingdon, v. 5, n. 1, p. 7-74, 1998. DOI: <https://doi.org/10.1080/0969595980050102>.
- BLACK, Paul; WILIAM, Dylan. Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, Dordrecht, v. 21, n. 1, p. 5–31, fev. 2009. DOI: <https://doi.org/10.1007/s11092-008-9068-5>.
- BLACK, Paul; WILIAM, Dylan. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, v. 92, n. 1, p. 81–90, 2010.
- CAPINDING, Andie Tangonan. Revolutionizing pre-calculus education: Photomath's AI-powered mathematics tutorship. *Problems of Education in the 21st Century*, v. 81, n. 6, p. 758-775, 2023. DOI: <https://doi.org/10.33225/pec/23.81.758>.
- COSTA, Juliana Evaristo. *Narrativas interativas gamificadas (NIGs) aplicadas aos processos de ensino e aprendizagem de matemática financeira*. 2024. 102 f. Dissertação (Mestrado Profissional em Tecnologia Educacional) - Universidade Federal do Ceará, Instituto Universidade Virtual, Fortaleza, 2024.
- FERNANDES, Domingos. Para uma teoria da avaliação formativa. *Revista da Faculdade de Educação*, v. 32, n. 2, p. 219-236, 2006.
- FIGUEIRA, Priscila Virgínia Salles Teixeira; FREITAS, Patrícia Martins de. Relação entre Ansiedade Matemática, Memória de Trabalho e Controle Inibitório: uma meta-análise. *Bolema: Boletim de Educação Matemática*, Rio Claro, SP, v. 34, n. 67, p. 678-696, ago. 2020. DOI: <http://dx.doi.org/10.1590/1980-4415v34n67a16>.
- FRANCIS, Gregory. Replication, statistical consistency, and publication bias. *Journal of Mathematical Psychology*, v. 57, p. 153-169, 2013.
- GUMIERO, Bárbara Silva; PAZUCH, Vinícius. Teachers' knowledge mobilized in geometry lessons and contingency situations. *International Electronic Journal of Mathematics Education*, v. 16, n. 1, p. em0620, 2021. DOI: <https://doi.org/10.29333/iejme/9371>.

HATTIE, John. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. [S.l.]: Routledge, 2008. DOI: <https://doi.org/10.4324/9780203887332>.

HATTIE, John; TIMPERLEY, Helen. The power of feedback. *Review of Educational Research*, Washington, DC, v. 77, n. 1, p. 81-112, 2007. DOI: <https://doi.org/10.3102/003465430298487>.

HEDGES, Larry V.; OLKIN, Ingram. *Statistical methods for meta-analysis*. New York: Academic Press, 1985.

HIGGINS, Julian P. T.; THOMPSON, Simon G.; DEEKS, Jonathan J.; ALTMAN, Douglas G. Measuring inconsistency in meta-analyses. *British Medical Journal*, v. 327, n. 7414, p. 557-560, 6 set. 2003. doi:10.1136/bmj.327.7414.557

HOFFMANN, Jussara. *Avaliar para promover: as setas do caminho*. 4. ed. Porto Alegre: Mediação, 2001.

KARTONO; SHORA, Rena Yuliantika. Effectiveness of Process Oriented Guided Inquiry Learning with Peer Feedback on Achieving Students' Mathematical Reasoning Capabilities. *International Journal of Instruction*, v. 13, n. 3, p. 555-570, 2020. DOI: <https://doi.org/10.29333/iji.2020.13338a>.

KÜLTÜR, Yusuf Ziya; KUTLU, Mahmut Oğuz. The effect of formative assessment on high school students' mathematics achievement and attitudes. *Journal of Pedagogical Research*, v. 5, n. 4, p. 155-171, 2021. DOI: <https://doi.org/10.33902/JPR.2021474302>.

MORRIS, Scott; DESHON, Richard. Combining effect size estimates in meta-analysis with repeated measures and independent-group designs. *Psychological methods*, v. 7, p. 105-125, 2002. DOI: <https://doi.org/10.1037/1082-989X.7.1.105>.

PAEZ, Arsenio. Gray literature: an important resource in systematic reviews. *Journal of Evidence-Based Medicine*, [s.l.], v. 10, n. 3, p. 233-240, 2017. DOI: <https://doi.org/10.1111/jebm.12266>.

PAGE, Matthew J.; MCKENZIE, Joanne E.; BOSSUYT, Patrick M.; BOUTRON, Isabelle; HOFFMANN, Tammy C.; MULROW, Cynthia D.; SHAMSEER, Larissa; TETZLAFF, Jennifer M.; AKL, Elie A.; BRENNAN, Sue E.; CHOU, Roger; GLANVILLE, Julie; GRIMSHAW, Jeremy M.; HRÓBJARTSSON, Asbjörn; LALU, Manoj M.; LI, Tianjing; LODER, Elizabeth W.; MAYO-WILSON, Evan; MCDONALD, Steve; MCGUINNESS, Luke A.; STEWART, Lesley A.; THOMAS, James; TRICCO, Andrea C.; WELCH, Vivian A.; WHITING, Penny; MOHER, David. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, v. 372, n. 71, p. 1-8, 2021. DOI: <https://doi.org/10.1136/bmj.n71>.

PERRENOUD, Philippe. *Avaliação: da excelência à regulação das aprendizagens*. Porto Alegre: Artmed, 1999.

R CORE TEAM. R: A language and environment for statistical computing (versão 4.4) [Software de computador]. 2024. Disponível em: <https://cran.r-project.org>. Acesso em: 1 jul. 2025.

RODRÍGUEZ-MARTÍNEZ, José Antonio; GONZÁLEZ-CALERO, José Antonio; DEL OLMO-MUÑOZ, Javier; ARNAU, David; TIRADO-OLIVARES, Sergio. Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions. *British Journal of Educational Technology*, v. 54, n. 1, p. 76-97, 2023. DOI: <https://doi.org/10.1111/bjet.13292>.

SADLER, D. Royce. Formative assessment and the design of instructional systems. *Instructional Science*, Dordrecht, v. 18, p. 119-144, 1989. DOI: <https://doi.org/10.1007/BF00117714>.

SCIPPO, Stefano; MADIAI, Serena; CUOMO, Stefano. Digital tessellation for geometry learning in primary school: A quasi-experimental study. *Journal of Information Technology Education: Research*, v. 24, art. 9, 2025. DOI: <https://doi.org/10.28945/5473>.

SLAVIN, Robert E. Evidence-Based Reform in Education: What Will It Take? *European Educational Research Journal*, v. 7, n. 1, p. 124-128, 2008. DOI: <https://doi.org/10.2304/eerj.2008.7.1.124>.

THE JAMOVI PROJECT. *Jamovi* (versão 2.6) [Software de computador]. 2024. Disponível em: <https://www.jamovi.org>. Acesso em: 1 jul. 2025.

VIECHTBAUER, Wolfgang. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, v. 36, p. 1-48, 2010.

CONTRIBUIÇÃO DOS AUTORES

Autor 1: Atuou na conceitualização; curadoria de dados; análise formal; investigação; metodologia; administração do projeto; software; visualização; redação – preparação do rascunho original; redação – revisão e edição.

Autor2: Participou da conceitualização; curadoria de dados; investigação; metodologia; visualização; redação – preparação do rascunho original.

Autor 3: Fez a supervisão; validação; redação – revisão e edição.

DISPONIBILIDADE DE DADOS

Este estudo utilizou apenas dados secundários provenientes de textos publicados. Todos os estudos incluídos na revisão estão listados nas referências.

CONFLITO DE INTERESSES

Os autores declaram não haver conflito de interesse.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.