

---

# Adoção acelerada, evidência escassa: a hierarquia de evidências e a inteligência artificial na educação básica

---

*Quatro perguntas, quatro tipos de prova e critérios para decidir em redes e escolas*

Marcos Cordioli

-----  
Memorando de pesquisa de Marcos Cordioli, integrante da pesquisa por ele conduzida, A aprendizagem na Era da Inteligência Artificial, que reúne estudos, textos e demais produções sobre inteligência artificial e educação e sobre inteligência artificial e aprendizagem. A análise registrada reflete o estado do conhecimento no momento da redação e permanece aberta à revisão diante de novas evidências.  
-----

## Resumo

A adoção da inteligência artificial nas escolas avança mais depressa do que a pesquisa capaz de avaliá-la. A análise de quatro materiais publicados entre 2025 e 2026 mostrou que cada um responde a uma pergunta distinta, pois a percepção docente informa a aceitação, o ensaio randomizado informa a eficácia comparada em contexto delimitado, o relato institucional informa a viabilidade e a revisão do campo informa o tamanho da lacuna. Um estudo experimental complementar, sobre tutores com e sem salvaguardas, demonstrou que o desenho da ferramenta modifica o efeito sobre a aprendizagem. Propõe-se um conjunto de critérios para que redes e escolas decidam sobre a adoção com proporcionalidade entre a força da decisão e a força da prova disponível.

Palavras-chave: Inteligência artificial; Hierarquia de evidências; Educação básica; Políticas educacionais; Decisão baseada em evidências.

## Abstract

The adoption of artificial intelligence in schools is moving faster than the research capable of evaluating it. An analysis of four materials published between 2025 and 2026 showed that each answers a different question: teacher perception informs acceptance, the randomized trial informs comparative efficacy in a bounded context, the institutional report informs feasibility, and the field review informs the size of the gap. A complementary experimental study on tutors with and without safeguards showed that tool design changes the effect on learning. A set of criteria is proposed so

that school systems and schools can decide on adoption with proportionality between the weight of the decision and the strength of the available proof.

Keywords: Artificial intelligence; Evidence hierarchy; Basic education; Education policy; Evidence-based decision-making.

Introdução

A inteligência artificial chegou às escolas antes de chegar aos laboratórios de avaliação. Segundo levantamento da RAND de 2025, citado por Barnard, Fesler e Loeb (2026), 54% dos estudantes dos anos finais do ensino fundamental e do ensino médio (middle e high school) nos Estados Unidos declararam usar a tecnologia em tarefas escolares, e 53% dos professores de língua inglesa, matemática e ciências a utilizaram em tarefas de ensino.

Os mesmos autores retomaram revisão recente que identificou mais de 800 trabalhos relevantes sobre inteligência artificial na educação básica, dos quais apenas vinte configuravam estudos causais de alta qualidade, e nenhum deles tratava de ferramentas voltadas ao estudante em escolas de educação básica estadunidenses. A defasagem entre a velocidade da adoção e a da pesquisa coloca gestores, professores e famílias diante de decisões que não podem esperar a evidência ideal.

A pergunta que orienta a análise indaga como decidir com racionalidade quando a evidência é escassa e heterogênea. A resposta começa pela distinção entre os tipos de prova, pois cada desenho de pesquisa responde a uma pergunta específica, e a decisão responsável associa cada pergunta ao tipo de prova que lhe corresponde.

1 | Quatro perguntas, quatro tipos de prova

Os quatro materiais analisados ilustram quatro desenhos de evidência. Um estudo da revista Education Sciences (RUIZ VIRUEL; SÁNCHEZ RIVAS; RUIZ PALMERO, 2025) mediu a percepção de 300 professores. Um ensaio randomizado cruzado da Scientific Reports (KESTIN et al., 2025) mediu o desempenho de 194 estudantes de física de Harvard. Um relatório do Hunt Institute (HART, 2025) descreveu o desenho institucional da Alpha School. A publicação de Barnard, Fesler e Loeb (2026) sintetizou o estado do campo. A Tabela 1 associa cada desenho à pergunta a que responde e aos limites do que permite afirmar.

Tabela 1 – Perguntas de decisão e tipos de prova correspondentes

| Pergunta de decisão                        | Tipo de prova     | Fonte ilustrativa                                | O que permite afirmar                                                | O que não permite afirmar                                       |
|--------------------------------------------|-------------------|--------------------------------------------------|----------------------------------------------------------------------|-----------------------------------------------------------------|
| A prática é bem aceita por quem a executa? | Percepção docente | Ruiz Viruel, Sánchez Rivas e Ruiz Palmero (2025) | Os professores avaliaram melhor a versão com inteligência artificial | Que os estudantes aprenderam mais ou se tornaram mais autônomos |
| A ferramenta                               | Ensaio            | Kestin et al.                                    | Maior ganho com                                                      | Que o efeito se                                                 |

|                                             |                      |                               |                                                               |                                                               |
|---------------------------------------------|----------------------|-------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|
| produz mais aprendizagem que a alternativa? | randomizado          | (2025)                        | tutor em conteúdo delimitado, em menos tempo                  | mantém em problemas abertos, em outras etapas e a longo prazo |
| O modelo é operacionalmente possível?       | Relato institucional | Hart (2025)                   | Que a organização do tempo em dois blocos funciona na prática | Que o resultado se deve à inteligência artificial             |
| Quanta evidência confiável existe?          | Revisão do campo     | Barnard, Fesler e Loeb (2026) | A dimensão da lacuna e um critério de avaliação               | Nada sobre o efeito de uma ferramenta específica              |

Fonte: elaboração própria, a partir das publicações citadas.

## 2 | Percepção docente e a pergunta sobre a aceitação

O estudo de Ruiz Viruel, Sánchez Rivas e Ruiz Palmero (2025) aplicou questionário em escala de cinco pontos a 300 professores que, por cinco meses, usaram a aprendizagem baseada em projetos em formato tradicional e, um ano depois, por mais cinco meses, em formato integrado à inteligência artificial. A versão com a tecnologia obteve média de 4,17, contra 3,49 da versão tradicional, com diferença estatisticamente significativa e tamanho de efeito grande ( $r = 0,56$ ).

O resultado informa que a prática encontrou boa aceitação entre os docentes, dado relevante para qualquer processo de implementação. Não informa, contudo, quanto os estudantes aprenderam, porque o instrumento mede impressão. O desenho sequencial, em que os mesmos professores avaliaram as duas condições em momentos separados por um ano, acrescenta a possibilidade de que a experiência acumulada e o entusiasmo com a novidade influenciem a avaliação. Os autores registraram ainda a infraestrutura insuficiente, as lacunas de formação digital, o risco de dependência excessiva e as preocupações com a privacidade dos dados.

## 3 | O ensaio randomizado e a pergunta sobre a eficácia

O ensaio de Kestin et al. (2025) ofereceu o tipo de prova mais robusto do corpus, pois estabeleceu causalidade por comparação direta, na mesma população, entre tutoria por inteligência artificial e aula ativa presencial. A pontuação mediana de pós-teste foi de 4,5 com o tutor e de 3,5 em sala, com ganhos de aprendizagem que superaram o dobro, tamanhos de efeito entre 0,63 e 1,3 desvio-padrão e mediana de 49 minutos de estudo.

A robustez causal convive com validade externa restrita. Os próprios autores advertiram que os resultados podem não valer em contextos que exigem síntese complexa de múltiplos conceitos e pensamento crítico de ordem superior. O experimento envolveu estudantes universitários, dois conteúdos de física, uma

---

instituição e um período de duas semanas. A decisão de uma rede de ensino sobre estudantes mais jovens, em componentes curriculares diversos e por anos letivos inteiros, extrapola o que o desenho autoriza.

O estudo de campo de Bastani et al. (2025), com quase mil estudantes de matemática do ensino médio, acrescentou ao corpus segundo desenho causal e uma informação decisiva para a adoção. O acesso a um tutor baseado em GPT-4 melhorou o desempenho durante a prática nas duas condições testadas, mas os estudantes que usaram a versão sem salvaguardas tiveram resultado inferior, quando o acesso foi retirado, ao dos que nunca haviam contado com a ferramenta. A versão desenhada para proteger a aprendizagem mitigou em grande parte este efeito. A evidência causal mais próxima da educação básica indica, portanto, que o desenho da ferramenta altera o sinal do efeito.

#### **4 | O relato institucional e a pergunta sobre a viabilidade**

Hart (2025) descreveu a Alpha School, que reserva cerca de duas horas diárias a módulos acadêmicos adaptativos conduzidos por inteligência artificial e dedica as tardes a projetos orientados por professores denominados guias. O relato incluiu o pequeno número de estudantes, as turmas de idades mistas, os planos de aprendizagem elaborados com as famílias e a avaliação por portfólios compartilhados com famílias e professores.

Este tipo de material responde com propriedade à pergunta sobre a viabilidade, pois demonstra que a organização é operacionalmente possível. Não responde à pergunta sobre a eficácia, porque não apresentou dados mensuráveis de resultado, não contou com grupo de comparação e não separa o efeito da inteligência artificial do efeito de variáveis como o porte reduzido, o acompanhamento individualizado e a seleção das famílias que optam pela escola. O próprio relato reconheceu riscos de isolamento do estudante, de viés algorítmico e de barreiras de acessibilidade. O entusiasmo em torno de modelos desta natureza ultrapassa, com frequência, o que o tipo de material comporta.

#### **5 | A revisão do campo e a pergunta sobre a lacuna**

Barnard, Fesler e Loeb (2026) não somaram evidência a favor ou contra a tecnologia. Delimitaram a evidência existente e propuseram critério de avaliação que desloca a pergunta de se a inteligência artificial funciona para se a ferramenta amplia o acesso do estudante a experiências de aprendizagem de eficácia já conhecida, como a instrução personalizada, a resolução de problemas complexos, os projetos colaborativos significativos, a discussão substantiva e a relação duradoura com um educador, ou se as substitui por versão mais rápida e mais pobre.

O critério tem a vantagem de poder ser aplicado antes que a pesquisa causal esteja completa. Ele transforma o conhecimento acumulado sobre as condições da aprendizagem em teste de plausibilidade para cada ferramenta, e coloca a prova do lado de quem propõe a adoção.

---

## 6 | A política pública diante da lacuna

O guia da UNESCO para a inteligência artificial generativa na educação e na pesquisa, publicado em setembro de 2023, adotou abordagem centrada no ser humano, que promove a agência humana, a inclusão, a equidade, a igualdade de gênero e a diversidade cultural e linguística (MIAO; HOLMES, 2023). O documento recomendou aos governos a adoção de normas de proteção de dados e privacidade, a fixação da idade mínima de 13 anos para o uso das ferramentas em sala de aula e a formação dos educadores, e registrou que menos de 10% das escolas e universidades consultadas dispunham de políticas formais sobre o tema.

A convergência entre o guia internacional e a revisão de Stanford indica que a escassez de evidência não justifica nem a paralisia nem a adoção sem critério. Aponta para a regulação proporcional, que combina proteção de dados, formação docente e avaliação contínua dos efeitos.

## 7 | Critérios para decidir em redes e escolas

A análise sustenta uma proposta de cinco critérios, que expressam a inferência do autor a partir das fontes e não constituem resultado das pesquisas citadas. O primeiro critério é a proporcionalidade, segundo a qual a força da prova exigida cresce com a abrangência e a irreversibilidade da decisão. Um piloto em uma turma admite prova mais frágil do que a contratação de uma plataforma para toda a rede.

O segundo critério associa a pergunta ao tipo de prova. Quando o fornecedor afirma que a ferramenta melhora a aprendizagem, a resposta adequada é um estudo com grupo de comparação, e a satisfação de usuários ou o relato de uma escola-modelo não atende à exigência. O terceiro critério examina o desenho da ferramenta quanto a salvaguardas, o que inclui verificar se o sistema sustenta o raciocínio do estudante ou entrega a resposta pronta, e se há indícios de dependência quando o acesso é retirado.

O quarto critério preserva as experiências de eficácia comprovada. A adoção deve ampliar o acesso do estudante à instrução personalizada, à resolução de problemas complexos, aos projetos colaborativos, à discussão e à relação com o educador, e não substituir estas experiências. O quinto critério exige que a implementação seja acompanhada de formação docente, proteção de dados e avaliação do percurso, com métricas de aprendizagem definidas antes do início e comparação com grupo ou período de referência.

## 8 | Limitações

O corpus é reduzido, não resultou de busca sistemática e concentra-se no contexto estadunidense e em língua inglesa. Os critérios propostos não foram testados empiricamente e dependem de adaptação aos sistemas de ensino brasileiros, cujas condições de infraestrutura, formação docente e financiamento diferem das registradas nas fontes. A transposição dos resultados ao país exige estudos locais.

---

## Considerações finais

A hierarquia de evidências não ensina a descartar os materiais mais frágeis, e sim a atribuir a cada um a pergunta que pode responder. A percepção docente informa a aceitação, o ensaio randomizado informa a eficácia em contexto delimitado, o relato institucional informa a viabilidade e a revisão do campo informa a lacuna.

Redes e escolas decidem melhor quando exigem de cada proposta a prova compatível com a decisão que ela sustenta. Em um campo em que a adoção precede a evidência, esta disciplina protege os estudantes, que não devem servir de experimento sem avaliação, e protege também a própria tecnologia, cujo valor depende de ser demonstrado.

## Declaração sobre o uso de inteligência artificial

A inteligência artificial generativa foi utilizada como ferramenta de apoio à organização das informações fornecidas para este artigo. O uso ocorreu como auxílio ao trabalho intelectual, sem substituir a análise dos documentos e das fontes. A seleção e a interpretação das informações, assim como a revisão e a responsabilidade pela versão final do texto, cabem ao autor. Ao final, a IA foi utilizada para a correção ortográfica e semântica do texto e também para gerar o resumo e as palavras-chave.

## Referências

BARNARD, Cristina; FESLER, Lily; LOEB, Susanna. Artificial intelligence in K-12 schools. AEPF Live Handbook, jun. 2026. Disponível em: <https://livehandbook.org/k-12-education/miscellaneous/k-12-education/school-resources/artificial-intelligence-in-k%E2%80%9312-schools/>. Acesso em: 4 out. 2026.

BASTANI, H.; BASTANI, O.; SUNGU, A.; GE, H.; KABAKCI, Ö.; MARIMAN, R. 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: 10.1073/pnas.2422633122. Disponível em: <https://doi.org/10.1073/pnas.2422633122>. Acesso em: 4 out. 2026.

HART, Jaalil. AI tutoring in schools: how personalized learning technology is changing K-12 education in 2025. Washington, DC: The Hunt Institute, 11 jun. 2025. Disponível em: <https://hunt-institute.org/resources/2025/06/ai-tutoring-alpha-school-personalized-learning-technology-k-12-education/>. Acesso em: 4 out. 2026.

KESTIN, G.; MILLER, K.; KLALES, A.; MILBOURNE, T.; PONTI, G. AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports, v. 15, 2025. DOI: 10.1038/s41598-025-97652-6. Disponível em: <https://www.nature.com/articles/s41598-025-97652-6>. Acesso em: 4 out. 2026.

MIAO, Fengchun; HOLMES, Wayne. Guidance for generative AI in education and research. Paris: UNESCO, 2023. Disponível em: <https://discovery.ucl.ac.uk/id/eprint/10176438/>. Acesso em: 4 out. 2026.

---

RUIZ VIRUEL; SÁNCHEZ RIVAS; RUIZ PALMERO. The role of artificial intelligence in project-based learning: teacher perceptions and pedagogical implications. Education Sciences, Basel, v. 15, n. 2, art. 150, 2025. DOI: 10.3390/educsci15020150. Disponível em: <https://www.mdpi.com/2227-7102/15/2/150>. Acesso em: 4 out. 2026.

## **Sobre o autor**

Marcos Cordioli é mestre em Educação pela PUC-SP, palestrante, consultor e empresário. Atua há décadas como pesquisador e professor nas áreas de políticas educacionais, currículo e gestão da educação brasileira, com especial atenção às Diretrizes Curriculares Nacionais e às Diretrizes Operacionais estabelecidas pelo Conselho Nacional de Educação (CNE).

A trajetória de estudos sobre a legislação e as políticas curriculares brasileiras teve início em 1997, com a publicação de sua primeira obra sobre os Parâmetros Curriculares Nacionais e os temas transversais. Desde então, publicou livros e centenas de artigos, além de participar de palestras, cursos, pesquisas e projetos educacionais em diferentes regiões do Brasil e no exterior.

Em 2011 e 2012, atuou como consultor ad hoc da relatoria do Plano Nacional de Educação na Câmara dos Deputados, participando de processos de discussão e sistematização de propostas para o planejamento da educação brasileira.

Foi secretário de Cultura de Curitiba e secretário de Inovação de Maringá. Também atuou como professor em cursos de graduação, especialização e mestrado e como consultor de instituições e sistemas de ensino.

Atualmente, concentra sua principal atividade de pesquisa no projeto "Políticas de Gestão de Currículos e Conteúdos Educacionais no Brasil pós-LDB (1996–atualidade)", dedicado à organização, análise e interpretação da trajetória das políticas curriculares brasileiras, de seus documentos normativos e dos processos de formulação e implementação conduzidos pelo Ministério da Educação, pelo Conselho Nacional de Educação e pelos sistemas de ensino.

Como segunda frente de pesquisa e atuação, dedica-se às transformações da educação diante das novas tecnologias, especialmente às possibilidades de utilização da inteligência artificial nos processos de ensino e aprendizagem, na formação de educadores e na gestão das instituições e dos sistemas de ensino.

Sobre esta temática, Marcos Cordioli é autor, entre outras publicações, de:

CORDIOLLI, Marcos. Currículo Escolar: Teorias e Práticas. Pinhais: Ed. Melo, 2005.

CORDIOLLI, Marcos. Para entender os PCNs: os temas transversais. Curitiba: Módulo, 1997.

CORDIOLLI, Marcos. Sistemas de ensino e políticas educacionais no Brasil. Curitiba: Dialógica, 2011.

---

Contato: [marcos.cordioli@gmail.com](mailto:marcos.cordioli@gmail.com) | (41) 99682-9123 | [@marcoscordioli](https://www.instagram.com/marcoscordioli) | [marcoscordioli.com.br](https://marcoscordioli.com.br) | Linktree <https://linktr.ee/marcos.cordioli>