



**INTELIGÊNCIA ARTIFICIAL NA
AVALIAÇÃO ESCOLAR: PERCEPÇÕES,
ESTRATÉGIAS E FORMAÇÃO**

ARTIFICIAL INTELLIGENCE IN SCHOOL
ASSESSMENT: PERCEPTIONS, STRATEGIES,
AND TEACHER EDUCATION

Bruno de Souza Santos

Especialista no Ensino de Geografia pela Universidade
Estadual Paulista Júlio de Mesquita Filho (UNESP)

E-mail: brunognr@gmail.com | ORCID:

<https://orcid.org/0009-0005-1782-2910>

Jussara do Carmo Badelli

Mestranda em Educação com Especialização em ESG
(Governança, Responsabilidade Social e Meio-
Ambiente), American Global Tech University (AGTU)

E-mail: jbadelli@live.com

Resumo

Este estudo investiga as percepções e os autorrelatos de práticas de professores da educação básica frente ao uso da inteligência artificial (IA) generativa na avaliação escolar. A pesquisa parte da hipótese de que a ausência de mediação formativa estruturada pode polarizar a prática docente. Participaram 40 professores de escolas públicas e privadas de São Paulo, selecionados por amostragem por conveniência, que responderam a um questionário on-line misto. Os dados quantitativos foram tratados por estatística descritiva, e as respostas abertas, por análise de conteúdo temática, com validação por juízes independentes ($\kappa = 0,82$). Os resultados indicam que 78% dos participantes identificaram indícios de uso de IA por estudantes. Desses, 65% usaram detectores, mas apenas 45% desse subgrupo os consideraram confiáveis. A análise cruzada revelou que docentes com até cinco anos de experiência tenderam a usar detectores automatizados, enquanto aqueles com mais de 15 anos priorizaram estratégias humanas, como a oralidade. Identificou-se forte demanda por protocolos éticos institucionais e formação continuada prática. Conclui-se que a universidade deve co-construir, com as redes de ensino, projetos-piloto que transformem a IA em aliada da práxis pedagógica, superando a lógica punitiva.

Palavras-chave: Inteligência artificial generativa. Avaliação formativa. Formação continuada. Prática docente. Autoria.

Abstract

This study investigates the perceptions and self-reported practices of basic education teachers regarding the use of generative artificial intelligence (AI) in school assessment. The research is based on the hypothesis that the absence of structured formative mediation may polarize teaching practice. Forty teachers from public and private schools in São Paulo participated in the

study; they were selected through convenience sampling and answered a mixed online questionnaire. Quantitative data were analyzed using descriptive statistics, while open-ended responses were examined through thematic content analysis, with coding validated by independent judges ($\kappa = 0.82$). The results indicate that 78% of participants had identified signs of AI use by students. Among them, 65% had used AI detectors, but only 45% of this subgroup considered them reliable. Cross-analysis showed that teachers with up to five years of experience tended to use automated detectors, whereas those with more than 15 years prioritized human strategies, such as oral assessment. There was a strong demand for institutional ethical protocols and practical continuing education. The study concludes that universities should co-construct pilot projects with education systems to transform AI into an ally of pedagogical praxis, moving beyond a punitive logic.

Keywords: Generative artificial intelligence. Formative assessment. Continuing education. Teaching practice. Authorship.

1 INTRODUÇÃO

A incorporação de modelos generativos de inteligência artificial (IA) ao cotidiano escolar desafia os fundamentos epistemológicos da avaliação da aprendizagem. Sabe-se que o plágio tradicional já exigia estratégias de verificação de autoria, mas a produção automatizada de textos por algoritmos sofisticados complica o processo avaliativo, dissociando, por vezes, o produto final do esforço cognitivo do estudante. Este artigo trata das percepções e dos autorrelatos de práticas de professores da educação básica sobre o uso da IA em suas práticas avaliativas.

O objetivo é analisar essas percepções e as demandas por formação continuada, empregando uma metodologia de abordagem mista concorrente. A relevância do tema reside na necessidade concreta de subsidiar políticas de formação que respondam às necessidades da sala de aula, evitando tanto a

adoção acrítica quanto a rejeição sumária da tecnologia.

2 MATERIAL E MÉTODOS

2.1 Caracterização da pesquisa

Trata-se de uma pesquisa de abordagem mista concorrente (CRESWELL; PLANO CLARK, 2017), na qual dados quantitativos e qualitativos foram coletados simultaneamente. O estudo combina estatística descritiva e análise temática de respostas abertas, sem pretensão de estabelecer relações causais ou generalizações estatísticas para além do grupo participante.

O instrumento de coleta foi um questionário on-line estruturado em quatro blocos: caracterização profissional, experiência com detecção, uso pessoal e duas questões abertas. Os dados quantitativos foram analisados por estatística descritiva. As respostas abertas foram submetidas à análise de conteúdo temática (BARDIN, 2016), com codificação validada por dois juízes independentes, obtendo-se coeficiente Kappa de Cohen de $\kappa = 0,82$ (LANDIS; KOCH, 1977).

2.2 Área de estudo e público-alvo

Participaram 40 professores da educação básica, recrutados entre janeiro e março de 2026, por meio de divulgação em grupos profissionais de redes sociais e listas de e-mail de secretarias de educação do estado de São Paulo. A taxa de resposta foi de 23%. Os critérios de inclusão foram estar em efetivo exercício da docência e possuir conhecimento sobre ferramentas de IA generativa.

As falas dos participantes citadas neste artigo são tratadas como dados primários de

pesquisa e identificadas por códigos, preservando-se a identidade dos respondentes. No tratamento dos dados, não são apresentados nomes, escolas ou outros elementos que permitam a identificação direta dos participantes. As respostas são citadas como dados primários do questionário, com indicação do código do participante e do ano da coleta.

3 REFERENCIAL TEÓRICO

A avaliação escolar deve ser mediadora e formativa, considerando tanto o processo quanto o produto da aprendizagem (HOFFMANN, 2011). Com a IA, a distinção entre processo e produto torna-se complexa, amplificando um problema já existente com o plágio tradicional. A avaliação desmistificada exige clareza sobre o objeto avaliado: se o objetivo é a produção textual autônoma, a delegação à máquina configura fraude; se é a curadoria crítica, o uso da IA pode ser legítimo (HADJI, 2001). Bacich e Moran (2018) e Wiggins (1990) defendem que a avaliação autêntica e as metodologias ativas são caminhos para integrar a tecnologia, exigindo que o professor seja designer de experiências de aprendizagem. A resistência docente, conforme apontam Holmes, Bialik e Fadel (2019), está menos na tecnologia em si do que na ausência de formação adequada e de diretrizes institucionais claras.

A IA generativa representa um salto qualitativo em relação às tecnologias educacionais anteriores. Diferentemente de sistemas adaptativos que personalizam o ensino com base em respostas predefinidas, modelos generativos, como ChatGPT,

Gemini e Copilot, produzem textos, códigos e resoluções a partir de comandos dos usuários. Essa capacidade desafia mecanismos tradicionais de verificação de autoria, pois o produto final pode ser tecnicamente impecável sem refletir o processo cognitivo do estudante.

No contexto brasileiro, o Referencial para o Uso e Desenvolvimento Responsáveis de Inteligência Artificial na Educação (BRASIL, 2026) destaca a necessidade de letramento digital crítico, indo além do domínio instrumental. O documento enfatiza que a formação docente deve abranger não apenas o uso técnico das ferramentas, mas também a reflexão sobre ética, autoria, vieses algorítmicos e impactos pedagógicos. Essa perspectiva alinha-se à concepção freireana de que a tecnologia não é neutra e deve ser criticamente apropriada para não reproduzir desigualdades (FREIRE, 2004).

A formação continuada, nesse contexto, deve promover a reciprocidade de saberes entre escola e universidade, superando o modelo de capacitação pontual e tecnicista. Professores precisam de espaços para compartilhar experiências, construir protocolos éticos compartilhados e desenvolver estratégias avaliativas que integrem a IA de forma crítica, preservando a autoria e o pensamento crítico dos estudantes.

4 RESULTADOS E DISCUSSÃO

Dos 40 participantes, 78% (n = 31) afirmaram já ter identificado indícios de uso de IA por estudantes, como vocabulário rebuscado incompatível e estrutura perfeita sem profundidade. Entre esse subgrupo, 65% (n = 20) usaram

detectores de IA, mas apenas 45% (n = 9) desse total os consideraram confiáveis, relatando falsos positivos e falta de transparência. Quanto ao uso pessoal, 72% (n = 29) utilizam IA no planejamento, mas apenas 33% (n = 13) na correção, citando reserva ética e medo de erros como barreiras.

Um exemplo ilustrativo foi narrado pelo Participante 07:

“Um aluno do 9º ano entregou uma redação sobre mudanças climáticas com vocabulário extremamente rebuscado, usando termos como ‘antropoceno’ e ‘feedback loops’ que ele nunca tinha usado antes. O detector de IA indicou 87% de probabilidade de texto gerado por máquina. Chamei o aluno para conversar, mas ele negou veementemente e disse que tinha pesquisado muito. Fiquei numa situação muito desconfortável, porque não tinha como provar nada. Acabei aceitando o trabalho, mas depois descobri que ele tinha usado o ChatGPT para gerar o texto-base e depois fez algumas modificações superficiais. O problema é que eu não tinha um protocolo claro da escola sobre como lidar com isso, e acabei deixando passar. Isso me fez perceber que precisamos de diretrizes institucionais.” (PARTICIPANTE 07, resposta ao questionário, 2026)

Esse relato ilustra as críticas de Holmes, Bialik e Fadel (2019) sobre a opacidade algorítmica dos detectores e a inadequação dessas ferramentas para decisões de alto impacto. A ausência de protocolos institucionais gera insegurança tanto para o professor quanto para o estudante.

A análise cruzada com o tempo de magistério revelou um padrão distintivo (Tabela 1). Docentes com até cinco anos de experiência tenderam ao uso de detectores automatizados (62,5%), enquanto aqueles com mais de 15 anos priorizaram a reconfiguração metodológica via



oralidade (71,4%). Isso sugere, no grupo investigado, uma associação descritiva entre tempo de magistério e estratégias relatadas. A interpretação não constitui efeito causal nem pode ser generalizada para além da amostra. A hipótese de que professores mais experientes, formados em contexto pré-digital (PRENSKY, 2001), podem confiar mais em repertórios de interação direta requer investigação específica.

A oralidade emergiu como estratégia central de enfrentamento (n = 22). O Participante 18 descreveu:

“Depois que percebi que os detectores de IA não eram confiáveis, decidi mudar completamente minha abordagem avaliativa. Em vez de pedir trabalhos escritos para casa, comecei a fazer defesas orais em sala de aula. Os alunos precisam apresentar seus trabalhos em 5 minutos e depois responder perguntas minhas e dos colegas. No começo eles reclamaram muito, dizendo que era ‘volta para a Idade Média’, mas depois perceberam que isso realmente ajuda a aprender. Um aluno que tinha entregue um trabalho excelente sobre a Revolução Francesa, quando foi chamado para defender oralmente, não conseguiu explicar nem o conceito de ‘burguesia’. Até eu soube que ele tinha usado IA. Em vez de punir, usei isso como oportunidade pedagógica: conversamos sobre autoria, sobre o que significa realmente aprender, e ele reescreveu o trabalho com minha orientação. Essa abordagem tem funcionado muito bem, porque desloca o foco do produto para o processo.” (PARTICIPANTE 18, resposta ao questionário, 2026)

Essa narrativa exemplifica o deslocamento do foco do produto para o processo, princípio central da avaliação mediadora (HOFFMANN, 2011). Wiggins (1990) argumenta que a avaliação autêntica exige que o estudante demonstre não apenas o que sabe, mas como e por que sabe. A

oralidade pode contribuir para verificar a mobilização de saberes em tempo real, embora não deva ser tratada como solução única ou universal.

Sobre as demandas, 87,5% (n = 35) pediram suporte estrutural. O Participante 22 sintetizou:

“O que eu mais preciso é de um protocolo claro da escola e da rede de ensino sobre como lidar com suspeitas de uso de IA. Hoje cada professor faz do seu jeito, e isso gera injustiça. Um colega meu deu zero para um aluno baseado num relatório de detector, e o aluno recorreu na direção porque o detector errou. A direção não soube o que fazer, porque não tem nenhuma norma. Eu preciso saber: posso usar detector? Se der positivo, o que eu faço? Chamo o aluno? Chamo os pais? Dou uma segunda chance? Como eu documento isso? Também preciso de materiais para ensinar meus alunos a usarem IA de forma ética, não como substituta do pensamento. A universidade poderia ajudar a desenvolver esses protocolos e materiais, em parceria com a rede, porque sozinha a escola não dá conta.” (PARTICIPANTE 22, resposta ao questionário, 2026)

Esse relato evidencia a necessidade de diretrizes institucionais, em consonância com o Referencial do MEC (BRASIL, 2026). Ao serem questionados se gostariam de proibir a IA, 83% (n = 33) responderam “não”, confirmando a busca por um uso crítico e orientado. Isso exige formação que una reflexão crítica com suporte prático, promovendo reciprocidade de saberes entre escola e universidade.

Tabela 1 – Estratégias adotadas frente à suspeita de uso de IA por tempo de magistério

Tempo de magistério	Uso de detectores (n)	Reconfiguração via oralidade (n)	Total (n)
Até 5 anos	5	3	8
6 a 15 anos	10	12	18
16 a 25 anos	4	7	11
Mais de 25	1	3	4



anos			
Total	20	25	40

Nota: categorias não mutuamente excludentes. Fonte: dados da pesquisa (2026).

5 CONSIDERAÇÕES FINAIS

Este estudo analisou as percepções docentes sobre a IA na avaliação escolar. Os resultados demonstram que os professores estão reconfigurando suas práticas, com destaque para o retorno da oralidade como estratégia de verificação da aprendizagem processual. A experiência docente esteve associada, no grupo investigado, à escolha de estratégias distintas, com professores mais experientes recorrendo com maior frequência a abordagens humanas.

Conclui-se que a formação continuada não pode se restringir ao treinamento técnico. Como implicação deste estudo, universidades e redes de ensino podem co-construir projetos-piloto de Laboratórios de Prática Avaliativa, com carga horária oficial para desenvolver protocolos éticos compartilhados. Futuras pesquisas devem ampliar a amostra e aprofundar a análise cruzada entre tempo de magistério e estratégias adotadas, mitigando os vieses de autorrelato identificados nesta investigação.

REFERÊNCIAS

BACICH, Lilian; MORAN, José (orgs.). *Metodologias ativas para uma educação inovadora: uma abordagem teórico-prática*. Porto Alegre: Penso, 2018. 260 p.

BARDIN, Laurence. *Análise de conteúdo*. Edição revista e ampliada. Tradução de Luís Antero Reto e Augusto Pinheiro. São Paulo: Edições 70, 2016.

BRASIL. Ministério da Educação. Secretaria de Gestão da Informação, Inovação e Avaliação de Políticas Educacionais. *Referencial para o uso e desenvolvimento responsáveis de inteligência artificial na educação*. Brasília: MEC, 2026. Disponível em: [https://www.gov.br/mec/pt-](https://www.gov.br/mec/pt-br/media/segape/referencial-oficial-pt.pdf)

[br/media/segape/referencial-oficial-pt.pdf](https://www.gov.br/mec/pt-br/media/segape/referencial-oficial-pt.pdf). Acesso em: 13 set. 2026.

CRESWELL, John W.; PLANO CLARK, Vicki L. *Designing and conducting mixed methods research*. 3. ed. Thousand Oaks: SAGE Publications, 2017.

FREIRE, Paulo. *Pedagogia da autonomia: saberes necessários à prática educativa*. 30. ed. São Paulo: Paz e Terra, 2004.

HADJI, Charles. *A avaliação desmistificada*. Tradução de Patrícia C. Ramos. Porto Alegre: Artmed, 2001.

HOFFMANN, Jussara. *Avaliação: mito e desafio: uma perspectiva construtivista*. 41. ed. Porto Alegre: Mediação, 2011. 104 p.

HOLMES, Wayne; BIALIK, Maya; FADEL, Charles. *Artificial intelligence in education: promises and implications for teaching and learning*. Boston, MA: Center for Curriculum Redesign, 2019.

LANDIS, J. Richard; KOCH, Gary G. *The measurement of observer agreement for categorical data*. Biometrics, v. 33, n. 1, p. 159-174, mar. 1977. DOI: <https://doi.org/10.2307/2529310>.

PRENSKY, Marc. *Digital Natives, Digital Immigrants Part 1. On the Horizon*, [S. l.], v. 9, n. 5, p. 1-6, 2001. DOI: <https://doi.org/10.1108/10748120110424816>.

WIGGINS, Grant. *The case for authentic assessment*. Practical Assessment, Research & Evaluation, [S. l.], v. 2, n. 2, p. 1-3, 1990. DOI: <https://doi.org/10.7275/ffb1-mm19>.