

Análise Comportamental de Textos Argumentativos em Sistemas Tutores Inteligentes: Estudo de Caso no Exército Brasileiro

Behavioral Analysis of Argumentative Texts in Intelligent Tutoring Systems: A Case Study in the Brazilian Army

Cláudia Rödel Bosaipo Sales da Silva¹, Sarah Vitória Luiz Vanderei², Alberto Sulaiman Sade Junior³, Ronaldo Ribeiro Goldschmidt⁴

1. Doutoranda de Engenharia de Defesa do Instituto Militar de Engenharia (PGED-IME). ORCID: [0000-0002-6694-6750](https://orcid.org/0000-0002-6694-6750). 2. Mestranda em Engenharia, Sistemas e Computação do Instituto Militar de Engenharia (PGSC-IME). ORCID: [0009-0006-5493-795x](https://orcid.org/0009-0006-5493-795x). 3. Pós-doutorando de Sistemas e Computação do Instituto Militar de Engenharia (PGSC-IME). ORCID: [0009-0007-3510-9531](https://orcid.org/0009-0007-3510-9531). 4. Doutor em Engenharia Elétrica. Docente do curso de Pós-Graduação de Engenharia de Defesa do IME. [0000-0003-1688-0586](https://orcid.org/0000-0003-1688-0586).

rodel.claudia@ime.eb.br

Palavras-chave

Aprendizagem Baseada em Competências
 Inteligência Artificial
 Modelos de Linguagem de Grande Porte
 Sistemas Tutores Inteligentes

Keywords

Competency-Based Learning
 Artificial Intelligence
 Large-Scale Language Models
 Intelligent Tutoring Systems

Artigo recebido em: 15.10.2025.

Aprovado para publicação em: 07.11.2025.

Resumo:

A Aprendizagem Baseada em Competências (AC) visa promover o desenvolvimento integrado dos conhecimentos, habilidades e atitudes dos alunos, todos essenciais para a prática profissional. Embora os Sistemas Tutores Inteligentes (STI) tenham apoiado a AC, fornecendo instrução personalizada, a maioria das pesquisas nessa área tem se concentrado nas dimensões de conhecimento e habilidade, deixando a dimensão atitudinal amplamente inexplorada. Embora alguns estudos tenham investigado a inferência de atitudes e comportamentos a partir de dados não estruturados, como textos argumentativos, nenhum o fez no contexto dos STI aplicados à AC. Neste contexto, o presente estudo analisa os resultados preliminares do feedback comportamental gerado por Modelos de Linguagem de Grande Porte (LLM) com base em textos argumentativos escritos por alunos militares em um curso técnico de aviação oferecido pelo Exército Brasileiro. Os resultados sugerem que os LLM são promissores como possíveis ferramentas para apoiar a avaliação atitudinal em ambientes STI por meio da análise da escrita dos alunos.

Abstract:

Competency-Based Learning (CBL) aims to foster the integrated development of students' knowledge, skills, and attitudes, all of which are essential for professional practice. Although Intelligent Tutoring Systems (ITS) have supported CBL by providing personalized instruction, most research in this area has focused on the knowledge and skill dimensions, leaving the attitudinal dimension largely unexplored. While some studies have investigated the inference of attitudes and behaviors from unstructured data, such as argumentative texts, none have done so within the context of ITS applied to CBL. In this context, the present study analyzes preliminary results of behavioral feedback generated by Large Language Models (LLM) based on argumentative texts written by military students in a technical aviation course offered by the Brazilian Army. The findings suggest that LLM hold promise as possible tools to support attitudinal assessment within ITS environments through the analysis of student writing.

INTRODUÇÃO

A aprendizagem baseada em competências (AC) vai além da simples assimilação de conteúdos ao buscar desenvolver junto aos alunos, de forma integrada, as dimensões de *Conhecimentos* (o quê se conhece), *Habilidades* (o saber fazer) e *Atitudes* (o saber ser), essenciais à atuação profissional nos dias atuais (PERRENOUD, 1999). Essa perspectiva é sintetizada pelo acrônimo CHA, que destaca a interdependência dessas três dimensões na formação integral dos estudantes (ZABALA; ARNAU, 2020). Para potencializar essa abordagem, o uso de Tecnologias da Informação e Comunicação (TIC) tem se mostrado promissor (ELIAS et al., 2020; SANTIAGO et al., 2018; MELO, 2018). Entre elas, destacam-se os Sistemas Tutores Inteligentes (STI), que fazem uso de inteligência artificial (IA) para proporcionar um ensino personalizado (GAVIDIA; ANDRADE, 2003), podendo usar diferentes mídias, tais como textos (GRAESSER et al., 2004; SILVA et al., 2024), áudios (WOOLF, 2010), imagens (AZEVEDO et al., 2009) ou mapas mentais (SILVA et al., 2024).

A maior parte dos estudos sobre STI se concentra no desenvolvimento de conhecimentos e habilidades junto aos alunos. Em geral, a dimensão atitudinal é pouco explorada. Neste cenário, é comum o uso de informações estruturadas extraídas de ambientes virtuais de aprendizagem como, por exemplo, datas de entrega de atividades e acesso aos recursos computacionais oferecidos. Tais informações são utilizadas como insumos para inferir acerca de aspectos comportamentais, tais como pontualidade e engajamento (HECKMAN; KAUTZ, 2013; ANDRADE; DU, 2007). Por outro lado, embora existam trabalhos voltados à inferência de atitudes e comportamentos a partir de dados não estruturados como textos dissertativo-argumentativos (GILARDI; ALIZADEH; KUBLI, 2023; HENKLAIN; CARMO, 2013; IMAMOVIC et al., 2024; YU et al., 2024), nenhum deles investigou tal possibilidade no contexto de STI na AC.

Diante do exposto, o presente trabalho teve como objetivo analisar os resultados preliminares obtidos a partir de *feedbacks* sobre aspectos comportamentais inferidos por um STI na AC do EB, com base em textos dissertativo-argumentativos escritos por discentes. Obtidos a partir da aplicação de Modelos de Linguagem (LLMs), tais resultados foram comparados com os *feedbacks* providos por instrutores humanos a partir de interações presenciais junto aos discentes do estudo de caso realizado. Tal estudo envolveu um curso técnico da área de Aviação do Exército Brasileiro (EB). A escolha por um curso neste contexto deveu-se ao fato de que o EB trabalha com AC em seus cursos/estágios desde o ano de 2012 (Comando do Exército, 2012).

Também foi determinante para a escolha do referido curso, o fato de que os autores da pesquisa sobre o EBTutor, um STI na AC do EB, disponibilizaram, para fins de pesquisa e aprimoramento, o código fonte do sistema (SILVA et al., 2024). Os resultados preliminares obtidos no estudo de caso apontaram para o potencial dos LLMs, como ferramentas de apoio, que podem ser integradas ao STI na avaliação de atitudes e aspectos comportamentais dos discentes, a partir de textos discursivos, em ambientes educacionais.

Este artigo se encontra organizado em mais quatro seções. O EBTutor encontra-se resumido na Seção 2. A Seção 3 aborda a metodologia empregada no presente estudo, detalhando o protocolo adotado. A seguir, na Seção 4 são apresentados os experimentos e os resultados obtidos a partir da pesquisa realizada. As considerações finais são, por último, relatadas na Seção 5.

EBTUTOR

O EBTutor é um sistema tutor inteligente, baseado na arquitetura WEB, para a AC do EB. Foi implementado como um *plugin*¹, integrado ao ambiente virtual de aprendizagem da Instituição, o EBAula, e hospedado na intranet do EB (EBNet). Desenvolvido em PHP (ULLMAN, 2009), utilizou o *framework* Adianti (DALL'OGLIO, 2012) e adotou o PostgreSQL (MILANI, 2008).

A arquitetura do EBTutor foi projetada a partir de uma adaptação do modelo típico de um sistema de *feedback* automatizado, proposto por (DEEVA et al., 2021). Os componentes da arquitetura do EBTutor são apresentados na Figura 1 e resumidos a seguir. Mais detalhes sobre o EBTutor e seu funcionamento podem ser obtidos em (SILVA et al., 2024).

Modelo do Domínio: Responsável por organizar o conhecimento de uma determinada área de estudo, sendo composto por três elementos principais. O primeiro é o conjunto de assuntos a serem aprendidos, definidos conforme os conteúdos da área em questão. Para cada assunto, o modelo inclui o conjunto de habilidades que o aluno deve desenvolver. Além disso, o Modelo de Domínio contempla um conjunto ordenado de questionários, com itens com respostas do tipo múltipla-escolha ou dissertativas. Estes avaliam tanto o domínio de conceitos quanto a capacidade de execução de habilidades específicas exigidas para sua resolução. Cabe ressaltar que, conforme (SILVA et al., 2024), o EBTutor se concentra nas dimensões de conhecimentos e habilidades, caracterizando sua limitação quanto à dimensão atitudinal.

Modelo do Aluno: Componente responsável por registrar o histórico de evidências de aprendizagem de cada aluno ao longo da trajetória formativa. Considera-se evidência de aprendizagem qualquer informação fornecida pelo aluno diretamente ao sistema, como respostas a questionários ou postagens em fóruns, além de dados observados a partir das interações com as plataformas, como frequência de acesso e cumprimento de prazos.

Modelo de Feedback Automatizado: Responsável por gerenciar todas as interações entre o aluno e o EBTutor. Sempre que deseja, o aluno pode solicitar um *feedback* indicando o questionário sobre o qual deseja receber informações. A partir dessa solicitação, o EBTutor consulta o EBAula para obter as evidências de aprendizagem necessárias à formulação do *feedback*, enviando-as para o modelo de Geração de *Feedback*.

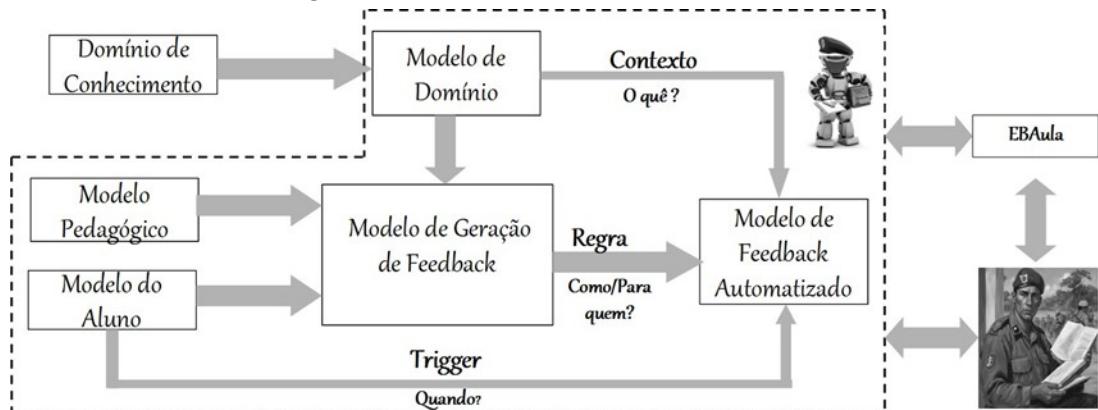
Modelo de Geração de Feedback: Elabora o *feedback* a ser fornecido ao aluno, com base em uma solicitação feita após a realização de um questionário no EBAula. No que diz respeito aos conhecimentos e habilidades, o modelo analisa as respostas fornecidas pelo aluno a cada item do questionário. Caso o aluno necessite de auxílio, o modelo identifica os conceitos necessários, bem como as habilidades que deveriam ser dominadas. Com essas informações, o *feedback* é construído fornecendo orientações de estudo, incluindo os conceitos e habilidades relacionados, com o objetivo de ajudar o aluno a superar suas dificuldades. Após sua geração, o *feedback* é encaminhado ao Modelo de Feedback Automatizado, que o apresenta ao aluno. É neste modelo que podem ser incorporados os LLMs. Diferente dos demais *feedbacks* (conhecimentos e habilidades) providos pelo EBTutor, os *feedbacks* sobre os aspectos comportamentais podem ser dados com o apoio dos LLM, ao analisar as respostas discursivas na forma dissertativa-argumentativa.

Modelo Pedagógico: O Modelo Pedagógico é o componente do EBTutor que incorpora o conhecimento pedagógico necessário à avaliação da aprendizagem e à formulação de *feedbacks* personalizados para os alunos. Para tanto, fundamenta-se em abordagens educacionais como, por exemplo, o *Feedback-Based Learning* (HATTIE; TIMPERLEY, 2007).

METODOLOGIA

O estudo de caso foi realizado em curso da Aviação do EB, na disciplina “Liderança e Gestão”, ofertada para 14 alunos militares. O curso tem como objetivo desenvolver nos discentes militares atitudes e valores tais quais: Comunicação, Decisão, Responsabilidade e Coragem Moral. Vale a pena ressaltar, no estudo de caso realizado, os *feedbacks* não foram fornecidos aos alunos. Eles foram gerados, analisados e comparados com as avaliações providas pelos instrutores humanos. Tal decisão foi tomada para preservar os discentes, evitando expô-los a possíveis alucinações e resultados indesejáveis gerados pelos modelos de linguagem.

Figura 1: Visão Macro-Funcional do EB Tutor.



Fonte (SILVA et al., 2024). Arquitetura do EB Tutor

O LLM aplicado no estudo de caso foi o ChatGPT 4.0, escolhido por sua capacidade de identificar padrões linguísticos que podem indicar traços comportamentais e atitudes (PETERS; MATZ, 2024). Além disso, este LLM demonstrou familiaridade com o contexto da AC (MICHALON; CAMACHO-ZUÑIGA, 2023), o que o torna apto a aprender os princípios das *Normas para Desenvolvimento e Avaliação dos Conteúdos Atitudinais* (NDACA)², empregadas no EB (Departamento de Educação e Cultura do Exército, 2019).

Assim sendo, o LLM escolhido teve condições de interpretar e classificar os padrões comportamentais/attitudinais dos alunos, por meio dos textos dissertativos por eles produzidos. O processo envolveu a elaboração de uma metodologia estruturada em 5 (cinco) fases descritas a seguir:

Fase 1 - Estruturação e Preparação do Corpus. Os dados brutos, compostos pelas respostas abertas dos 14 estudantes militares a oito perguntas do curso de “Liderança e Gestão”, foram estruturados e preparados. As respostas foram reorganizadas em um formato com identificadores anônimos (*'userid'*, *'coluna'* e *'response-text'*) para garantir a análise adequada pelo LLM, respeitando a LGPD.

Fase 2 - Mapeamento de Avaliação Humana. Paralelamente, dois avaliadores experientes avaliaram de forma independente as atitudes dos alunos com base na observação longitudinal e na interação presencial durante o curso. Cada avaliador atribuiu pontuações de 1,0 a 10,0 para quatro construtos: Comunicação, Decisão, Responsabilidade (atitudes) e Coragem Moral (valor). Devido à alta qualidade do desempenho dos alunos, as pontuações se concentraram entre 8,0 e 10,0. As avaliações foram registradas em planilhas separadas para comparação posterior.

Fase 3 - Engenharia de *Prompt* e configuração do LLM. Foram desenvolvidos dois *prompts zero-shot* para o GPT-4, baseados nas descrições do NDACA: um para o valor Coragem Moral e outro focado nas três atitudes. Cada *prompt* foi estruturado para obter uma única pontuação por construto para cada aluno, com resultados em formato tabular ordenado por *'userId'*. Para assegurar a independência total entre as análises humanas e da máquina, o LLM analisou apenas as respostas textuais, sem acesso a dados de treinamento ou avaliações humanas.

Fase 4 - Execução, calibração e mapeamento conceitual do LLM. Os *prompts* foram executados e os resultados iniciais revelaram maior dispersão do que nas avaliações humanas, o que levou a um processo de recalibração. Para garantir a comparabilidade, os *prompts* foram ajustados para restringir os resultados do modelo em um intervalo mais focado - de 8,0 a 10,0, melhorando a consistência da pontuação e o alinhamento com a distribuição observada.

Após a geração da pontuação, os resultados foram normalizados e mapeados em três faixas conceituais para melhorar a interpretabilidade e apoiar a comparação categórica entre os avaliadores: A (9,3–10,0): Excelente alinhamento atitudinal / B (8,6–9,2): Desempenho bom, mas não excelente / C (8,0–8,5): Expressão satisfatória, mas limitada.

Fase 5 - Estruturação dos resultados. A fase final concentrou-se na consolidação dos dados da avaliação em formatos estruturados para apoiar a análise comparativa apresentada na próxima seção. Foram produzidas duas tabelas de resumo para refletir a natureza dupla dos resultados da avaliação: Tabela 1, que exibe as pontuações numéricas brutas atribuídas pelo Avaliador Humano 1 (H1), Avaliador Humano 2 (H2) e LLM para cada um dos quatro construtos; e Tabela 2, que apresenta os mesmos dados em um formato discretizado, usando o framework conceitual A-B-C.

Para promover a rastreabilidade e a consistência na interpretação dos resultados entre dimensões e avaliadores, os cabeçalhos das colunas seguem o formato X-Y, em que X denota o construto e Y identifica o avaliador.

Tabela 1: Pontuações numéricas brutas atribuídas pelos Avaliadores Humanos 1 e 2 e o LLM

Aluno	C-H1	C-H2	C-LLM	D-H1	D-H2	D-LLM	R-H1	R-H2	R-LLM	MC-H1	MC-H2	MC-LLM
1	10,00	9,00	8,70	9,67	9,67	8,40	10,00	10,00	9,20	9,67	9,67	8,70
2	10,00	9,33	9,90	9,80	9,80	8,40	10,00	9,00	8,10	9,67	9,67	9,90
3	9,33	10,00	9,50	9,50	9,50	8,60	9,33	10,00	9,20	9,67	9,67	9,50
4	10,00	10,00	9,20	9,67	9,67	9,00	9,33	8,67	8,30	10,00	10,00	9,20
5	9,67	10,00	8,30	9,33	9,33	8,90	9,33	10,00	8,10	10,00	10,00	8,30
6	9,00	10,00	8,30	9,50	9,50	8,60	10,00	10,00	9,90	9,50	9,50	8,30
7	10,00	9,33	8,10	9,67	9,67	9,20	10,00	9,33	9,90	9,67	9,67	8,10
8	9,67	9,67	9,70	9,67	9,67	8,30	9,33	10,00	9,60	10,00	10,00	9,70
9	9,00	8,67	9,20	8,50	8,50	8,60	10,00	10,00	8,60	9,50	9,50	9,20
10	10,00	9,00	9,40	9,67	9,67	8,70	10,00	10,00	8,20	9,67	9,67	9,40
11	9,33	10,00	8,00	9,67	9,67	8,90	9,33	10,00	9,40	9,67	9,67	8,00
12	10,00	9,33	9,90	9,50	9,50	9,60	10,00	9,67	8,90	9,67	9,67	9,90
13	9,67	10,00	9,70	9,50	9,50	8,40	9,33	10,00	8,20	10,00	10,00	9,70
14	10,00	10,00	8,40	9,67	9,67	9,00	9,67	9,00	9,00	9,67	9,67	8,40

Tabela 2: Dados discretizados nas categorias A–B–C

Aluno	C-H1	C-H2	C-LLM	D-H1	D-H2	D-LLM	R-H1	R-H2	R-LLM	MC-H1	MC-H2	MC-LLM
1	A	B	B	A	A	C	A	A	B	A	A	B
2	A	A	A	A	A	C	A	B	C	A	A	A
3	A	A	A	A	A	B	A	A	B	A	A	A
4	A	A	B	A	A	B	A	B	C	A	A	B
5	A	A	C	A	A	B	A	A	C	A	A	C
6	B	A	C	A	A	B	A	A	A	A	A	C
7	A	A	C	A	A	B	A	A	A	A	A	C
8	A	A	A	A	A	C	A	A	A	A	A	A
9	B	B	B	C	C	B	A	A	B	A	A	B
10	A	B	A	A	A	B	A	A	C	A	A	A
11	A	A	C	A	A	B	A	A	A	A	A	C
12	A	A	A	A	A	A	A	A	B	A	A	A
13	A	A	A	A	A	C	A	A	C	A	A	A
14	A	A	C	A	A	B	A	B	B	A	A	C

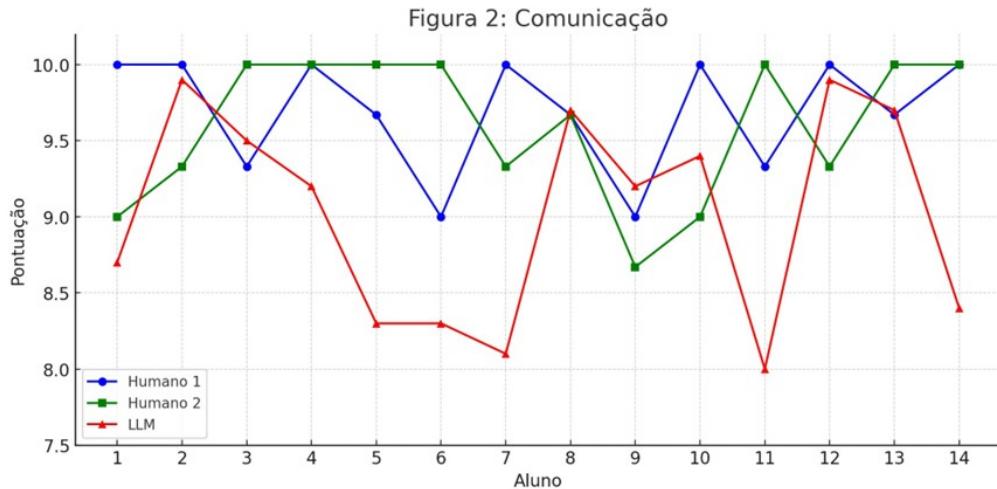
RESULTADOS E DISCUSSÃO

Esta seção está organizada em duas subseções. A primeira compara as pontuações dos avaliadores numericamente, e a segunda analisa a concordância com base nas categorias conceituais.

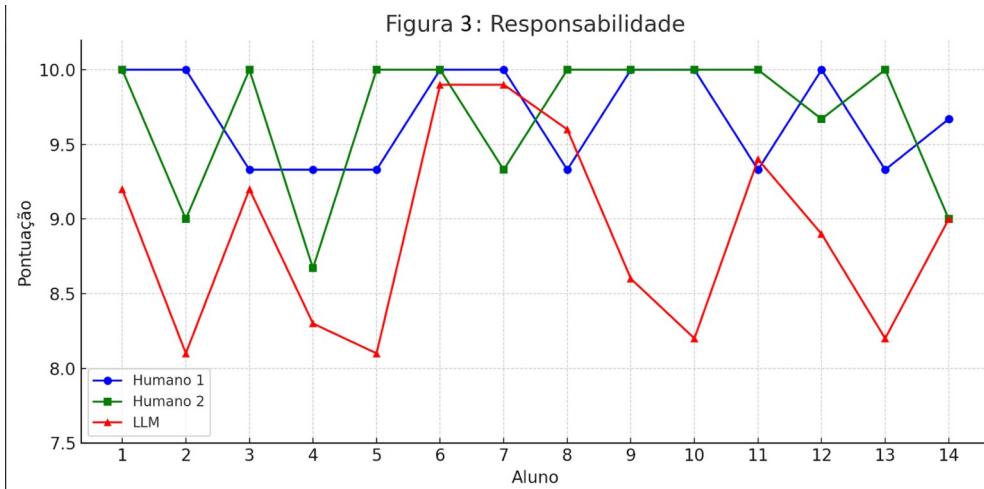
COMPARAÇÃO DO NÍVEL DE PONTUAÇÃO

As Figuras 2 a 5 mostram as tendências das pontuações dos alunos para cada constructo. Os avaliadores Humanos são mostrados em azul e verde e o LLM em vermelho.

A Figura 2 (Comunicação) revela um alinhamento próximo entre os dois avaliadores humanos, com pontuações altas e estáveis. O LLM atribui pontuações mais baixas em geral, especialmente para os alunos 1, 5, 6, 7 e 11, demonstrando um padrão de pontuação mais conservador.



A Figura 3 (Responsabilidade) ilustra forte alinhamento entre os avaliadores humanos, enquanto o LLM diverge em vários casos, notadamente para os alunos 2, 4, 5 e 13, aos quais atribui pontuações mais baixas.



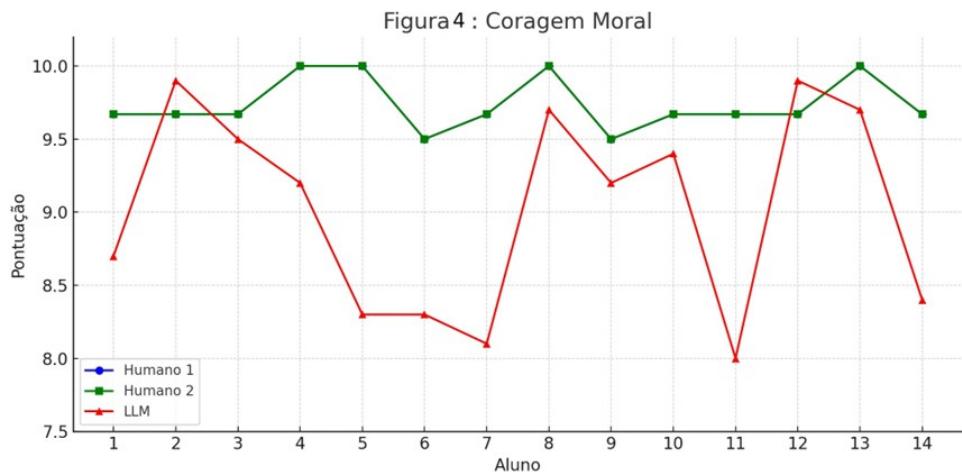
A Figura 4 (Coragem Moral) confirma uma alta concordância entre os seres humanos e destaca a tendência do LLM de obter pontuações mais baixas, particularmente para os alunos 1, 5, 7 e 11.

Em suma, as Figuras 2 e 3 ilustram graficamente os resultados apresentados na Tabela 1. Para quantificar essas relações, foram calculados os coeficientes de correlação de Pearson entre os três avaliadores. A Tabela 3 apresenta esses coeficientes por dimensão. Como esperado, as correlações mais fortes ocorrem entre o Humano 1 e o Humano 2. Em contrapartida, as correlações entre o LLM e os avaliadores humanos são geralmente baixas, sugerindo uma concordância linear limitada.

Tabela 3: Coeficientes de correlação de Pearson por construto

Dimensão	H1 vs H2	H1 vs LLM	H2 vs LLM
Comunicação	-0,134	0,202	-0,251
Decisão	1	0,04	0,04
Responsabilidade	-0,06	0,129	0,202
Coragem Moral	1	0,22	0,22

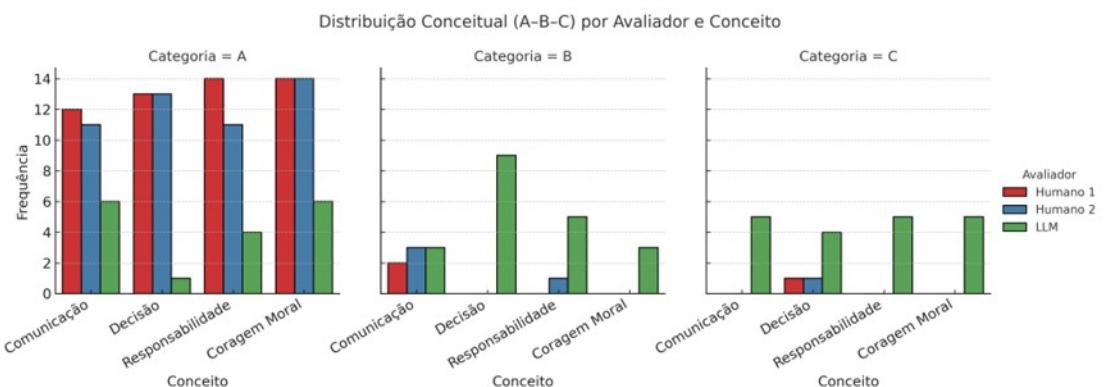
Esses resultados destacam uma diferença metodológica fundamental entre a avaliação baseada em observação e a avaliação baseada em texto. Enquanto os avaliadores humanos mostram alta consistência — particularmente em construções com expressão comportamental clara — o LLM demonstra maior variabilidade e uma tendência para pontuações médias.



COMPARAÇÃO EM NÍVEL CONCEITUAL (A–B–C)

Para facilitar a comparação categórica, as pontuações originais foram transformadas em níveis conceituais usando o esquema A–B–C descrito na Seção 3: A Figura 5 apresenta a distribuição das etiquetas conceituais entre os quatro constructos. Os avaliadores humanos mostram forte convergência nas categorias A e B, particularmente em Decisão e Coragem Moral. O LLM, em contraste, favorece as categorias B e C, reforçando seu comportamento conservador na pontuação.

Figura 5: Distribuição conceitual (A–B–C) por avaliador e construto
Figura da comparação da distribuição conceitual (A–B–C) por avaliador e construto



Para avaliar a concordância categórica, calculamos a porcentagem de alunos aos quais cada par de avaliadores atribuiu o mesmo conceito, conforme ilustrado na Tabela 4. Observou-se concordância total entre os avaliadores humanos em Decisão e Coragem Moral. A concordância entre o LLM e qualquer um dos avaliadores humanos foi consistentemente menor, particularmente em Comunicação e Responsabilidade.

Tabela 4: Concordância entre avaliadores (%) por categoria conceitual

Dimensão	H1 = H2	H1 = LLM	H2 = LLM	% H1 = H2	% H1 = LLM	% H2 = LLM
Comunicação	11	7	7	78,6	50	50
Decisão	14	1	1	100	7,1	7,1
Responsabilidade	11	4	5	78,6	28,6	35,7
Coragem Moral	14	6	6	100	42,9	42,9

Esses resultados indicam que, embora o LLM possa se aproximar das categorizações humanas em determinados contextos, particularmente em conceitos como Responsabilidade, que são mais baseados no comportamento, ele apresenta limitações perceptíveis em domínios que exigem sensibilidade afetiva ou raciocínio moral.

CONSIDERAÇÕES FINAIS

Foi identificada uma carência de estudos quanto à inferência de atitudes e comportamentos discente a partir de dados não estruturados, como textos argumentativos, especialmente no contexto de Sistemas Tutores Inteligentes aplicados à Aprendizagem Baseada em Competências.

Diante desse cenário, este artigo apresentou os resultados preliminares obtidos a partir da análise de textos dissertativos-argumentativos escritos por alunos de um curso técnico da área de Aviação do EB. A análise foi conduzida por meio do Modelo de Linguagem ChatGPT 4.0, e seus resultados foram comparados com os *feedbacks* elaborados por instrutores humanos, que acompanharam presencialmente os discentes ao longo do curso. Cabe destacar que foi utilizado o referencial NDACA como base normativa para ambas as avaliações.

O LLM demonstrou ser ferramenta promissora de apoio à avaliação formativa, ao gerar *feedbacks* personalizados e contextualizados, sem a identificação de alucinações. Os resultados revelaram um padrão conservador nas respostas, indicando robustez suficiente para sua incorporação junto aos STIs. Como sua contribuição principal, este estudo pode demonstrar a viabilidade da utilização de LLM para inferência de atitudes, a partir da produção escrita dos alunos.

Atualmente, estão sendo realizados experimentos para integrar os LLMs aos STIs que forneçam *feedback* em tempo real aos alunos. Paralelamente, está em andamento a coleta das percepções dos estudantes sobre a utilidade e eficácia desses *feedbacks*, visando validar e aprimorar o modelo.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível superior - Brasil (CAPES) - Código de financiamento 001 (bolsa de mestrado - proc. 88887.196477/2025-00).

NOTAS

1. Software que adiciona funcionalidade específica a programa ou sistema maior.
2. Estas Normas se destinam aos cursos e estágios gerais das Linhas de Ensino Militar Bélico, Complementar e de Saúde, realizados nos Estabelecimentos de Ensino (Estb Ens) e Organizações Militares (OM) com encargos de ensino subordinados e/ou vinculados ao Departamento de Educação e Cultura do Exército (DECEEx), com exceção da Diretoria de Educação Preparatória e Assistencial (DEPA), com legislação própria.

REFERÊNCIAS

- ANDRADE, H.; DU, Y. Student responses to criteria-referenced self-assessment. *Assessment & evaluation in higher education*, Taylor & Francis, v. 32, n. 2, p. 159–181, 2007.
- AZEVEDO, R. et al. **Metatutor**: A metacognitive tool for enhancing self-regulated learning. In: *AAAI fall symposium: Cognitive and metacognitive educational systems*. [S.l.: s.n.], 2009. p. 14–19.
- SILVA, C.R.B.S. et al. Análise Comportamental de Textos Argumentativos em Sistemas Tutores Inteligentes: Estudo de Caso no Exército Brasileiro. Pleiade, 19(49): 56-64, Out.-Dez., 2025 DOI: 10.32915/pleiade.v19i49.1197

COMANDO DO EXÉRCITO. **Portaria nº 137**, de 28 de fevereiro de 2012: Aprova a Diretriz para o Projeto de Implantação do Ensino por Competências no Exército Brasileiro. 2012. Disponível em: <<https://www.eb.mil.br>> [Acesso conforme aplicável]. Publicado no Diário Oficial da União, 2012.

DALL'OGLIO, P. **Adianti framework para php**. Lajeado: Edição do autor, 2012.

DEEVA, G. et al. A review of automated feedback systems for learners: Classification framework, challenges and opportunities. **Computers & Education**, Elsevier, v. 162, p. 104094, 2021.

DEPARTAMENTO DE EDUCAÇÃO E CULTURA DO EXÉRCITO. **Normas para Desenvolvimento e Avaliação de Conteúdos Atitudinais**. 2019. EB60-N-05.013.

ELIAS, A. et al. Avaliar: **Sistema para autoria e acompanhamento de recursos avaliativos**. In: SBC. *Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*. [S.I.], 2020. p. 97–101.

GAVIDIA, J. J. Z.; ANDRADE, L. C. V. d. **Sistemas tutores inteligentes**. Trabalho de Conclusão da Disciplina de IA, Universidade Federal do Rio de Janeiro. Rio de Janeiro–RJ: UFRJ, 2003.

GILARDI, F.; ALIZADEH, M.; KUBLI, M. Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences, National Academy of Sciences*, v. 120, n. 30, p. e2305016120, 2023.

GRAESSER, A. C. et al. Autotutor: A tutor with dialogue in natural language. **Behavior Research Methods, Instruments, & Computers**, Springer, v. 36, n. 2, p. 180–192, 2004.

HATTIE, J.; TIMPERLEY, H. The power of feedback. **Review of educational research**, Sage Publications Sage CA: Thousand Oaks, CA, v. 77, n. 1, p. 81–112, 2007.

HECKMAN, J. J.; KAUTZ, T. **Fostering and measuring skills**: Interventions that improve character and cognition. National Bureau of economic research, 2013.

HENKLAIN, M. H. O.; CARMO, J. d. S. Contributions of behavior analysis to education: an invitation for dialogue. **Cadernos de Pesquisa**, SciELO Brasil, v. 43, p. 704–723, 2013.

IMAMOVIC, M. et al. **Using chatgpt for annotation of attitude within the appraisal theory**: Lessons learned. In: *Proceedings of The 18th Linguistic Annotation Workshop (LAW-XVIII)*. [S.I.: s.n.], 2024. p. 112–123.

MELO, G. A. M. S. **Ballgorithm** - uma ferramenta introdutória para conceitos de programação. Universidade Federal do Maranhão, 2018.

MICHALON, B.; CAMACHO-ZUÑIGA, C. Chatgpt, a brand-new tool to strengthen timeless competencies. In: FRONTIERS MEDIA SA. **Frontiers in Education**. [S.I.], 2023. v. 8, p. 1251163.

MILANI, A. **Postgresql-guia do programador**. [S.I.]: Novatec Editora, 2008.

PERRENOUD, P. Construir competências é virar as costas aos saberes. **Revista Pátio**, Porto Alegre: ARTMED, ano, v. 3, p. 15–19, 1999.

PETERS, H.; MATZ, S. C. Large language models can infer psychological dispositions of social media users. **PNAS nexus**, Oxford University Press US, v. 3, n. 6, p. pgae231, 2024.

SANTIAGO, J. M. S. et al. **Mathquiz**: A game app for m-learning. In: SBC. *Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*. [S.I.], 2018. p. 69–72.

SILVA, C. S. da et al. Ebtutor: **Uma proposta de sistema tutor inteligente na aprendizagem baseada em competências no exército brasileiro**. In: SBC. *Simpósio Brasileiro de Informática na Educação (SBIE)*. [S.I.], 2024. p. 1674–1688.

ULLMAN, L. **PHP for the Web: Visual QuickStart Guide**. [S.I.]: Peachpit Press, 2009.

WOOLF, B. P. **Building intelligent interactive tutors**: Student-centered strategies for revolutionizing e-learning. [S.I.]: Morgan Kaufmann, 2010.

YU, D. et al. Assessing the potential of llm-assisted annotation for corpus-based pragmatics and discourse analysis: The case of apology. **International Journal of Corpus Linguistics**, John Benjamins Publishing Company Amsterdam/Philadelphia, v. 29, n. 4, p. 534–561, 2024.

ZABALA, A.; ARNAU, L. **Métodos para ensinar competências**. [S.I.]: Penso Editora, 2020.