REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

TUTORIA ADAPTATIVA BASEADA EM LLM: AVALIAÇÃO PERCEPTIVA DA PLATAFORMA QUESTIA NO APOIO AO RACIOCÍNIO LÓGICO

LLM-BASED ADAPTIVE TUTORING: A PERCEPTUAL EVALUATION OF THE QUESTIA PLATFORM IN SUPPORTING LOGICAL REASONING

Ciências Exatas e da Terra • 20/11/2026

REGISTRO DOI: 10.70773/revistatopicos/794930903

Alisson Daniel Costa Nunes[1]

Marcos Vinicius Oliveira Sobrinho[2]

RESUMO

Este artigo analisa a percepção de estudantes do Instituto Federal do Maranhão (IFMA) – Campus Monte Castelo, em São Luís (MA), acerca da contribuição da plataforma educacional adaptativa QuestIA, baseada em *Large Language Models* (LLMs), para o desenvolvimento do raciocínio lógico e a aprendizagem em disciplinas de ciências exatas. A pesquisa é aplicada, de abordagem mista (qualiquantitativa), descritiva e de campo. Após quatro semanas de uso da plataforma, 20 discentes responderam, de forma anônima e voluntária, a um questionário eletrônico. O instrumento contemplou cinco assertivas em escala *Likert* de cinco pontos, um item sobre a frequência de acesso e duas questões abertas. Os dados quantitativos foram submetidos à estatística descritiva e à correlação de Spearman, enquanto o *corpus* qualitativo foi tratado via análise de conteúdo. Os resultados indicam uma percepção majoritariamente favorável: a média geral atingiu 4,30 em uma escala de 5, com concordância entre 85 e 95 por cento por afirmativa, destacando-se a identificação de erros de raciocínio. A frequência de uso correlacionou-se de modo positivo e moderado com a percepção de benefício, sobretudo no aprimoramento da compreensão dos conteúdos. Qualitativamente, os usuários valorizaram o design, a interface e a gamificação do sistema, sugerindo melhorias na velocidade de resposta e a inclusão de tradução instantânea. Considerando o delineamento transversal e de grupo único, os achados expressam associações perceptivas, delegando a investigação de causalidade a estudos futuros. Conclui-se que há indícios favoráveis à adoção da QuestIA como suporte pedagógico no ensino tecnológico, recomendando-se estudos quase-experimentais com medidas objetivas de desempenho.

Palavras-chave: Inteligência artificial; Large Language Model; Aprendizagem adaptativa; Raciocínio lógico; Ensino técnico.

ABSTRACT

This article analyzes the perceptions of students at the Federal Institute of Maranhão (IFMA) – Monte Castelo Campus, in São Luís (MA), regarding the contribution of the adaptive educational platform QuestIA, based on *Large Language Models* (LLMs), to the development of logical reasoning and learning in exact science subjects. The research is applied, descriptive and field-based, with a mixed-methods (qualitative-quantitative) approach. After four weeks of using the platform, 20 students answered an electronic questionnaire anonymously and voluntarily. The instrument included five statements rated on a five-point *Likert* scale, one item on access frequency and two open-ended questions. Quantitative data were analyzed using descriptive statistics and Spearman's correlation, while the qualitative *corpus* was examined through content analysis. Results indicate a predominantly favorable perception: the overall mean reached 4.30 on a 5-point scale, with agreement rates between 85 and 95 percent per statement, especially regarding the identification of reasoning errors. Frequency of use showed a positive, moderate correlation with perceived benefit, especially concerning improved comprehension of course content. Qualitatively, users appreciated the system's design, interface and gamification elements, and suggested faster response times and an instant translation feature. Given the cross-sectional, single-group design, the findings reflect perceptual associations, leaving the investigation of causality to future research. It is concluded that there are favorable indications for adopting QuestIA as a pedagogical aid in technological education, and quasi-experimental studies with objective performance measures are recommended.

Keywords: Artificial intelligence; Large Language Model; Adaptive learning; Logical reasoning; Technical education.

1. INTRODUÇÃO

O baixo desempenho em matemática é um dos problemas mais persistentes da educação brasileira. No Programa Internacional de Avaliação de Estudantes (Pisa) de 2022, o Brasil obteve média de 379 pontos na área, e 73% dos estudantes avaliados ficaram abaixo do nível 2, patamar que a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) considera o mínimo de proficiência. Entre os países-membros da OCDE, esse percentual foi de apenas 31% (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [Inep], 2023). Esse quadro repercute diretamente no ensino técnico, no qual o raciocínio lógico é a base fundamental para as disciplinas de exatas e para a formação profissional.

O raciocínio lógico — entendido como a capacidade de extrair conclusões a partir de informações prévias, reconhecer padrões e organizar etapas na resolução de problemas — perpassa a matemática, a física, a programação e as demais disciplinas tecnológicas. No ensino técnico, em que o estudante precisa articular a base científica à prática profissional, lacunas nessa base podem comprometer o acompanhamento das disciplinas e a autonomia de estudo. O apoio individualizado, apontado como a resposta mais direta a esse tipo de dificuldade, encontra barreiras na elevada proporção de estudantes por docente e no tempo limitado das aulas, o que torna premente investigar recursos tecnológicos capazes de complementar o trabalho do professor.

Nos últimos anos, a inteligência artificial (IA) generativa despontou como uma alternativa viável para ampliar o apoio individualizado ao estudante. Os *Large Language Models* (LLMs), modelos de

linguagem de grande escala capazes de gerar texto em linguagem natural (Brown *et al.*, 2020), permitem construir plataformas capazes de explicar conteúdos, debater respostas e adaptar atividades ao ritmo de cada aluno (Kasneci *et al.*, 2023). Ao oferecer *feedback* imediato e personalizado, essas ferramentas aproximam-se do modelo de tutoria individual, historicamente identificado por Bloom (1984) como um dos mais eficazes para a aprendizagem.

A rápida difusão dessas tecnologias trouxe, contudo, preocupações inerentes à sua adoção. A revisão sistemática de Zawacki-Richter *et al.* (2019) sobre aplicações de IA no ensino superior concluiu que boa parte dos estudos carece de reflexão pedagógica e ética, além de contar com pouca participação de educadores em seu desenho. Esse diagnóstico reforça a necessidade de pesquisas empíricas que ouçam o estudante e avaliem as plataformas em contextos reais de ensino, desconstruindo a premissa determinista de que a tecnologia, por si só, garante a melhoria da aprendizagem.

Nessa mesma direção, Lodge, Thompson e Corrin (2023) delineiam uma agenda de pesquisa para a IA generativa na educação superior, o que reforça que o campo ainda está em construção e que há muito a investigar sobre seu uso pedagógico. É nesse cenário que se insere a QuestIA, plataforma educacional adaptativa baseada em LLM, desenvolvida pelo próprio pesquisador para apoiar o estudo de raciocínio lógico e matemática por meio de exercícios interativos, feedback imediato e elementos de gamificação. Nas fontes consultadas para este trabalho, são ainda escassos os estudos empíricos que avaliem o uso de plataformas desse tipo em instituições públicas de ensino técnico, o que constitui a lacuna que esta pesquisa busca começar a preencher.

A relevância do estudo é de três ordens. Do ponto de vista social, testa uma alternativa de apoio ao estudo acessível a estudantes de uma instituição pública. Do ponto de vista acadêmico, produz dados empíricos originais sobre o uso de uma plataforma baseada em LLM em um contexto real de ensino técnico. Do ponto de vista tecnológico, fornece ao desenvolvedor evidências de uso que orientam o aprimoramento da plataforma.

A pesquisa justifica-se, ainda, por uma razão de método. Antes de investir em estudos de maior custo, como experimentos com medidas objetivas de desempenho, o desenvolvedor de uma plataforma educacional precisa conhecer a percepção de quem a utiliza. Um levantamento de percepção, embora limitado quanto às inferências que permite, é uma etapa inicial viável, que ajuda a refinar o produto e a formular com mais precisão as perguntas das etapas seguintes.

Diante disso, o problema de pesquisa é: de que forma o uso da plataforma QuestIA, baseada em LLM, é percebido pelos estudantes do IFMA Campus Monte Castelo como contribuição para o desenvolvimento do raciocínio lógico e para a aprendizagem em disciplinas de exatas? Partiu-se de duas hipóteses: H1, a QuestIA é percebida como um recurso que apoia a aprendizagem de lógica porque oferece exercícios interativos e feedback imediato; e H2, os estudantes percebem a plataforma como útil e fácil de usar, o que se expressa em avaliações majoritariamente favoráveis de utilidade e de usabilidade.

O objetivo geral é analisar a percepção de estudantes do IFMA Campus Monte Castelo sobre a contribuição da QuestIA para o

desenvolvimento do raciocínio lógico e para a aprendizagem em disciplinas de exatas. Os objetivos específicos são:

a. caracterizar a frequência de uso da plataforma durante quatro semanas;

b. avaliar a percepção dos estudantes quanto à melhora do entendimento dos conteúdos, à identificação de erros e à confiança para resolver exercícios;

c. avaliar a percepção quanto à motivação, à autonomia nos estudos e à usabilidade da plataforma;

d. verificar a associação entre a frequência de uso e a percepção de benefício; e

e. levantar os pontos fortes e as sugestões de melhoria apontados pelos participantes.

Trata-se de pesquisa aplicada, de abordagem mista, descritiva e de campo, com coleta por formulário eletrônico respondido por 20 estudantes ao final de quatro semanas de uso. Convém explicitar desde já que o estudo não teve pré-teste nem grupo de controle e não mediu notas; seus resultados expressam, portanto, a percepção dos participantes e associações entre variáveis, e não efeitos causais da plataforma. O artigo está organizado em seis seções: após esta introdução, a seção 2 apresenta a fundamentação teórica; a seção 3 apresenta a plataforma QuestIA; a seção 4 descreve a metodologia; a seção 5 traz os resultados e a discussão; e a seção 6 apresenta a conclusão.

Por essas características, o trabalho deve ser entendido como um estudo piloto, exploratório e de viabilidade, com n = 20. Seu propósito é obter indícios iniciais sobre a aceitação da QuestIA e orientar estudos futuros, e não generalizar os achados para a população de estudantes do campus ou de outras instituições.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Raciocínio Lógico e Resolução de Problemas

O raciocínio lógico pode ser entendido como o processo pelo qual se extraem conclusões a partir de informações dadas, segundo regras de inferência. Johnson-Laird (2006) sustenta que as pessoas raciocinam construindo e examinando modelos mentais das situações descritas pelas premissas, e que boa parte dos erros decorre de não se considerarem todos os modelos possíveis. Sob essa perspectiva, identificar em que ponto o raciocínio deixou de considerar uma possibilidade é parte central do aprendizado, o que aproxima o estudo da lógica de atividades em que o estudante recebe retorno sobre o passo em que errou.

Polya (1945) propôs um método clássico de resolução de problemas em quatro etapas: compreender o problema, estabelecer um plano, executar o plano e examinar a solução obtida. O autor defende que o ensino de matemática deve desenvolver a capacidade de resolver problemas, e não apenas a memorização de procedimentos. A etapa final, de retrospecto, é aquela em que o estudante confronta o resultado com o enunciado e compreende onde errou, o que dialoga diretamente com a dimensão deste estudo sobre a identificação de erros e a compreensão das questões.

Wing (2006) argumenta que o pensamento computacional, que envolve decompor problemas, reconhecer padrões e abstrair, é uma habilidade fundamental, a ser somada à leitura, à escrita e à aritmética. Embora o conceito seja mais amplo do que o raciocínio lógico tratado neste artigo, ambos compartilham operações mentais que se exercitam por meio de problemas estruturados e de prática repetida. Como o desenvolvimento dessas habilidades depende de prática com retorno, plataformas que oferecem muitos exercícios com feedback imediato são candidatas naturais a apoiar o estudo, desde que o retorno ajude o estudante a raciocinar, e não apenas a conhecer a resposta correta.

2.2. Large Language Models: Fundamentos Técnicos

Os Large Language Models (LLMs) são modelos de linguagem de grande escala, baseados em inteligência artificial, capazes de processar e gerar linguagem natural com elevado grau de coerência contextual (Brown et al., 2020). Sua base é a arquitetura Transformer, proposta por Vaswani et al. (2017), que utiliza um mecanismo de atenção (self-attention) para ponderar a importância de diferentes partes do texto ao processar cada elemento, o que permite capturar dependências de longo alcance na linguagem.

O desenvolvimento de um LLM envolve duas etapas centrais. No pré-treinamento, o modelo é exposto a grandes volumes de texto para aprender a prever o próximo elemento da sequência (Radford et al., 2019). No ajuste fino (fine-tuning), ele é adaptado a domínios ou tarefas específicas; uma técnica relevante nessa etapa é o aprendizado por reforço a partir de feedback humano (RLHF), que alinha as respostas do modelo a preferências humanas, tornando-as mais úteis e seguras (Ouyang et al., 2022). Esses modelos

demonstraram capacidade de generalizar para tarefas diversas com poucos exemplos ou nenhum exemplo específico, fenômeno conhecido como aprendizado few-shot e zero-shot (Brown et al., 2020).

Apesar do desempenho expressivo, os LLMs têm limitações. A principal delas é a alucinação, isto é, a geração de informações incorretas com aparência de veracidade (Zhang et al., 2023). Por isso, o uso educacional exige atenção à confiabilidade do conteúdo gerado e à possível presença de vieses nos dados de treinamento.

2.3. Large Language Models na Educação

No campo educacional, os LLMs despertam interesse crescente por seu potencial de ampliar o acesso a experiências de aprendizagem personalizadas. Kasneci et al. (2023) mapearam oportunidades e desafios do uso desses modelos na educação, citando aplicações como a geração de explicações de código e a formulação de questões avaliativas. Os autores ressaltam que o uso deve ser orientado por princípios pedagógicos claros, pois o acesso fácil a respostas prontas pode prejudicar o desenvolvimento do pensamento crítico e gerar dependência tecnológica, além de levantar questões de honestidade acadêmica.

No plano das políticas educacionais, Miao e Holmes (2023), em documento orientador da UNESCO, afirmam que a adoção responsável da IA generativa exige diretrizes pedagógicas específicas, de modo que essas tecnologias ampliem, e não substituam, o raciocínio autônomo dos estudantes. Nesse contexto, a arquitetura RAG (Retrieval-Augmented Generation), proposta por Lewis et al. (2020), é particularmente relevante: ela permite que o

LLM fundamente suas respostas em materiais específicos, como o conteúdo de um curso, reduzindo o risco de alucinações e aumentando a confiabilidade das informações oferecidas ao estudante.

Revisões recentes ajudam a dimensionar essa produção. Shi et al. (2026) sistematizam aplicações empíricas, benefícios e desafios dos LLMs na educação. Carvalho et al. (2026), em revisão sistemática de 35 estudos sobre LLMs e aprendizagem adaptativa no ensino superior de Computação, identificaram seis modalidades de uso, entre elas a tutoria conversacional e o feedback automatizado, e observaram que 25 estudos (71%) relatam evidências positivas e 8 (23%), resultados mistos, restando 2 com efeitos neutros ou negativos. Os autores notam ainda uma transição do uso genérico de chatbots para aplicações mais estruturadas, integradas a ambientes virtuais de aprendizagem. Esses números se referem ao ensino superior de Computação e não podem ser transpostos diretamente ao ensino técnico.

Na matemática, Pardos e Bhandari (2024) relatam que a ajuda gerada pelo ChatGPT produziu ganhos de aprendizagem equivalentes aos da ajuda elaborada por tutores humanos em habilidades matemáticas. Em revisão integrativa de 29 estudos publicados em 2023 e 2024, Vogel, Ramos e Franzoni (2025) apontam, como benefícios, a personalização da aprendizagem e a tutoria individualizada e, como desafios, as questões éticas, a desigualdade de acesso e as limitações técnicas, além da escassez de estudos de longo prazo.

2.4. Sistemas Tutores Inteligentes e Aprendizagem Adaptativa

Antes dos LLMs, a ideia de apoio individualizado por computador já era explorada pelos sistemas tutores inteligentes (STI), programas que mantêm um modelo do conhecimento do estudante e ajustam problemas, dicas e retornos a esse modelo. Anderson et al. (1995) descrevem lições do desenvolvimento de tutores cognitivos para matemática e programação, entre elas a importância de oferecer retorno no momento do erro e de manter o estudante resolvendo problemas de forma ativa.

VanLehn (2011) comparou a efetividade da tutoria humana, dos sistemas tutores inteligentes e de outras formas de apoio e concluiu que os sistemas que acompanham o estudante passo a passo na resolução alcançaram efeitos próximos aos da tutoria humana, superiores aos de sistemas que apenas avaliam a resposta final. Em meta-análise, Kulik e Fletcher (2016) encontraram efeito positivo da maioria dos STI sobre a aprendizagem, em comparação com a instrução convencional. Esses estudos, porém, mediram a aprendizagem por testes de desempenho, o que a presente pesquisa não fez. Servem, portanto, como horizonte de investigação, e não como termo de comparação direta com os resultados aqui apresentados.

Luckin et al. (2016) defendem que a IA na educação pode oferecer a cada estudante um acompanhamento próximo ao de um tutor, liberando o professor para atividades que exigem julgamento, afeto e criatividade, mas ressaltam que a tecnologia deve ser concebida em parceria com educadores. A diferença que os LLMs introduzem em relação aos tutores tradicionais é a flexibilidade: enquanto os STI dependem de bases de problemas e de regras construídas por especialistas, um LLM é capaz de gerar explicações em linguagem natural para uma ampla variedade de perguntas, inclusive a pedido

do próprio estudante, como no relato de P11, que solicitou explicações mais simples. Em contrapartida, essa flexibilidade traz o risco de respostas imprecisas (Zhang et al., 2023), o que reforça a importância de mecanismos de fundamentação como a RAG (Lewis et al., 2020).

A aprendizagem adaptativa pode ser entendida como o ajuste do conteúdo, do ritmo, da dificuldade ou do suporte instrucional às necessidades do aprendiz (Spriggs; Lau; Passi, 2025). Nos estudos revisados por Carvalho et al. (2026), as estratégias mais frequentes foram o feedback contextual personalizado (9 estudos) e o andaime cognitivo (6 estudos), no qual o sistema conduz o estudante por pistas, decomposição de problemas e perguntas orientadoras. Essa lógica aparece em Ma et al. (2025), que propõem apoiar a aprendizagem de programação por meio da decomposição de problemas feita em conjunto pelo estudante e pelo LLM, e se aproxima das etapas de planejamento descritas por Polya (1945). Bassner, Lottner e Krusche (2025), por sua vez, investigam o impacto de tutores de IA sensíveis ao contexto do curso, em comparação com chatbots de uso geral, sobre a aprendizagem dos estudantes.

2.5. Feedback Formativo e Zona de Desenvolvimento Proximal

Um dos elementos centrais para compreender o potencial de plataformas adaptativas é o papel do feedback no processo de aprendizagem. Hattie e Timperley (2007) definem que o feedback eficaz deve responder a três questões: onde o estudante está em relação ao objetivo de aprendizagem, como está progredindo e quais são os próximos passos. Segundo os autores, o feedback é uma das intervenções pedagógicas de maior impacto sobre o desempenho acadêmico, mas sua efetividade depende da

qualidade, da especificidade e da imediatidade da informação fornecida ao aluno.

Shute (2008) acrescenta que o feedback formativo imediato reduz a carga cognitiva do estudante e evita que erros se consolidem na memória de longo prazo, sendo especialmente útil em atividades de raciocínio lógico e resolução de problemas, porque permite corrigir o raciocínio no momento em que o erro ocorre. Wiggins (2012), por sua vez, defende que o feedback eficaz precisa ser específico, descritivo e orientado à ação, distinguindo-se de simples confirmações de acerto ou erro.

A teoria da carga cognitiva ajuda a explicar por que esse apoio é importante. Sweller (1988) argumenta que a resolução de problemas por tentativa e erro impõe elevada carga à memória de trabalho de iniciantes e pode competir com a construção de esquemas de conhecimento. Isso justifica oferecer explicações, exemplos e dicas estruturadas enquanto o estudante ainda não domina o assunto, retirando o apoio de forma gradual.

A fundamentação da aprendizagem adaptativa remonta a Vygotsky (1978), que propôs o conceito de Zona de Desenvolvimento Proximal (ZDP): a distância entre o que o aprendiz realiza de forma independente, seu nível de desenvolvimento real, e o que consegue realizar com a orientação de alguém mais experiente, seu nível de desenvolvimento potencial. Plataformas adaptativas podem ser vistas como tentativas de oferecer esse apoio calibrado em larga escala.

O estudo de Bloom (1984), conhecido como o problema dos dois sigma, mostrou que estudantes submetidos à tutoria individual com

aprendizagem por domínio obtiveram desempenho cerca de dois desvios padrão acima do de estudantes em sala de aula convencional. O desafio colocado por Bloom, encontrar métodos de instrução coletiva tão eficazes quanto a tutoria individual, é o pano de fundo das plataformas adaptativas baseadas em IA.

2.6. Gamificação e Tecnologia Educacional

Deterding et al. (2011) definem gamificação como o uso de elementos de design de jogos em contextos que não são jogos, o que a distingue do simples uso de jogos educativos. Kapp (2012) argumenta que a abordagem é especialmente adequada a conteúdos que exigem raciocínio lógico e resolução de problemas, pois os ciclos de desafio, tentativa, erro e feedback imediato, típicos dos jogos, favorecem a consolidação do pensamento lógico-matemático. No Brasil, Fardo (2013) sustenta que ambientes gamificados podem promover competências cognitivas complexas, desde que os elementos de jogo estejam alinhados a objetivos pedagógicos claros.

As evidências sobre a gamificação, contudo, pedem cautela. Hamari, Koivisto e Sarsa (2014), ao revisarem estudos empíricos, constataram que os efeitos costumam ser positivos, mas dependem do contexto em que a gamificação é aplicada e das características dos usuários. Dichev e Dicheva (2017), em revisão crítica, apontam que ainda há incerteza sobre seus efeitos na educação e recomendam estudos mais rigorosos, com delineamentos capazes de separar o efeito dos elementos de jogo de outros fatores presentes na intervenção.

A personalização do ensino mediada por tecnologia é defendida também por Bacich, Tanzi Neto e Trevisani (2015), para quem ela

permite que cada estudante percorra rotas de aprendizagem distintas no mesmo ambiente escolar, respeitando ritmos individuais e promovendo maior equidade. Combinadas, essas perspectivas fundamentam a proposta da QuestIA: gamificação para estruturar o percurso em desafios progressivos, feedback imediato para orientar a correção de erros e um LLM para adaptar as explicações ao nível de cada estudante.

2.7. Motivação, Autoeficácia e Aceitação de Tecnologia

Três construtos ajudam a compreender o uso de uma plataforma como a QuestIA: a autoeficácia, a motivação e a aceitação de tecnologia. Bandura (1997) define a autoeficácia como a crença do indivíduo em sua capacidade de organizar e executar as ações necessárias para atingir determinados resultados, e identifica as experiências de êxito como a principal fonte dessa crença. Nesse sentido, atividades em que o estudante resolve problemas com apoio e vivencia acertos progressivos tendem a fortalecer a confiança para resolver exercícios.

Ryan e Deci (2000), na teoria da autodeterminação, afirmam que a motivação intrínseca é favorecida quando o ambiente apoia as necessidades psicológicas de autonomia, competência e pertencimento. Ambientes que permitem ao estudante escolher o ritmo, receber retorno sobre sua competência e controlar o próprio percurso são, portanto, candidatos a sustentar a motivação. Zimmerman (2002) acrescenta que a aprendizagem autorregulada envolve planejar, monitorar e avaliar o próprio desempenho, e que o retorno sobre o desempenho alimenta esse monitoramento. Esses referenciais ajudam a compreender a motivação e a autonomia nos estudos.

Por fim, Davis (1989) propôs que a aceitação de uma tecnologia depende sobretudo de duas percepções do usuário: a utilidade percebida e a facilidade de uso percebida. O modelo é útil como lembrete de que percepções importam para a adoção de uma ferramenta, mas elas não substituem a medida direta da aprendizagem, distinção que orienta a leitura dos resultados deste artigo.

2.8. A QuestIA à Luz do Referencial Teórico

Conforme descrito na introdução e detalhado na seção 3, a QuestIA reúne exercícios interativos de raciocínio lógico, ajuste da dificuldade conforme o desempenho, feedback imediato, elementos de gamificação e um tutor de apoio que pode recorrer a um LLM para gerar dicas e explicações. À luz do referencial apresentado, cada componente pode ser associado a um fundamento teórico. Os exercícios interativos relacionam-se à ideia de resolução ativa de problemas (Polya, 1945) e aos ciclos de desafio e tentativa da gamificação (Kapp, 2012). O feedback imediato e as explicações do tutor apoiam-se em Hattie e Timperley (2007) e Shute (2008). O ajuste da dificuldade ao desempenho do estudante aproxima-se do apoio calibrado da Zona de Desenvolvimento Proximal (Vygotsky, 1978) e da busca por métodos de instrução tão eficazes quanto a tutoria individual (Bloom, 1984). Os elementos de jogo e a interface relacionam-se à motivação e à experiência de uso (Deterding et al., 2011; Davis, 1989).

Essa associação é uma proposta de leitura, e não uma demonstração de que cada componente produza o efeito correspondente. O estudo não isolou o efeito de nenhum deles, e a plataforma foi avaliada em conjunto. A associação serve, antes, para organizar a

leitura dos componentes da plataforma e para orientar a discussão dos resultados na seção 5.

3. A PLATAFORMA QUESTIA

3.1. Visão Geral e Acesso

A QuestIA é uma plataforma web de treino de raciocínio lógico, desenvolvida pelo próprio pesquisador. Na versão utilizada neste estudo (v2.5), ela é distribuída como um único arquivo HTML, executado diretamente no navegador, sem necessidade de cadastro, login ou servidor próprio. A tela inicial exibe os pontos de experiência (XP) do estudante, o número de sessões realizadas e seu recorde de acertos, e dela o estudante inicia uma sessão adaptativa (Figura 1). A plataforma permite ainda alternar entre tema claro e escuro e, opcionalmente, conectar o tutor a um modelo de linguagem externo, como descrito na seção 3.4.

Figura 1. *Tela inicial da plataforma QuestIA*



Fonte: O autor (2026).

3.2. Banco de Questões

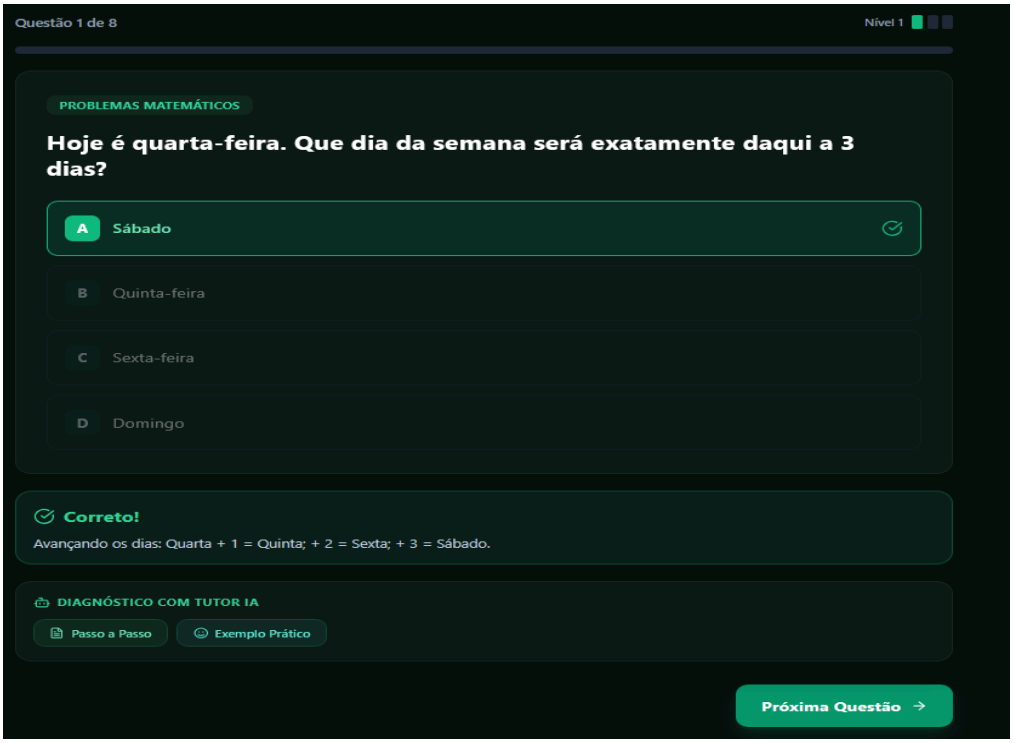
O conteúdo da plataforma é exclusivamente o raciocínio lógico. O banco reúne 20 questões de múltipla escolha, com quatro alternativas cada, distribuídas em três níveis de dificuldade (iniciante, intermediário e avançado) e em cinco categorias: sequências lógicas (5 questões), problemas matemáticos (5), proposições (4), dedução lógica (3) e lógica espacial (3). A categoria de problemas matemáticos reúne questões de aritmética e de raciocínio quantitativo apresentadas em forma de problema, como o cálculo de preços a partir de relações entre valores ou do número de apertos de mão entre um grupo de pessoas. As questões de proposições tratam da negação, da condicional e de inferências como o modus ponens e o modus tollens, além da falácia da afirmação do consequente. Cada questão acompanha uma explicação elaborada previamente, exibida ao estudante depois que ele responde, e a ordem das alternativas é embaralhada a cada apresentação.

3.3. Sessões, Adaptação de Nível E Feedback

Cada sessão é composta por oito questões, sorteadas do banco dentro do nível de dificuldade corrente, e começa sempre no nível 1. A adaptação segue uma regra simples e determinística: dois acertos consecutivos elevam o nível em uma unidade, até o máximo de 3, e dois erros consecutivos o reduzem, até o mínimo de 1. Quando não restam questões inéditas no nível, a plataforma sorteia entre as demais ainda não apresentadas na sessão. A adaptação, portanto, depende apenas do desempenho do estudante na sessão e não é realizada por um modelo de linguagem. Ao responder, o estudante recebe feedback imediato: a plataforma indica se a resposta está

correta e apresenta a explicação da questão (Figura 2). Ao final da sessão, é exibido um resumo com a pontuação, o XP ganho, a maior sequência de acertos, o desempenho por categoria e o histórico de níveis, com a possibilidade de revisar as questões respondidas.

Figura 2. *Questão com feedback imediato e explicação*



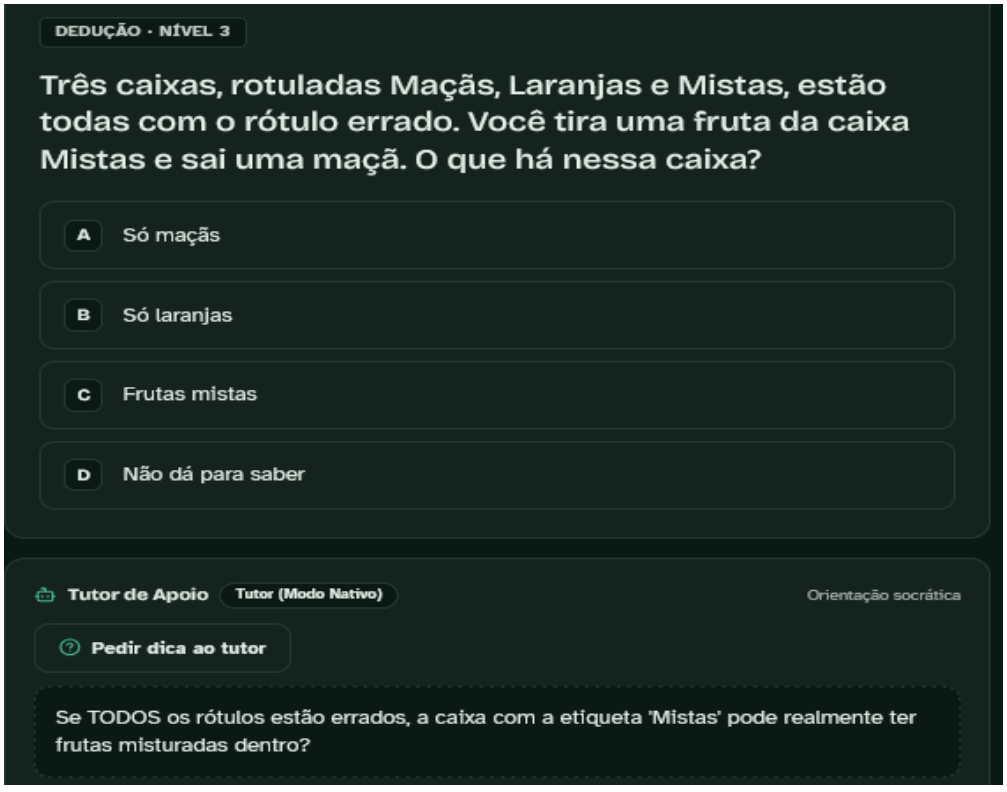
Fonte: O autor (2026).

3.4. Tutor de Apoio e Uso de Modelo de Linguagem

O tutor da plataforma oferece três modos de apoio: a dica socrática, solicitada antes da resposta e que formula uma pergunta orientadora sem revelar a solução; o passo a passo, que apresenta o diagnóstico lógico da questão; e o exemplo prático, que a relaciona a uma situação do cotidiano. Os dois últimos ficam disponíveis após a resposta (Figura 3). O tutor pode operar de duas formas. No modo nativo, que não depende de serviços externos de inteligência artificial, as mensagens são compostas a partir de textos pré-definidos e da explicação elaborada para cada questão. No modo com LLM, opcional, o estudante informa nas configurações uma

chave de API e a plataforma passa a enviar ao modelo Gemini 1.5 Flash, do Google, um prompt com o enunciado, o gabarito, a explicação oficial, a resposta do estudante e o modo solicitado, com a instrução de responder em português do Brasil e em até 75 palavras. Se não há chave configurada ou se a chamada ao modelo falha, a plataforma recorre ao modo nativo. O modelo é utilizado como disponibilizado pelo provedor, sem treinamento ou ajuste fino realizados pelo pesquisador e sem recuperação de documentos (RAG). Neste estudo, o tutor foi utilizado no modo com LLM, com a chave de API configurada pelo próprio pesquisador, e atuou como recurso auxiliar, sem participar da adaptação de nível.

Figura 3. *Tutor de apoio: dica socrática solicitada antes da resposta*



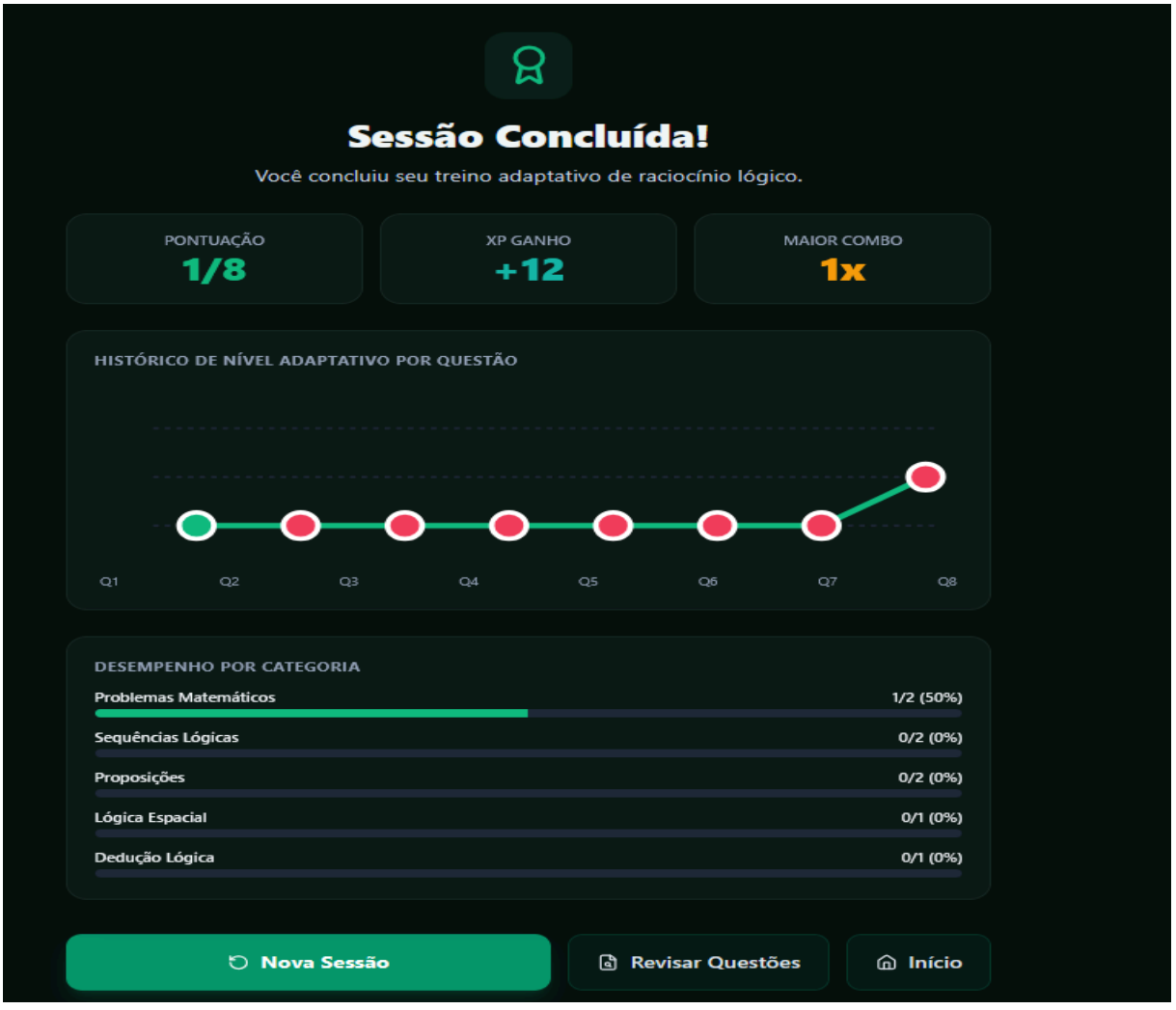
Fonte: O autor (2026).

3.5. Gamificação

A plataforma incorpora elementos de jogo ao percurso do estudante. Cada acerto rende pontos de experiência proporcionais

ao nível da questão (10, 20 ou 30 pontos), acrescidos de um bônus por sequência de acertos (combo), limitado a 10 pontos. Há seis conquistas, a saber: Primeira Vitória; Em Chamas, para três acertos seguidos; Imparável, para cinco acertos seguidos; Mestre da Lógica, ao alcançar o nível 3; Perfeição Cognitiva, para oito acertos em oito questões; e Mente Curiosa, para três consultas ao tutor. A plataforma também notifica as mudanças de nível e exibe um efeito de confete ao final de sessões com cinco ou mais acertos (Figura 4). Esses elementos buscam estruturar o percurso em desafios progressivos, conforme discutido na seção 2.6.

Figura 4. *Galeria de conquistas e resultado da sessão*



Fonte: O autor (2026).

3.6. Construção e Tecnologias

A plataforma foi desenvolvida em HTML, CSS e JavaScript, com o framework de estilos Tailwind CSS e os ícones da biblioteca Lucide, carregados por redes de distribuição de conteúdo, o que exige conexão com a internet. Não há back-end: o banco de questões e a regra de adaptação estão no próprio arquivo, e o progresso do estudante (XP, sessões, recorde, conquistas e preferências), bem como a chave de API, quando informada, são gravados apenas no armazenamento local do navegador. Por essa razão, a plataforma não registra o uso de cada estudante de forma centralizada, e a frequência de uso analisada neste estudo foi obtida por autorrelato no formulário (seção 4.3).

3.7. Disponibilização aos Estudantes

A plataforma foi disponibilizada aos estudantes do curso de Eletroeletrônica do IFMA Campus Monte Castelo como ferramenta complementar de estudo durante quatro semanas. O acesso foi restrito a esses estudantes e enviado a eles pelo aplicativo WhatsApp, sem necessidade de cadastro. A apresentação da atividade ficou a cargo do próprio pesquisador, que também desenvolveu a plataforma. O uso foi livre: os estudantes utilizaram a QuestIA por conta própria, sem vínculo com uma disciplina específica, e, ao final do período, responderam ao questionário eletrônico.

4. METODOLOGIA

4.1. Tipo de Pesquisa

Quanto à natureza, esta pesquisa é aplicada, pois busca gerar conhecimento prático sobre o uso da QuestIA em um contexto real de ensino. Quanto à abordagem, é mista, com predominância

quantitativa: as respostas em escala Likert foram tratadas por estatística descritiva e correlação, e as respostas abertas foram analisadas qualitativamente. Quanto aos objetivos, é descritiva, com uma análise exploratória da associação entre frequência de uso e percepção. Quanto aos procedimentos, é uma pesquisa de campo do tipo levantamento (survey), de corte transversal, realizada diretamente com estudantes do IFMA Campus Monte Castelo (Gil, 2002; Marconi; Lakatos, 2003).

A opção pela abordagem mista segue a lógica descrita por Creswell (2010), segundo a qual a combinação de dados quantitativos e qualitativos permite uma compreensão mais ampla do problema do que cada tipo isoladamente. Neste estudo, os dados quantitativos, obtidos nas afirmativas fechadas, mostram a direção e a intensidade das percepções, e os dados qualitativos, obtidos nas questões abertas, revelam aspectos que as afirmativas não contemplam, como a opinião sobre o design e as sugestões de melhoria. Em termos de delineamento, trata-se de um estudo exploratório de corte transversal, com um único grupo e uma única medição, sem manipulação de variáveis, o que determina o alcance das inferências permitidas.

Em razão do tamanho da amostra e da ausência de grupo de controle, a pesquisa é caracterizada como estudo piloto exploratório de viabilidade. Esse enquadramento orienta a análise: os resultados são apresentados como indicativos e voltados à formulação de hipóteses, e os testes estatísticos têm caráter exploratório.

4.2. População e Amostra

A pesquisa foi realizada no IFMA Campus Monte Castelo, em São Luís (MA). A população do estudo é composta pelos estudantes do campus que utilizaram a QuestIA durante quatro semanas, como ferramenta complementar de estudo no contexto de disciplinas de exatas, com foco em raciocínio lógico e matemática. A plataforma é de acesso web, sem cadastro ou login, e foi disponibilizada aos estudantes nas condições descritas na seção 3.7. Seu conteúdo reúne 20 questões de múltipla escolha de raciocínio lógico e de problemas matemáticos, distribuídas em sequências lógicas, problemas matemáticos, proposições, dedução lógica e lógica espacial, com sessões de oito questões e ajuste de dificuldade conforme o desempenho (seções 3.2 e 3.3). O uso, portanto, teve caráter de prática e revisão complementar às aulas. A amostra é não probabilística, por conveniência, e reúne os 20 estudantes que aceitaram responder ao formulário.

Como a amostra reúne apenas os estudantes que aceitaram responder, é possível que tenham participado sobretudo aqueles mais envolvidos com a plataforma, o que configura um possível viés de autosseleção. Esse aspecto deve ser considerado na leitura dos resultados, pois impede generalizar as percepções aqui descritas para todos os estudantes do campus.

4.3. Instrumento de Coleta de Dados

O instrumento foi um formulário eletrônico, elaborado no Google Forms, com oito questões, precedidas pelo Termo de Consentimento Livre e Esclarecido (TCLE), conforme o Quadro 1. A Questão 1 mede a frequência de uso da plataforma em quatro categorias. As Questões 2 a 6 são afirmativas avaliadas em escala Likert de cinco pontos, de 1 (discordo totalmente) a 5 (concordo totalmente). As Questões 7 e 8

são abertas. A escolha do questionário apoia-se em Gil (2002) e Marconi e Lakatos (2003), que o consideram um instrumento eficaz para obter dados padronizados e comparáveis em pesquisas de campo.

Quadro 1. *Questões do formulário aplicado aos estudantes*

Nº	Questão ou afirmativa	Tipo e opções de resposta
TCLE	Você leu e compreendeu o TCLE e aceita participar voluntariamente desta pesquisa?	Fechada: sim
1	Com que frequência você utilizou a plataforma QuestIA durante as quatro semanas?	Fechada: 1 vez; 2 a 3 vezes; 4 a 5 vezes; todos os dias
2	Após utilizar a QuestIA, melhorei meu entendimento dos conteúdos de raciocínio lógico e matemática.	Likert de 1 a 5
3	A QuestIA me ajudou a identificar meus erros e compreender melhor as questões.	Likert de 1 a 5
4	Após utilizar a QuestIA, sinto-me mais confiante para resolver exercícios de raciocínio lógico e matemática.	Likert de 1 a 5
5	O uso da QuestIA aumentou minha motivação e autonomia nos estudos.	Likert de 1 a 5
6	A plataforma QuestIA foi fácil de usar, e suas explicações foram claras e úteis.	Likert de 1 a 5
7	Como você avalia sua experiência de aprendizagem utilizando a	Aberta

	plataforma QuestIA durante as quatro semanas?	
8	O que você mais gostou na QuestIA e o que gostaria que fosse melhorado?	Aberta

Fonte: O autor (2026).

As afirmativas foram associadas às dimensões discutidas no referencial teórico: Q2 e Q3 relacionam-se ao feedback formativo e à identificação do erro (Polya, 1945; Hattie; Timperley, 2007; Shute, 2008); Q4, à autoeficácia (Bandura, 1997); Q5, à motivação e à aprendizagem autorregulada (Ryan; Deci, 2000; Zimmerman, 2002); e Q6, à facilidade de uso percebida (Davis, 1989). As afirmativas Q2 e Q4 mencionam raciocínio lógico e matemática porque o banco de questões da plataforma inclui a categoria de problemas matemáticos (seção 3.2).

A escala de cinco pontos segue a técnica proposta por Likert (1932) para a mensuração de atitudes, na qual o respondente indica seu grau de concordância com afirmativas sobre o objeto avaliado. As afirmativas foram redigidas em primeira pessoa, em sentido positivo e com linguagem acessível, de modo a manter o formulário curto e de resposta rápida. A redação em sentido positivo facilita a compreensão, mas pode favorecer a concordância dos respondentes, aspecto retomado entre as limitações do estudo. As questões abertas foram incluídas para captar percepções que não se limitam às categorias previstas nas afirmativas.

4.4. Procedimentos de Coleta

Após as quatro semanas de uso da plataforma, o link do formulário foi disponibilizado aos estudantes, que o responderam uma única vez, de forma anônima e voluntária. As respostas foram registradas entre 2 e 5 de novembro de 2026. Para a apresentação dos resultados, os participantes foram numerados de P1 a P20 pela ordem de registro das respostas.

4.5. Tratamento e Análise dos Dados

Os dados foram exportados do Google Forms em planilha e processados em linguagem Python 3.12, com as bibliotecas pandas 3.0 e SciPy 1.18. As respostas em escala Likert foram convertidas em valores de 1 a 5 e descritas por frequências absolutas e relativas, média, desvio padrão amostral e percentual de concordância (soma das respostas 4 e 5). A consistência interna das cinco afirmativas foi estimada pelo alfa de Cronbach, que resultou em 0,85. A associação entre a frequência de uso (codificada de 1 a 4) e cada afirmativa foi examinada pelo coeficiente de correlação de postos de Spearman, adequado a variáveis ordinais e a amostras pequenas, com nível de significância de 5%. Por se tratar de análise exploratória com 20 participantes, não se aplicou correção para comparações múltiplas, e os valores de p devem ser lidos como indicativos.

As respostas abertas (40 respostas, sendo 20 para cada questão) foram analisadas pela técnica de análise de conteúdo de Bardin (2011), em três fases: pré-análise, com leitura flutuante do material; exploração do material, com codificação, tendo cada resposta como unidade de registro, e agrupamento em categorias definidas a posteriori; e tratamento dos resultados, com contagem das categorias e interpretação à luz do referencial teórico. A codificação foi realizada por um único pesquisador.

A consistência interna, estimada pelo alfa de Cronbach (1951), indica o grau em que os itens de uma escala variam em conjunto. O valor obtido, 0,85, situa-se na faixa que Tavakol e Dennick (2011) consideram adequada, mas deve ser lido com cautela: Taber (2018) lembra que o alfa depende do número de itens e da homogeneidade do que se mede, e que não constitui, isoladamente, evidência de validade do instrumento. O coeficiente de correlação de postos de Spearman (1904), por sua vez, mede a associação monotônica entre duas variáveis a partir de seus postos, razão pela qual não pressupõe distribuição normal e é apropriado a variáveis ordinais, como as respostas em escala Likert e a frequência de uso. O Quadro 2 resume a correspondência entre os objetivos específicos, as questões do formulário e as técnicas de análise empregadas.

Quadro 2. *Correspondência entre objetivos específicos, questões do formulário e técnicas de análise*

Objetivo específico	**Questões**	**Técnica de análise**
a) Caracterizar a frequência de uso da plataforma	Q1	Frequências absolutas e relativas
b) Avaliar a percepção sobre entendimento, identificação de erros e confiança	Q2, Q3 e Q4	Distribuição das respostas, média, desvio padrão e percentual de concordância
c) Avaliar a percepção sobre motivação, autonomia e usabilidade	Q5 e Q6	Distribuição das respostas, média, desvio padrão e percentual de concordância
d) Verificar a associação entre frequência de uso e percepção de benefício	Q1 com Q2 a Q6	Correlação de postos de Spearman
e) Levantar pontos fortes e sugestões de melhoria	Q7 e Q8	Análise de conteúdo (Bardin, 2011)

Fonte: O autor (2026).

4.6. Aspectos Éticos e Uso de Inteligência Artificial

A participação foi voluntária e anônima. O formulário iniciava com o TCLE, e todos os 20 participantes registraram a concordância em participar. Ferramentas de inteligência artificial generativa (Claude, da Anthropic) foram utilizadas como apoio à estruturação e à redação do texto, à revisão linguística e ao cálculo das estatísticas; a definição do problema, a coleta dos dados, a conferência dos resultados e as interpretações são de responsabilidade do autor.

5. RESULTADOS E DISCUSSÕES

Esta seção apresenta os resultados da coleta e, em seguida, discute-os à luz da fundamentação teórica. Como o estudo não teve pré-teste nem grupo de controle, os achados referem-se à percepção dos 20 participantes e não permitem inferir efeito da plataforma sobre o desempenho.

5.1. Frequência de Uso da Plataforma

A Tabela 1 mostra que 12 dos 20 participantes (60%) usaram a QuestIA com frequência alta, de 4 a 5 vezes ou todos os dias; 6 (30%) usaram de 2 a 3 vezes; e 2 (10%) usaram uma única vez. A categoria mais frequente foi a de 4 a 5 vezes (40%). O instrumento não especificou a unidade de tempo dessas categorias (por semana, por exemplo), o que limita a interpretação da variável.

Tabela 1. *Frequência de uso da plataforma QuestIA durante as quatro semanas*

Frequência de uso	n	%
1 vez	2	10
2 a 3 vezes	6	30
4 a 5 vezes	8	40
Todos os dias	4	20
Total	**20**	**100**

Fonte: O autor (2026).

A distribuição indica que a maior parte da amostra teve contato repetido com a plataforma, condição necessária para que as percepções relatadas nas seções seguintes se baseiem em experiência de uso, e não apenas em uma impressão inicial. Os dois participantes que a utilizaram uma única vez, por outro lado, têm base de experiência limitada, o que justifica examinar se a percepção varia com a frequência, objetivo tratado na seção 5.3.

5.2. Percepção Sobre Aprendizagem, Confiança, Autonomia e Usabilidade

A Tabela 2 e o Gráfico 1 apresentam as respostas às cinco afirmativas. As médias variaram de 4,15 a 4,40, e a média geral por participante foi 4,30 (desvio padrão de 0,60). Nenhum participante assinalou “discordo totalmente”. O alfa de Cronbach de 0,85 indica boa consistência interna entre as afirmativas, embora a amostra pequena torne essa estimativa pouco precisa.

Tabela 2. *Distribuição das respostas (n) e medidas descritivas das afirmativas Q2 a Q6 (escala de 1 a 5; n = 20)*

Afirmativa	1	2	3	4	5
Q2 Melhora do entendimento dos conteúdos	0	0	3	9	8
Q3 Identificaç	0	0	1	11	8

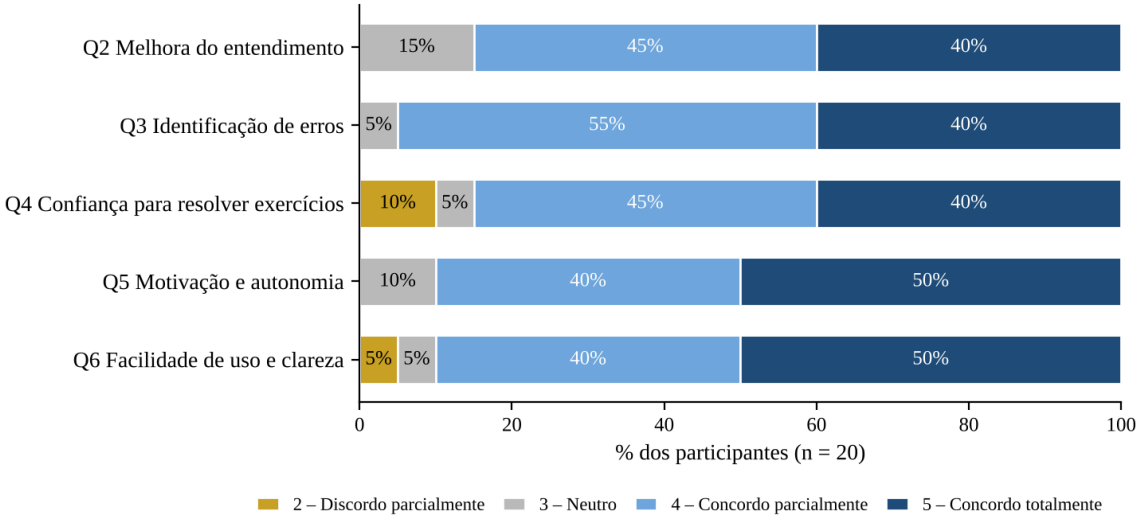
⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: https://revistatopicos.com.br/artigos/tutoria-adaptativa-baseada-em-llm-avaliacao-perceptiva-da-plataforma-questia-no-apoio-ao-raciocinio-logico?noblockage

Fonte: O autor (2026).

Nota: DP = desvio padrão amostral; Concord. = percentual de respostas 4 (concordo parcialmente) e 5 (concordo totalmente).

A afirmativa com maior concordância foi a de identificação de erros e compreensão das questões (Q3), com 95% de respostas 4 ou 5 e média de 4,35. A de motivação e autonomia (Q5) teve a maior média (4,40), com metade dos participantes concordando totalmente, percentual igual ao de facilidade de uso e clareza das explicações (Q6), cuja média foi 4,35. A melhora do entendimento (Q2) alcançou 85% de concordância (média de 4,25), com três respostas neutras. A confiança para resolver exercícios (Q4) teve a menor média (4,15) e a maior dispersão (desvio padrão de 0,93). Os três únicos registros de discordância parcial do conjunto, dois em Q4 e um em Q6, vieram de apenas dois participantes.

Gráfico 1. *Distribuição percentual das respostas às afirmativas Q2 a Q6*



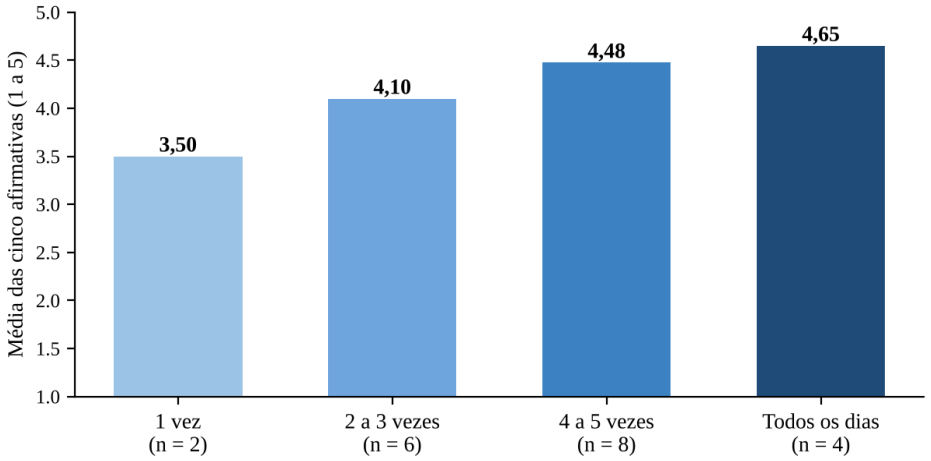
Fonte: O autor (2026).

Nota: nenhuma afirmativa recebeu resposta 1 (discordo totalmente).

5.3. Frequência de Uso e Percepção de Benefício

Participantes que usaram a plataforma quatro vezes ou mais (n = 12) tiveram média geral de 4,53, contra 3,95 dos que a usaram até três vezes (n = 8). O Gráfico 2 detalha esse padrão: a média cresce a cada categoria de frequência, de 3,50 (1 vez) a 4,65 (todos os dias).

Gráfico 2. *Média geral das afirmativas Q2 a Q6, por frequência de uso*



Fonte: O autor (2026).

Nota: as categorias "1 vez" (n = 2) e "todos os dias" (n = 4) têm poucos casos, e suas médias são instáveis.

A Tabela 3 confirma a associação: a correlação de Spearman entre a frequência de uso e a média geral foi positiva e moderada (ρ = 0,51; p = 0,021). Entre as afirmativas, a associação foi mais forte para a melhora do entendimento (Q2; ρ = 0,68; p < 0,001) e para a identificação de erros (Q3; ρ = 0,56; p = 0,010), foi limítrofe para a confiança (Q4; ρ = 0,43; p = 0,059) e foi fraca e não significativa para motivação e autonomia (Q5) e para usabilidade (Q6). Essa associação não indica causalidade: é possível que estudantes mais satisfeitos tenham usado mais a plataforma, e não o contrário.

Tabela 3. *Correlação de Spearman entre a frequência de uso e as afirmativas (n = 20)*

Afirmativa	**ρ de Spearman**	**p**
Q2 Melhora do entendimento dos conteúdos	0,68	< 0,001
Q3 Identificação de erros e compreensão das questões	0,56	0,010
Q4 Confiança para resolver exercícios	0,43	0,059

Q5 Motivação e autonomia nos estudos	0,20	0,397
Q6 Facilidade de uso e clareza das explicações	0,24	0,314
Média geral das cinco afirmativas	**0,51**	**0,021**

Fonte: O autor (2026).

Nota: análise exploratória, sem correção para comparações múltiplas.

5.4. Respostas Abertas

A Questão 7 pedia a avaliação da experiência de aprendizagem com a plataforma. A Tabela 4 mostra que 9 respostas (45%) foram avaliações globais positivas em uma palavra ou expressão, como “Excelente”, “Muito boa” ou “Ótima”, e que 7 (35%) descreveram ganho percebido de aprendizagem ou de desempenho. Entre estas, P4 afirmou que a plataforma “me ajudou a compreender melhor os assuntos mais complexos”, e P11 relatou estar “aprendendo bastante, peço pra simplificar e fica bom”, o que indica que o estudante pediu ao sistema explicações mais simples. Uma resposta (P16) associou a plataforma à busca de conhecimento e aprimoramento, e três (15%) não permitiram classificação (“Sim”, “Todas as etapas”, “No raciocínio lógico”).

Tabela 4. *Categorias das respostas à Questão 7 (avaliação da experiência de aprendizagem)*

Categoria	**n**	**%**	**Exemplo**
Avaliação global positiva	9	45	“Excelente”

Ganho percebido de aprendizagem ou desempenho	7	35	“Muito aprendizado”
Motivação para buscar aprimoramento	1	5	“buscar conhecimento e aprimoramento”
Resposta não classificável	3	15	“Todas as etapas”
Total	**20**	**100**	

Fonte: O autor (2026).

A Questão 8 perguntava o que o participante mais gostou e o que gostaria que fosse melhorado. A Tabela 5 mostra que o aspecto mais citado foi o visual, isto é, design, interface e tema (7 respostas; 35%). Seis participantes (30%) expressaram satisfação geral ou não apontaram melhorias (“Está ótimo”, “Nada”). Duas respostas trataram do desempenho e da velocidade de resposta da plataforma, e duas, da qualidade das respostas geradas (“Respostas mais humanizadas” e a organização das ideias). Houve ainda uma menção ao aprendizado gamificado (P2) e uma menção à gratuidade da plataforma, acompanhada da sugestão de incluir tradução instantânea (P3). Como a questão reúne dois pedidos, a maioria das respostas curtas, como “Interface” ou “Design”, não permite distinguir elogio de sugestão.

Tabela 5. *Categorias das respostas à Questão 8 (o que gostou e o que melhorar)*

Categoria	**n**	**%**	**Exemplo**

Design, interface e tema	7	35	“A interface”
Satisfação geral ou sem sugestão de melhoria	6	30	“Está ótimo”
Desempenho e velocidade de resposta	2	10	“Velocidade de resposta”
Qualidade e humanização das respostas	2	10	“Respostas mais humanizadas”
Gamificação	1	5	“Gostei do aprendizado gamificado”
Gratuidade e sugestão de tradução instantânea	1	5	(resposta parafraseada)
Sem resposta válida	1	5	“.”
Total	**20**	**100**	

Fonte: O autor (2026).

5.5. Síntese dos Resultados Segundo os Objetivos Específicos

Antes da discussão, o Quadro 3 reúne os principais resultados organizados pelos objetivos específicos definidos na introdução, de modo a mostrar em que medida cada um foi contemplado pelos dados.

Quadro 3. *Síntese dos resultados por objetivo específico*

Objetivo específico	**Principais resultados**	**Seção**

a) Frequência de uso	60% usaram a plataforma 4 vezes ou mais (40% de 4 a 5 vezes e 20% todos os dias); 30% usaram de 2 a 3 vezes e 10% uma única vez.	5.1
b) Entendimento, erros e confiança	Concordância de 95% na identificação de erros (média 4,35), 85% na melhora do entendimento (4,25) e 85% na confiança para resolver exercícios (4,15).	5.2
c) Motivação, autonomia e usabilidade	Concordância de 90% em motivação e autonomia (média 4,40) e de 90% em facilidade de uso e clareza das explicações (4,35).	5.2
d) Frequência e percepção	Correlação positiva e moderada com a média geral ($\rho = 0,51$; $p = 0,021$), mais forte em entendimento ($\rho = 0,68$) e identificação de erros ($\rho = 0,56$); confiança limítrofe ($p = 0,059$); motivação e usabilidade sem associação significativa.	5.3
e) Pontos fortes e sugestões	Design, interface e tema foram os aspectos mais citados (35%); houve menção à gamificação; as sugestões foram maior velocidade de resposta, respostas mais humanizadas e tradução instantânea.	5.4

Fonte: O autor (2026).

5.6. Discussão

No conjunto, os dados indicam uma percepção amplamente favorável à QuestIA. A maior concordância recaiu sobre a identificação de erros e a compreensão das questões (95%), o que é coerente com o papel atribuído ao feedback por Hattie e Timperley (2007), para quem o retorno ao estudante deve indicar onde ele está e como seguir, e por Shute (2008), que destaca o valor do feedback imediato em atividades de raciocínio lógico. É preciso, contudo,

separar percepção de efeito: os estudantes relatam que a plataforma ajuda a entender os erros, mas o estudo não verificou se isso se traduziu em melhor desempenho.

Estudos de revisão sobre sistemas tutores inteligentes (VanLehn, 2011; Kulik; Fletcher, 2016) relatam efeitos positivos sobre a aprendizagem, medidos por testes de desempenho. Os resultados desta pesquisa são compatíveis com eles quanto à direção, já que a percepção é favorável, mas diferem quanto à natureza da evidência: perceber que uma ferramenta ajuda não equivale a demonstrar que ela melhora o desempenho. Estudantes podem avaliar bem uma ferramenta que não altera suas notas, e o inverso também é possível. Por isso, os achados devem ser tomados como ponto de partida para estudos com medidas objetivas, e não como prova de eficácia.

A associação mais forte entre frequência de uso e percepção ocorreu nas afirmativas ligadas ao entendimento (Q2; $\rho = 0{,}68$) e à identificação de erros (Q3; $\rho = 0{,}56$). Uma interpretação possível, que este estudo não permite testar, é que o benefício do feedback se acumula com a prática repetida. A confiança (Q4) teve associação moderada ($\rho = 0{,}43$; $p = 0{,}059$), menor que a de Q2 e Q3, e a motivação e autonomia (Q5; $\rho = 0{,}20$) e a usabilidade (Q6; $\rho = 0{,}24$) não apresentaram associação significativa. Com $n = 20$, porém, os testes têm baixo poder estatístico, e a ausência de significância é inconclusiva: não permite afirmar que essas dimensões independam da frequência de uso, nem que se formem logo nos primeiros contatos. Também não há base nos dados para afirmar que alguma dimensão dependa mais do uso continuado do que as demais.

Os níveis de concordância podem ser lidos à luz de referenciais já apresentados, de forma apenas interpretativa. Bandura (1997) afirma

que a autoeficácia se fortalece principalmente por experiências repetidas de êxito, mas o estudo não acompanhou os acertos dos estudantes e, por isso, não pode avaliar esse mecanismo; a confiança teve a menor média entre as afirmativas (4,15), ainda assim com 85% de concordância. A média mais alta de motivação e autonomia (Q5) é compatível com a ideia de Ryan e Deci (2000) de que ambientes que oferecem retorno sobre a competência e permitem controlar o ritmo apoiam a motivação, ressalvado que Q5 reúne dois aspectos, o que impede saber qual deles pesou mais.

Quanto à facilidade de uso e à clareza das explicações (Q6), a concordância de 90% é compatível com o modelo de Davis (1989), que considera a facilidade de uso percebida um dos determinantes da aceitação de uma tecnologia. O modelo, porém, não trata de como essa percepção varia com a frequência de uso e, por isso, não é empregado aqui para explicar a ausência de associação ($\rho = 0{,}24$; $p = 0{,}314$), que, com $n = 20$, é inconclusiva. Como a afirmativa mistura facilidade de uso e clareza das explicações, não é possível saber qual dos dois aspectos sustentou a concordância.

O relato de P11, que pede à plataforma explicações mais simples, ilustra a adaptação das explicações ao nível do estudante, ideia próxima do apoio calibrado descrito por Vygotsky (1978) na Zona de Desenvolvimento Proximal. Trata-se, porém, de um relato isolado, e o estudo não mediu o nível real dos estudantes nem a adequação do apoio. Do mesmo modo, a presença de comentários sobre o design e o aprendizado gamificado é compatível com a visão de Kapp (2012) e Fardo (2013) de que elementos de jogo favorecem o engajamento, mas os dados não separam o efeito da gamificação do efeito do LLM, e os elogios ao design se referem à experiência de uso, não à aprendizagem.

Quanto à gamificação, a literatura recomenda cautela: os efeitos costumam ser positivos, mas dependem do contexto e dos usuários (Hamari; Koivisto; Sarsa, 2014), e ainda há incerteza sobre o que de fato produz o resultado quando os elementos de jogo vêm combinados a outras intervenções (Dichev; Dicheva, 2017). No presente estudo, a única menção explícita ao aprendizado gamificado partiu de um participante, e os elogios ao design e à interface foram bem mais frequentes. É plausível que a atratividade visual contribua para a percepção favorável, mas os dados não permitem afirmar que ela contribua para a aprendizagem. Deve-se considerar, ainda, a possibilidade de que a percepção positiva tenha sido influenciada por viés de autosseleção e de desejabilidade social, já que os respondentes sabiam que o autor da pesquisa é também o desenvolvedor da plataforma.

A alta positividade das respostas, com média geral de 4,30 e concordância de 85% a 95% em todas as afirmativas, exige leitura crítica. Dois fatores podem ter contribuído para ela. O primeiro é o efeito de novidade, isto é, o entusiasmo inicial que costuma acompanhar o contato com uma tecnologia nova e que tende a diminuir com o uso prolongado; como o levantamento ocorreu ao final de apenas quatro semanas, não se sabe se a avaliação favorável se manteria em períodos maiores. O segundo é a desejabilidade social, agravada pelo fato de a plataforma ter sido desenvolvida pelo próprio pesquisador: ainda que o formulário fosse anônimo, os participantes sabiam quem era o autor e podem ter evitado críticas. Por isso, as médias elevadas devem ser lidas como indício de boa aceitação inicial, e não como medida do valor pedagógico da ferramenta.

A literatura recente reforça essa cautela. Na revisão de Carvalho et al. (2026), 20 dos 35 estudos usaram surveys, escalas Likert ou entrevistas, o que mostra que a avaliação de percepção é comum na área, e apenas 11 empregaram pré-teste, pós-teste ou desempenho final. Os autores apontam, como lacunas, as intervenções curtas, os dados autorrelatados, a ausência de grupos de controle e a escassez de estudos longitudinais, características que o presente estudo também apresenta. Estudos sobre a percepção de estudantes de computação em diferentes países, como os de Asgari et al. (2024) sobre ferramentas digitais e de Haque e Hundhausen (2025) sobre IA generativa, ilustram esse tipo de abordagem, enquanto Gambo et al. (2026) combinam métodos para investigar o efeito do ChatGPT sobre resultados e desempenho de estudantes de programação. Há, ainda, alertas quanto aos riscos: Fan et al. (2025) discutem a chamada preguiça metacognitiva associada ao uso de IA generativa, e Vogel, Ramos e Franzoni (2025) listam a dependência e a redução da autonomia dos estudantes entre os riscos do uso de LLMs.

As sugestões dos participantes, como respostas mais humanizadas, melhor organização das ideias, maior velocidade de resposta e tradução instantânea, apontam caminhos concretos de evolução da plataforma e dialogam com as limitações próprias dos LLMs descritas por Zhang et al. (2023). O instrumento não investigou se o uso da QuestIA gerou dependência de respostas prontas, risco destacado por Kasneci et al. (2023) e pela orientação da UNESCO (Miao; Holmes, 2023); esse é um ponto a ser monitorado em estudos futuros.

Quanto às hipóteses, a percepção de que a QuestIA melhorou o entendimento (85%) e ajudou a identificar erros (95%) está alinhada à H1, que associa o apoio à aprendizagem à interatividade e ao

feedback imediato. Como o estudo não mediu a velocidade de aprendizagem nem comparou a plataforma a outros recursos, a H1 é sustentada apenas quanto à percepção. A H2 também encontra apoio: a utilidade percebida, expressa nas afirmativas sobre entendimento e identificação de erros, e a usabilidade, expressa em Q6, receberam concordância entre 85% e 95%. A associação entre frequência e percepção de benefício ($\rho = 0{,}51$) é um indício de que o uso mais frequente acompanha percepções mais favoráveis, o que não equivale a melhora de desempenho.

5.7. Implicações Pedagógicas e Recomendações

Os resultados trazem implicações para o desenvolvimento da plataforma, para o uso pedagógico e para pesquisas futuras. Para o desenvolvimento, as sugestões dos participantes indicam prioridades concretas: reduzir o tempo de resposta, tornar as explicações mais naturais e bem organizadas e incluir recursos de tradução. Recomenda-se, ainda, que a própria plataforma registre dados de uso, como número de sessões, de exercícios resolvidos e de acertos ao longo do tempo. Esses registros substituiriam a frequência autodeclarada, que neste estudo não teve unidade de tempo definida, e permitiriam relacionar o uso ao desempenho com maior precisão.

Para o uso pedagógico, os achados sugerem que a QuestIA seja tratada como complemento ao trabalho do professor, e não como substituta. Conforme alertam Kasneci et al. (2023) e Miao e Holmes (2023), o acesso fácil a respostas prontas pode prejudicar o pensamento crítico e gerar dependência. Uma forma de mitigar esse risco é combinar o uso da plataforma com atividades em que o estudante precise justificar seus passos de raciocínio e discutir os

erros em sala, de modo que o retorno automático alimente a mediação docente, e não a dispense. Também é recomendável que os estudantes sejam informados sobre os limites da ferramenta, em especial sobre a possibilidade de respostas incorretas (Zhang et al., 2023), e que os dados coletados pela plataforma sejam tratados com transparência e com os cuidados exigidos à proteção de dados de estudantes.

Para a pesquisa, recomenda-se um delineamento quase experimental, com pré-teste, pós-teste e grupo de controle, que permita estimar o efeito da plataforma sobre o desempenho em raciocínio lógico. Convém ampliar e caracterizar a amostra, separar em itens distintos os aspectos hoje reunidos em Q5 e Q6, formular as questões abertas de modo que elogios e sugestões sejam registrados em campos próprios, incluir itens redigidos em sentido negativo para reduzir o efeito de aquiescência e verificar a concordância entre codificadores na análise de conteúdo. Entrevistas ou grupos focais poderiam, ainda, aprofundar a compreensão das categorias que emergiram das respostas abertas.

5.8. Limitações do Estudo

Os resultados devem ser interpretados com cautela, em razão das seguintes limitações:

a. a amostra é pequena (n = 20) e não probabilística, e não há informações sociodemográficas, como idade, curso, turma e gênero, porque o formulário anônimo não as coletou, o que constitui uma fragilidade do desenho da coleta;

b. não houve pré-teste nem grupo de controle, o que impede atribuir os resultados à plataforma;

c. os dados são autorrelatados, sem notas nem registros de desempenho da plataforma, que guarda o progresso apenas no navegador de cada estudante;

d. a plataforma foi desenvolvida pelo próprio pesquisador, o que, somado à desejabilidade social e ao possível efeito de novidade, pode ter elevado as avaliações;

e. as afirmativas Q5 e Q6 reúnem dois construtos cada (motivação e autonomia; facilidade de uso e clareza das explicações) em uma única pergunta, o que impede análises separadas de cada um deles;

f. a Questão 8 pergunta ao mesmo tempo o que agradou e o que melhorar, e a Questão 1 não especificou a unidade de tempo da frequência de uso;

g. a análise de conteúdo foi conduzida por um único pesquisador, sem verificação de concordância entre codificadores;

h. a plataforma ajusta a dificuldade por uma regra de desempenho e dispõe de tutor que pode operar com ou sem modelo de linguagem, e o estudo não isolou o efeito de cada recurso; e

i. todas as afirmativas do formulário foram redigidas em sentido positivo, o que pode favorecer a concordância dos respondentes.

6. CONCLUSÃO

O estudo responde ao problema de pesquisa ao indicar que, na percepção dos estudantes do IFMA Campus Monte Castelo, a QuestIA contribui para o desenvolvimento do raciocínio lógico e para a aprendizagem em disciplinas de exatas. Os participantes avaliam de modo favorável a melhora do entendimento, a identificação de erros, a confiança, a motivação, a autonomia e a usabilidade da plataforma.

O objetivo geral e os cinco objetivos específicos são atingidos. A maior parte dos participantes usa a plataforma com frequência alta. A percepção de benefício é majoritariamente favorável. A frequência de uso associa-se de forma positiva e moderada à percepção de benefício, com maior força na melhora do entendimento e na identificação de erros. O design, a interface e a gamificação aparecem como pontos fortes, e a tradução instantânea, as respostas mais humanizadas e a maior velocidade de resposta aparecem como sugestões.

A hipótese H1 é confirmada quanto à percepção: os dados sustentam que os estudantes veem na interatividade e no feedback imediato um apoio à aprendizagem, sem que a plataforma tenha sido comparada a outros recursos. A hipótese H2 é confirmada: a utilidade e a usabilidade da plataforma são avaliadas de modo majoritariamente favorável. A pesquisa não coleta notas e, por isso, não permite afirmar efeito sobre o rendimento.

A principal contribuição teórica é reunir evidência empírica inicial sobre a percepção de estudantes de ensino técnico público a respeito de uma plataforma baseada em LLM. A contribuição prática é oferecer ao desenvolvedor indicações para evoluir a plataforma,

em especial a velocidade de resposta, a qualidade e a organização das explicações e os recursos de tradução.

Os resultados têm limites: a amostra é pequena e por conveniência, o levantamento é transversal e conta com um único grupo, os dados são autorrelatados e a plataforma é desenvolvida pelo próprio pesquisador. Esses limites impedem afirmar que a QuestIA causa melhora de desempenho. Estudos futuros devem adotar delineamento quase experimental, com pré-teste, pós-teste e grupo de controle, medidas objetivas de desempenho, amostra maior e caracterizada e verificação de possível dependência de respostas prontas.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON, John R. et al. Cognitive tutors: lessons learned. **The Journal of the Learning Sciences,** v. 4, n. 2, p. 167–207, 1995.

ASGARI, M. et al. Students' perspectives on using digital tools in programming courses: a cross country case study between Sweden and Taiwan. **Discover Education,** v. 3, n. 1, 2024. DOI: 10.1007/s44217-024-00170-2.

BACICH, Lilian; TANZI NETO, Adolfo; TREVISANI, Fernando de Mello. **Ensino híbrido:** personalização e tecnologia na educação. Porto Alegre: Penso, 2015.

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

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

BASSNER, P.; LOTTNER, A.; KRUSCHE, S. Towards understanding the impact of context-aware AI tutors and general-purpose AI chatbots on student learning. In: KOLI CALLING INTERNATIONAL CONFERENCE ON COMPUTING EDUCATION RESEARCH, 25., 2025, Koli. **Proceedings** [...]. [S. l.]: ACM, 2025. DOI: 10.1145/3769994.3770025.

BLOOM, Benjamin S. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. **Educational Researcher,** v. 13, n. 6, p. 4–16, 1984.

BROWN, Tom B. et al. Language models are few-shot learners. **Advances in Neural Information Processing Systems,** v. 33, p. 1877–1901, 2020.

CARVALHO, D. R. A. et al. Large Language Models (LLMs) e aprendizagem adaptativa no ensino superior de computação: uma revisão sistemática da literatura. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 37., 2026, [s. l.]. **Anais** [...]. [S. l.]: SBC, 2026. p. 160–175. DOI: 10.5753/sbie.2026.26807.

CRESWELL, John W. **Projeto de pesquisa:** métodos qualitativo, quantitativo e misto. 3. ed. Porto Alegre: Artmed, 2010.

CRONBACH, Lee J. Coefficient alpha and the internal structure of tests. **Psychometrika,** v. 16, n. 3, p. 297–334, 1951.

DAVIS, Fred D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS Quarterly,** v. 13, n. 3, p. 319–340, 1989.

DETERDING, Sebastian et al. From game design elements to gamefulness: defining gamification. In: INTERNATIONAL ACADEMIC

MINDTREK CONFERENCE, 15., 2011, Tampere. **Proceedings [...].** New York: ACM, 2011. p. 9–15.

DICHEV, Christo; DICHEVA, Darina. Gamifying education: what is known, what is believed and what remains uncertain: a critical review. **International Journal of Educational Technology in Higher Education,** v. 14, art. 9, 2017.

FAN, Y. et al. Beware of metacognitive laziness: effects of generative artificial intelligence on learning motivation, processes, and performance. **British Journal of Educational Technology,** v. 56, n. 2, p. 489–530, 2025. DOI: 10.1111/bjet.13544.

FARDO, Marcelo Luis. A gamificação aplicada em ambientes de aprendizagem. **Renote:** Revista Novas Tecnologias na Educação, v. 11, n. 1, 2013.

GAMBO, O. et al. Impact of ChatGPT on learning outcomes and performance of students in computer programming courses: a mixed-method approach. **Discover Education,** v. 5, art. 164, 2026. DOI: 10.1007/s44217-026-01142-4.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Atlas, 2002.

HAMARI, Juho; KOIVISTO, Jonna; SARSA, Harri. Does gamification work? A literature review of empirical studies on gamification. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 47., 2014, Waikoloa. **Proceedings** [...]. Piscataway: IEEE, 2014. p. 3025–3034.

HAQUE, S.; HUNDHAUSEN, C. Generative AI access, usage, and perceptions: an empirical comparison of computing students in the United States and Bangladesh. In: ACM CONFERENCE ON INTERNATIONAL COMPUTING EDUCATION RESEARCH, 2025, Charlottesville. **Proceedings** [...]. [S. l.]: ACM, 2025. v. 1. DOI: 10.1145/3702652.3744231.

HATTIE, John; TIMPERLEY, Helen. The power of feedback. **Review of Educational Research,** v. 77, n. 1, p. 81–112, 2007.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Divulgados os resultados do Pisa 2022.** Brasília, DF: Inep, 2023. Disponível em: https://www.gov.br/inep/pt-br/assuntos/noticias/acoes-internacionais/divulgados-os-resultados-do-pisa-2022. Acesso em: 5 nov. 2026.

JOHNSON-LAIRD, Philip N. **How we reason.** Oxford: Oxford University Press, 2006.

KAPP, Karl M. **The gamification of learning and instruction.** San Francisco: Pfeiffer, 2012.

KASNECI, Enkelejda et al. ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences,** v. 103, art. 102274, 2023.

KULIK, James A.; FLETCHER, J. D. Effectiveness of intelligent tutoring systems: a meta-analytic review. **Review of Educational Research,** v. 86, n. 1, p. 42–78, 2016.

LEWIS, Patrick et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. **Advances in Neural Information**

Processing Systems, v. 33, p. 9459–9474, 2020.

LIKERT, Rensis. A technique for the measurement of attitudes. **Archives of Psychology,** v. 22, n. 140, p. 1–55, 1932.

LODGE, J. M.; THOMPSON, K.; CORRIN, L. Mapping out a research agenda for generative artificial intelligence in tertiary education. **Australasian Journal of Educational Technology,** v. 39, n. 1, p. 1–8, 2023.

LUCKIN, Rose et al. **Intelligence unleashed:** an argument for AI in education. London: Pearson, 2016.

MA, S. et al. DBox: scaffolding algorithmic programming learning through learner-LLM co-decomposition. In: CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2025, Yokohama. **Proceedings** [...]. [S. l.]: ACM, 2025. DOI: 10.1145/3706598.3713748.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica.** 5. ed. São Paulo: Atlas, 2003.

MIAO, Fengchun; HOLMES, Wayne. **Guidance for generative AI in education and research.** Paris: UNESCO, 2023.

OUYANG, Long et al. Training language models to follow instructions with human feedback. **Advances in Neural Information Processing Systems,** v. 35, p. 27730–27744, 2022.

PARDOS, Z. A.; BHANDARI, S. ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. **PLOS ONE,** v. 19, n. 5, p. e0304013, 2024.

POLYA, George. **How to solve it:** a new aspect of mathematical method. Princeton: Princeton University Press, 1945.

RADFORD, Alec et al. Language models are unsupervised multitask learners. **OpenAI Blog,** v. 1, n. 8, p. 1–24, 2019.

RYAN, Richard M.; DECI, Edward L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. **American Psychologist,** v. 55, n. 1, p. 68–78, 2000.

SHI, Y. et al. Large language models in education: a systematic review of empirical applications, benefits, and challenges. **Computers and Education: Artificial Intelligence,** v. 10, p. 100529, 2026. DOI: 10.1016/j.caeai.2025.100529.

SHUTE, Valerie J. Focus on formative feedback. **Review of Educational Research,** v. 78, n. 1, p. 153–189, 2008.

SPEARMAN, Charles. The proof and measurement of association between two things. **The American Journal of Psychology,** v. 15, n. 1, p. 72–101, 1904.

SPRIGGS, K.; LAU, M. C.; PASSI, K. Personalizing education through an adaptive LMS with integrated LLMs. **arXiv,** 2025. DOI: 10.48550/arXiv.2502.08655.

SWELLER, John. Cognitive load during problem solving: effects on learning. **Cognitive Science,** v. 12, n. 2, p. 257–285, 1988.

TABER, Keith S. The use of Cronbach's alpha when developing and reporting research instruments in science education. **Research in Science Education,** v. 48, n. 6, p. 1273–1296, 2018.

TAVAKOL, Mohsen; DENNICK, Reg. Making sense of Cronbach's alpha. **International Journal of Medical Education,** v. 2, p. 53–55, 2011.

VANLEHN, Kurt. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. **Educational Psychologist,** v. 46, n. 4, p. 197–221, 2011.

VASWANI, Ashish et al. Attention is all you need. **Advances in Neural Information Processing Systems,** v. 30, 2017.

VOGEL, D.; RAMOS, A. M.; FRANZONI, A. M. B. Transformando a educação com Large Language Models (LLMs): benefícios, limitações e perspectivas. **Revista Caderno Pedagógico,** Curitiba, v. 22, n. 4, p. 1–20, 2025. DOI: 10.54033/cadpedv22n4-013.

VYGOTSKY, Lev S. **Mind in society:** the development of higher psychological processes. Cambridge: Harvard University Press, 1978.

WIGGINS, Grant. Seven keys to effective feedback. **Educational Leadership,** v. 70, n. 1, p. 10–16, 2012.

WING, Jeannette M. Computational thinking. **Communications of the ACM,** v. 49, n. 3, p. 33–35, 2006.

ZAWACKI-RICHTER, Olaf et al. Systematic review of research on artificial intelligence applications in higher education: where are the educators? **International Journal of Educational Technology in Higher Education,** v. 16, art. 39, 2019.

ZHANG, Yue et al. Siren's song in the AI ocean: a survey on hallucination in large language models. **arXiv preprint,**

arXiv:2309.01219, 2023.

ZIMMERMAN, Barry J. Becoming a self-regulated learner: an overview. **Theory Into Practice,** v. 41, n. 2, p. 64–70, 2002.

---

[1] Discente do Curso Superior de Sistemas de Informação do Centro Universitário Santa Terezinha (CEST), São Luís, MA. E-mail: acesse o artigo original para visualizar o e-mail

[2] Docente do Curso Superior de Sistemas de Informação do Centro Universitário Santa Terezinha (CEST), São Luís, MA, orientador. E-mail: acesse o artigo original para visualizar o e-mail