



AVALIAÇÃO FORMATIVA E AUTORIA ESTUDANTIL COM INTELIGÊNCIA ARTIFICIAL NA EDUCAÇÃO BÁSICA

Leonardo Lima Lopes

leonardolima1722@gmail.com

Doutor em Educação pela Saint Alcuin of York Anglican College

ISSN 3086-8033 · Vol. 2 (2024) · Publicado em 20 de maio de 2024 · Artigo de Revisão · Ciências Humanas

RESUMO

Este artigo analisa a avaliação da aprendizagem e a autoria estudantil em atividades mediadas por inteligência artificial na Educação Básica. Adota revisão narrativa com análise de estudos experimentais e referenciais sobre aprendizagem, diálogo e educação digital. O problema central consiste na possibilidade de produtos bem elaborados não representarem compreensão ou participação intelectual do estudante. Pesquisas vinculadas a Oxford, Harvard, Cambridge, Columbia e Stanford são discutidas segundo seus contextos e limites, sem equiparar resultados universitários e escolares. Propõe-se avaliar processos por meio de produções iniciais, justificativas de revisão e tarefas de transferência, preservando recursos de acessibilidade. A autoria é compreendida como responsabilidade pelas escolhas e pelos argumentos, com explicitação da assistência utilizada. Conclui-se que a avaliação precisa produzir evidências sobre o que o estudante consegue explicar e realizar, em vez de depender apenas do acabamento do produto ou de suspeitas acerca de sua origem.

Palavras-chave: Avaliação formativa. Autoria estudantil. Inteligência artificial. Aprendizagem. Educação Básica.

ABSTRACT

This article examines formative assessment and student authorship in AI-mediated activities in basic education. A narrative review addresses the difference between polished outputs and evidence of student understanding. Research associated with Oxford, Harvard, Cambridge, Columbia, and Stanford is interpreted within its methodological limits. The proposed approach combines initial work, explanations of revisions, and transfer tasks while retaining accessibility supports. Authorship involves responsibility for decisions and arguments and disclosure of assistance. Assessment should establish what students can explain and accomplish rather than rely solely on final products or suspicions about their origin.

Keywords: Formative assessment. Student authorship. Artificial intelligence. Learning.

1. INTRODUÇÃO

Um texto correto pode resultar de compreensão, revisão orientada ou simples aceitação de uma resposta externa. Na presença de sistemas generativos, essas possibilidades ficam menos distinguíveis quando a avaliação considera somente o produto final. O problema não é exclusivo da IA: toda tarefa pode receber ajuda, mas a rapidez e a abrangência da assistência ampliam a distância possível entre aparência e aprendizagem. A avaliação precisa tornar essa distância examinável.

O objetivo deste artigo é discutir como observar aprendizagem e autoria em atividades apoiadas por inteligência artificial na Educação Básica. A questão orientadora é quais evidências permitem distinguir execução assistida e apropriação do conhecimento. Sustenta-se que a avaliação deve acompanhar decisões e revisões, sem converter a sala de aula em ambiente de suspeita permanente. O foco é a qualidade da informação pedagógica produzida para orientar o ensino.

Autoria é definida como responsabilidade progressiva por escolhas, argumentos e formas de apresentação. Ela pode envolver colaboração e recursos de apoio, desde que o estudante compreenda o que apresenta e identifique contribuições relevantes. Essa definição evita equiparar autoria a trabalho solitário. Também impede que a simples edição superficial de um texto gerado seja tomada como demonstração suficiente de participação intelectual.

2. PROCEDIMENTOS DE ANÁLISE

O estudo constitui revisão narrativa, realizada em outubro de 2026 por consulta a publicações científicas e fontes institucionais.

As expressões de busca abrangeram avaliação, aprendizagem com tutoria, IA generativa, erros e diálogo. Foram selecionados trabalhos que ajudassem a distinguir o resultado observado de sua interpretação. A seleção é intencional e não constitui levantamento exaustivo de intervenções ou catálogo de ferramentas de detecção.

O corpus reúne pesquisas experimentais recentes, um estudo sociotécnico, uma revisão psicológica e uma contribuição conceitual sobre pedagogia. A leitura considerou população, tipo de tarefa, condição de comparação e momento da avaliação. A pré-publicação de Stanford foi consultada em sua primeira versão e não é apresentada como artigo de periódico. O estudo de Harvard é universitário e recebe delimitação explícita ao informar a discussão escolar.

A síntese foi organizada em torno de validade das evidências, autoria e desenho de tarefas. Os procedimentos propostos são analíticos e não foram testados neste trabalho. Não foram coletadas produções de estudantes nem estimadas taxas de fraude. O artigo discute como ampliar a informação disponível ao professor e reconhece que nenhuma evidência isolada permite reconstituir integralmente o processo de aprendizagem.

3. O QUE UMA TAREFA PERMITE INFERIR

Uma avaliação só pode sustentar conclusões compatíveis com suas condições. Se a tarefa permite assistência integral, um bom resultado pode indicar uso competente de recursos, mas não demonstra domínio independente do conteúdo. Se o objetivo é avaliar colaboração com ferramentas, essa condição pode ser apropriada. Quando a finalidade é verificar compreensão individual, são necessárias outras oportunidades de explicação. A validade depende da coerência entre objetivo e inferência.

Eynon e Young (2021), com participação de pesquisadora de Oxford, mostram que diferentes atores constroem sentidos e finalidades para a IA educacional. Essa análise sociotécnica não estuda instrumentos avaliativos escolares. Sua pertinência aqui está em exigir clareza sobre o que a instituição reconhece como sucesso.

A escolha de indicadores pode alterar o trabalho pedagógico. Se a plataforma valoriza a quantidade de atividades concluídas, professores e estudantes podem organizar esforços em torno da conclusão. Isso não prova que o indicador cause perda de aprendizagem, mas revela uma hipótese a verificar. A escola deve perguntar se o avanço registrado corresponde a explicações mais consistentes, reconhecimento de relações e capacidade de enfrentar novas situações.

Kestin et al. (2025), em Harvard, realizaram um experimento com 194 estudantes de Física universitária e compararam tutoria estruturada de IA e aprendizagem ativa em aula. Houve vantagem da condição de IA nas medidas investigadas. Esse resultado é circunscrito ao desenho e não demonstra emancipação ou eficácia universal na Educação Básica.

Bastani et al. (2025), no Ensino Médio, observaram diferença entre melhora assistida e desempenho posterior sem o recurso em uma modalidade sem salvaguardas. O estudo reforça a necessidade de separar as condições de prática e de avaliação. Não implica que toda ajuda seja prejudicial ou que todas as tarefas devam ocorrer sem ferramentas.

Propõe-se, portanto, explicitar duas perguntas em cada sequência: o que o aluno consegue realizar com apoio e o que consegue reconstruir depois, preservadas as adaptações de acesso. As respostas não precisam coincidir. Uma diferença pode indicar aprendizagem ainda em desenvolvimento e orientar nova intervenção. Seu uso adequado é formativo, não uma justificativa automática para punição ou para retirada de assistência necessária.

4. ERRO E DEVOLUTIVA COMO EVIDÊNCIAS DE PROCESSO

Metcalfe (2017), de Columbia, discute benefícios de erros acompanhados de correção e análise do raciocínio. A revisão não investiga ferramentas generativas. Para a avaliação, sua contribuição é sustentar o interesse pedagógico de observar tentativas e mudanças, em vez de registrar apenas o acerto final.

Uma atividade pode solicitar uma resposta breve antes de qualquer consulta. Após o estudo, o estudante apresenta nova versão e explica uma mudança relevante. O professor examina se a justificativa aborda o conteúdo ou apenas descreve uma alteração de palavras. A primeira produção não precisa receber nota classificatória: seu valor está em oferecer uma referência para compreender o percurso e planejar o apoio posterior.

O registro deve ser proporcional à finalidade. Exigir históricos completos de navegação ou transcrições extensas pode aumentar a exposição de dados e a carga de trabalho sem melhorar a interpretação. Em muitos casos, basta preservar um trecho inicial, uma intervenção significativa e uma explicação do estudante. A avaliação precisa produzir evidências úteis, não acumular informações que ninguém terá condições de analisar.

Wang et al. (2024), em pré-publicação de Stanford, investigam IA dirigida a tutores humanos e relatam melhora no domínio de tópicos. O resultado não valida correção automatizada ou atribuição automática de notas. Ele permite discutir assistência ao educador como configuração distinta da resposta entregue diretamente ao aluno.

O professor pode utilizar um sistema para gerar hipóteses de devolutiva e escolher a que considera adequada, após revisão. A recomendação deve apontar a relação a ser examinada, sem impor uma reescrita completa que retire do estudante a oportunidade de

decidir. Para avaliar a contribuição, pode-se observar se a nova versão resolve o problema identificado e se o aluno consegue explicar por que a mudança foi necessária.

Também é preciso admitir incerteza na interpretação. Um estudante pode compreender oralmente e registrar pouco por escrito, ou reproduzir uma explicação memorizada sem aplicá-la. Por isso, produções de modalidades diferentes podem complementar-se. O acompanhamento não deve exigir desempenho idêntico de comunicação, mas conservar o conceito e o nível de raciocínio que a atividade se propõe a avaliar.

5. AUTORIA E EXPLICITAÇÃO DA ASSISTÊNCIA

Wegerif e Casebourne (2026), de Cambridge, propõem integrar IA em uma perspectiva dialógica de educação. O trabalho é conceitual. Para este artigo, sua contribuição consiste em manter a participação no debate como finalidade formativa, sem reduzir aprendizagem à elaboração de um artefato individual.

A autoria pode ser acompanhada por uma declaração breve e adequada à idade: que ajuda foi recebida, em qual etapa e quais decisões permaneceram com o estudante. Essa declaração não funciona como prova absoluta, mas estabelece linguagem comum para discutir colaboração. A escola precisa esclarecer previamente o que é permitido. Não é coerente incentivar o uso de IA no planejamento e depois tratar qualquer assistência como infração sem distinção de finalidade.

Em uma produção argumentativa, pode ser permitido solicitar sugestões de perguntas e vedado delegar a construção integral do argumento. Em outra atividade, a turma pode analisar criticamente um texto gerado pelo professor. São tarefas com objetivos diferentes e critérios próprios. O contrato pedagógico deve ser compreensível, com exemplos do que conta como apoio, colaboração e substituição indevida, sem depender de regras implícitas.

A avaliação da autoria pode incluir a escolha de uma fonte e a explicação de sua pertinência. O estudante identifica o que o material permite afirmar e o que permanece sem sustentação. Uma referência sugerida por IA só deve ser utilizada após verificação. Quando o aluno não encontra o documento, a ausência pode tornar-se objeto de aprendizagem sobre limites do sistema. A aparência formal da citação não estabelece sua existência.

Miao e Holmes (2023) defendem agência humana e validação das aplicações educacionais. Essa orientação apoia a proposta de declarar a assistência e revisar os resultados, mas não equivale a um protocolo de comprovação de autoria. A escola continua responsável por construir procedimentos compatíveis com seus estudantes e suas finalidades.

Este artigo não recomenda inferir fraude a partir de estilo impessoal, vocabulário sofisticado ou erro incomum. Tais características não identificam, por si, a origem de uma produção. Diante de uma dúvida, a proposta é solicitar explicação, apresentar nova situação e reunir registros do percurso. Essa alternativa privilegia informação pedagógica e possibilidade de resposta do estudante, em vez de atribuição de culpa por aparência textual.

6. DESENHO DE TAREFAS E JUSTIÇA AVALIATIVA

Propõe-se uma sequência com três conjuntos de evidência: produção inicial, revisão comentada e aplicação a uma situação diferente. No primeiro, o professor identifica conhecimentos e dúvidas. No segundo, observa como o estudante utiliza apoios e justifica mudanças. No terceiro, verifica se reconhece o conceito em novo contexto. A sequência não elimina todas as incertezas, mas reduz a dependência do produto final isolado.

Em Ciências, uma turma pode explicar inicialmente por que certos materiais mudam de estado. Após estudar fontes e examinar uma resposta automatizada, os alunos revisam suas explicações. A tarefa final apresenta outra situação, com materiais e condições diferentes. O professor avalia relações conceituais, não apenas a presença de termos técnicos. A consulta pode ser permitida em algumas etapas e delimitada em outras, conforme o objetivo declarado.

Uma rubrica pode examinar precisão das relações, pertinência das evidências e capacidade de revisão. Os descritores devem mostrar o que diferencia uma explicação incompleta de outra mais elaborada. Recomenda-se evitar critérios que premiem apenas extensão, acabamento visual ou linguagem excessivamente formal. Esses elementos podem variar com acesso a ferramentas e não representam necessariamente a aprendizagem investigada.

A tarefa independente deve conservar recursos de acessibilidade. Um estudante que utiliza comunicação alternativa, leitor de tela ou apoio de registro não perde autoria por essa condição. A retirada do recurso pode alterar o objeto avaliado e produzir uma dificuldade de acesso indevidamente interpretada como lacuna de conhecimento. O limite relevante é a delegação do raciocínio central, e não a presença de qualquer auxílio.

As condições de acesso também precisam ser equivalentes. Se parte da turma dispõe de versões pagas e outra depende de conexão instável, a tarefa domiciliar pode incorporar diferenças externas ao objetivo. A escola pode oferecer momentos de realização com recursos comuns ou alternativas comparáveis. A avaliação deve documentar o que foi disponibilizado, para não atribuir ao esforço individual diferenças que o próprio desenho introduziu.

A viabilidade depende da carga de trabalho. Acompanhar processos não exige comentar integralmente todas as versões de todos os textos. O professor pode escolher pontos decisivos, utilizar devolutivas coletivas e alternar o foco de acompanhamento. A instituição precisa assegurar condições para esse trabalho. A automação pode auxiliar na organização, mas a decisão avaliativa exige critérios e possibilidade de revisão humana.

7. INTERPRETAÇÃO CONJUNTA DE DIFERENTES EVIDÊNCIAS

Considere-se uma atividade em que o estudante apresenta um texto final consistente, mas não explica a relação entre duas afirmações centrais. Esse contraste não permite concluir imediatamente que houve fraude ou ausência total de aprendizagem. Pode indicar participação limitada naquela parte, dificuldade de expressão ou compreensão instável. A resposta formativa é delimitar a relação e oferecer outra oportunidade de explicação. O professor precisa registrar a hipótese que está investigando e evitar que uma impressão global determine a interpretação de todas as produções posteriores.

Em outra situação, um texto pode ter problemas de organização, enquanto a explicação oral revela relações adequadas e uso pertinente de fontes. A avaliação deve separar o domínio conceitual e as necessidades de escrita, caso ambos façam parte dos objetivos. Uma única nota sem critérios pode ocultar essa diferença. Quando a IA auxilia na revisão linguística, o professor pode solicitar que o estudante identifique quais alterações modificaram apenas a forma e quais afetaram o argumento. O procedimento torna a assistência discutível sem pressupor que toda revisão comprometa a autoria.

É possível trabalhar com descritores qualitativos relacionados à tarefa. Uma justificativa pode apenas afirmar que a resposta está correta, indicar um exemplo sem relacioná-lo à conclusão ou estabelecer uma relação explícita entre evidência e argumento. Esses níveis descrevem produções, não identidades fixas dos estudantes. Sua utilização precisa ser acompanhada de exemplos e possibilidade de revisão. Não se propõe uma escala universal de pensamento crítico, nem se presume que o mesmo descritor mantenha significado idêntico em todas as disciplinas e etapas escolares.

A confiança declarada pelo estudante pode complementar a explicação, mas não substituí-la. Perguntar onde há dúvida pode orientar a devolutiva e revelar a necessidade de retomar uma relação. A instituição não deve utilizar a autoavaliação para punir quem reconhece dificuldade. Se declarar incerteza produz desvantagem, o procedimento pode estimular respostas defensivas e perder utilidade pedagógica. A autoavaliação precisa estar associada a oportunidades reais de estudo e a critérios que reconheçam a revisão fundamentada como parte do processo.

Por fim, a decisão avaliativa deve considerar o conjunto das evidências e explicitar o que permanece desconhecido. Um histórico de versões não prova toda a aprendizagem, assim como uma explicação satisfatória em um dia não demonstra retenção duradoura. Acompanhamentos em momentos posteriores podem informar a continuidade do ensino. As situações desta seção são exemplos analíticos, não casos coletados. Seu propósito é mostrar como hipóteses alternativas e oportunidades de resposta podem reduzir inferências excessivas, mantendo a avaliação vinculada à compreensão e ao desenvolvimento do estudante.

8. CONSIDERAÇÕES FINAIS

A presença da IA exige explicitar o que as tarefas demonstram sobre aprendizagem e autoria. Um produto bem elaborado continua relevante, mas é insuficiente quando sua produção pode ser amplamente delegada. A resposta pedagógica proposta é reunir evidências de tentativa, revisão e transferência, preservando condições de participação. Não se trata de excluir ferramentas de todas as atividades, mas de tornar interpretáveis as situações em que são utilizadas.

Autoria implica responsabilidade progressiva pelas escolhas e pelos argumentos, com reconhecimento da assistência. Regras prévias, exemplos claros e oportunidades de explicação oferecem condições mais consistentes do que suspeitas baseadas em aparência. O professor precisa distinguir o apoio que permite acesso daquele que substitui a operação intelectual avaliada.

A revisão não testa a eficácia do desenho proposto e reúne estudos de diferentes níveis educacionais. Investigações futuras podem comparar sua viabilidade, a qualidade das evidências produzidas e a carga de trabalho. O critério de êxito deve ser a capacidade da avaliação de orientar o ensino e tornar a aprendizagem visível, sem ampliar exclusão ou vigilância desnecessária.

REFERÊNCIAS

BASTANI, Hamsa; BASTANI, Osbert; SUNGU, Alp; GE, Haosen; KABAKCI, Özge; MARIMAN, Rei. **Generative AI without guardrails can harm learning: Evidence from high school mathematics**. Proceedings of the National Academy of Sciences, v. 122, n. 26, e2422633122, 2025. DOI: <https://doi.org/10.1073/pnas.2422633122>.

EYNON, Rebecca; YOUNG, Erin. **Methodology, Legend, and Rhetoric: The Constructions of AI by Academia, Industry, and Policy Groups for Lifelong Learning**. Science, Technology, & Human Values, v. 46, n. 1, p. 166-191, 2021. DOI: <https://doi.org/10.1177/0162243920906475>.

KESTIN, Greg; MILLER, Kelly; KLALES, Anna; MILBOURNE, Timothy; PONTI, Gregorio. **AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting**. Scientific Reports, v. 15,

art. 17458, 2025. DOI: <https://doi.org/10.1038/s41598-025-97652-6>.

METCALFE, Janet. **Learning from Errors**. Annual Review of Psychology, v. 68, p. 465-489, 2017. DOI: <https://doi.org/10.1146/annurev-psych-010416-044022>.

MIAO, Fengchun; HOLMES, Wayne. **Guidance for generative AI in education and research**. Paris: UNESCO, 2023. Disponível em: <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>. Acesso em: 16 out. 2026.

WANG, Rose E.; RIBEIRO, Ana T.; ROBINSON, Carly D.; LOEB, Susanna; DEMSZKY, Dora. **Tutor CoPilot: A Human-AI Approach for Scaling Real-Time Expertise**. arXiv, versão 1, 2024. Pré-publicação. Disponível em: <https://arxiv.org/abs/2410.03017v1>. Acesso em: 16 out. 2026.

WEGERIF, Rupert; CASEBOURNE, Imogen. **A dialogic theoretical foundation for integrating generative AI into pedagogical design**. British Journal of Educational Technology, v. 57, n. 3, p. 639-654, 2026. DOI: <https://doi.org/10.1111/bjet.70026>.



CC BY-NC-ND 4.0

Este artigo está licenciado sob uma Licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 Internacional (CC BY-NC-ND 4.0).