

Uso de Inteligência Artificial para geração de roteiros de ensino e estudo suportados pela ciência da aprendizagem

Maykell Alvim de França¹, Pedro Calais²

¹Faculdade XP Educação - IGTI- XPE

Caixa Postal 30360-680 N° 561– Belo Horizonte – MG – Brazil

maykell.franca@estudante.xpe.edu.br, pedro.calais@xpe.edu.br

Abstract. *One of the potential uses of generative artificial intelligence in education is leveraging the knowledge people already have to provide personalized learning through tailored teaching and study plans. In this project, using two different AIs, we experimented with prompts that embed well-known learning science theories, such as transfer of learning, problem-based learning, Bloom's taxonomy, and backwards design. We found that, for the most part, personalized teaching and study plans outperformed common prompt-generated plans in criteria such as storytelling, learning flow, adaptability, and inclusivity, enhancing the overall quality of the generated teaching plans.*

Resumo. Um dos potenciais uso de inteligência artificial generativa na Educação é o de aproveitar conhecimentos que as pessoas já possuem, e assim proporcionar um aprendizado personalizado por meio de roteiros de ensino e estudo adaptados. Neste projeto, com duas I.As diferentes, experimentamos *prompts* que embutem teorias conhecidas da ciência da aprendizagem, como transferência de aprendizado, aprendizagem baseada em problemas, taxonomia de Bloom e planejamento reverso, e descobrimos que, majoritariamente, planos de ensino e estudo personalizados se sobressaíram em critérios como *storytelling*, fluxo de aprendizado, adaptação e inclusão, aprimorando a qualidade dos planos de ensino gerados por *prompts* comuns.

1. Introdução

Desde os primórdios, a educação tem sido um pilar fundamental para o desenvolvimento da sociedade e do conhecimento. John Dewey destacou a importância do aprendizado experimental para o engajamento cívico [Theobald 2009], enquanto Jean Piaget enfatizou a interação entre sujeito e objeto na construção do conhecimento [Piaget, apud Becker 2010]. Howard Gardner, por sua vez, propôs a teoria das múltiplas inteligências, defendendo abordagens pedagógicas diversificadas [Gardner 1993].

No século XXI, os avanços tecnológicos e a inteligência artificial (IA) ampliaram as possibilidades pedagógicas, permitindo a personalização do ensino em larga escala. A *ciência da aprendizagem* tem explorado como ambientes adaptativos podem atender às necessidades individuais dos alunos, tornando o processo educacional mais eficaz. No entanto, apesar dessas inovações, estudos recentes [Salas-Pilco et al. 2022] indicam que educadores continuam sendo essenciais para fomentar criatividade e pensamento crítico. A interação entre tecnologia e presença humana na educação levanta, portanto, um questionamento central: como equilibrar esses elementos para maximizar o potencial

educacional? Este trabalho investiga essa questão por meio do uso de Modelos de Linguagem de Grande Escala (*LLMs – Large Language Models*), analisando o impacto da incorporação de teorias pedagógicas no *design de prompts* para a geração de planos de ensino e estudo. Foram utilizados dois modelos estado-da-arte – ChatGPT 4.0 e Gemini 2.0, para gerar e avaliar se *prompts* baseados em princípios da ciência da aprendizagem resultam em materiais mais claros, estruturados e alinhados aos objetivos educacionais.

Os resultados indicam que a inclusão de teorias pedagógicas nos *prompts*, a saber, a taxonomia de Bloom, o aprendizado baseado em problemas, o planejamento reverso e transferência de aprendizado aprimoram a qualidade dos planos gerados nas áreas de Programação e Ciência de Dados. Em alguns cenários, porém, *prompts* mais genéricos se mostraram mais adequados para públicos amplos. Além disso, identificamos diferenças na avaliação dos modelos: o Gemini 2.0 apresentou menor variabilidade nas notas atribuídas em comparação ao ChatGPT 4.0.

O restante do trabalho está estruturado da seguinte forma: a Seção 2 discute trabalhos relacionados; na Seção 3, detalhamos o planejamento e execução dos experimentos; a Seção 4 apresenta e analisa os resultados; e, por fim, a Seção 5 traz as conclusões e sugestões de trabalhos futuros.

2. Trabalhos Relacionados

A busca por métodos que melhorem a eficácia do ensino é uma preocupação constante na pesquisa pedagógica. Pesquisadores têm se concentrado em projetar estratégias de ensino para melhores efeitos educacionais [Kolodner 2002; Luckin et al. 2016; VanLehn 2006]. Desde que o ChatGPT fascinou o mundo em 2022, a atenção à inteligência artificial (IA) na educação ganhou destaque, pois melhorar seu desempenho envolve fornecer contexto sobre seu papel e a tarefa a ser realizada, como informações sobre o curso, a idade dos alunos e seus interesses [Levy; Pérez Albertos 2024]. A IA na educação é derivada de teorias de IA, ciências cognitivas e educação [Fiok et al. 2021; Nguyen et al. 2022; Zawacki-Richter et al. 2019], discutindo questões como: como um aluno pode ser auxiliado no aprendizado? Quais padrões eficazes de interação pedagógica devem ser usados? Quais conceitos errôneos os aprendizes têm [Lehmann 2024, Woolf 2009]? Estudos sobre IA como tutor inteligente mostram que ela pode ajudar alunos a associarem o conteúdo aprendido com conceitos que já compreendem. Iniciativas incluem o uso de IA para elaborar questões de exames [Kic-Drgas 2024], auxiliar em testes de software [Amalfitano 2023] e personalizar a experiência de aprendizado com *prompts* personalizados [Mattalo 2024]. Este projeto foca na criação de roteiros de aprendizado, utilizando a IA para explicar conceitos com base nos interesses dos alunos, promovendo conexões mais profundas entre o aprendizado e suas experiências. Contudo, há uma oportunidade de aprimorar essa abordagem aplicando conceitos pedagógicos mais específicos, e é isso que este projeto busca realizar.

3. Metodologia

Foram selecionadas duas *personas* que desejam utilizar inteligência artificial para ensinar ou aprender de forma mais eficiente – o professor, que deseja gerar um plano de ensino, e o estudante, que deseja gerar um plano de estudo.

Persona Professor: O professor de Ciência da Computação/Ciência de Dados busca ensinar introdução à programação em *Python* e aprendizado de máquina em um plano de ensino de duas horas. Ele deseja aprimorar seus métodos de ensino, adaptando o conteúdo para alunos com diferentes níveis de conhecimento e mantendo o interesse dos estudantes, especialmente aqueles que têm dificuldades. Ao utilizar o ChatGPT 4.0 e o Gemini 2.0, espera criar exemplos personalizados, gerar exercícios práticos desafiadores e obter explicações claras para conceitos complexos.

Persona Estudante: O estudante de ciência da computação ou ciência de dados, que nunca programou, tem facilidade com números e raciocínio lógico. Ele é curioso, autodidata e gosta de resolver desafios matemáticos, buscando aprender programação em *Python* e aprendizado de máquina desde o básico. Ao utilizar o ChatGPT 4.0 e o Gemini 2.0, espera receber explicações simples, exemplos práticos para iniciantes e exercícios básicos para praticar, além de *feedback* imediato sobre erros e dicas para melhorar seu aprendizado.

Selecionamos duas áreas de conhecimento para geração de planos de estudo (para o estudante) e de ensino (para o professor): programação e aprendizado de máquina. Desenhamos um *prompt* que considera quatro teorias clássicas da ciência de aprendizagem e ampliamiento utilizados no campo educacional, a saber:

Transferência de Aprendizado: para se referir à capacidade de aplicar conhecimentos, habilidades ou atitudes adquiridos em um contexto para resolver problemas ou realizar tarefas em outro contexto. Esse método é amplamente estudado na psicologia educacional e na ciência cognitiva [G. Steiner 2002], sendo essencial para o aprendizado significativo e a formação de competências transferíveis.

Backward Design Framework: O planejamento reverso é uma abordagem pedagógica que se concentra em planejar experiências de aprendizagem a partir dos resultados desejados, ou seja, começa pelo fim. Essa metodologia, baseada no *Harvard Project Zero* [Mardell et al, 2016], envolve três etapas principais: identificar os objetivos de aprendizagem, determinar as evidências que demonstrarão que os alunos alcançaram esses objetivos e, por fim, planejar as atividades de ensino que levarão os alunos a esses resultados. Ao adotar essa estratégia, os educadores garantem que o ensino seja intencional e focado, promovendo uma compreensão mais profunda e significativa do conteúdo pelos alunos.

Taxonomia de Bloom: Uma estrutura hierárquica desenvolvida por Benjamin Bloom [*Taxonomy of Educational Objectives*, 1956] e colaboradores para classificar os objetivos educacionais em diferentes níveis de complexidade e especificidade. Ela é amplamente utilizada no planejamento pedagógico, na avaliação da aprendizagem e no desenvolvimento de currículos. As etapas da hierarquia da Taxonomia de Bloom são: **lembrar, compreender, aplicar, analisar, avaliar e criar**, progredindo do nível mais básico de conhecimento para o mais complexo.

Problem-Based Learning: O *Problem-Based Learning* (PBL), ou Aprendizagem Baseada em Problemas, é uma abordagem pedagógica centrada no aluno, na qual o aprendizado ocorre por meio da resolução de problemas complexos e reais [A Brief History of Problem-based Learning, Schmidt, H.G. 2012]. Essa metodologia incentiva os alunos a desenvolverem habilidades de pensamento crítico, colaboração e autonomia. No PBL, os alunos são apresentados a um problema inicial e, em vez de receberem respostas prontas, são incentivados a identificar o problema, pesquisar informações relevantes, propor soluções baseadas em evidências e refletir sobre o processo de aprendizado.

Nesta seção, as Figuras 1 e 2 representam os planos de ensino tradicional e de aprendizagem ativa.

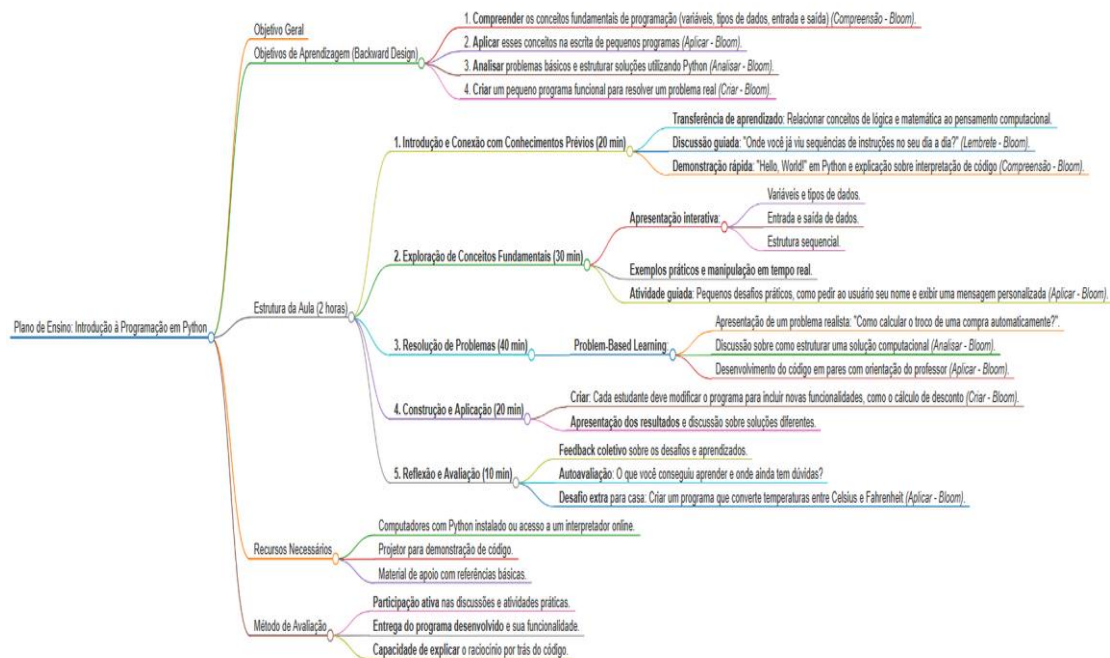


Figura 1: Representação visual de um plano de ensino tradicional



Figura 2: Representação visual de um plano de ensino gerado com teorias pedagógicas

A Figura 1 apresenta um plano de ensino tradicional, estruturado de forma sequencial, com explicações seguidas de exercícios práticos. Ele foca na introdução à sintaxe do Python, abordando variáveis, estruturas condicionais e loops, finalizando com desafios práticos e uma sessão de dúvidas. A Figura 2, por sua vez, adota uma abordagem baseada em aprendizagem ativa e resolução de problemas (PBL), começando com a conexão de conceitos prévios e explorando a programação de forma interativa. Os alunos são incentivados a discutir, analisar e construir soluções reais, promovendo maior engajamento.

Enquanto o primeiro plano prioriza a instrução direta, o segundo segue a Taxonomia de Bloom, focando na compreensão, aplicação, análise e criando soluções. Além disso, o segundo plano define critérios de avaliação claros, como participação ativa e capacidade de explicar o raciocínio. Dessa forma, ele tende a ser mais eficaz para desenvolver pensamento computacional e autonomia na programação.

Para avaliar quantitativamente os planos de ensino/estudo, empregamos o paradigma "*LLM as a Jury*" [Verga Pat *et al* 2024] para avaliar os planos de ensino gerados pelo GPT 4.0 e Gemini 2.0. O paradigma *LLM as a Jury* consiste em utilizar modelos de linguagem de grande escala (LLMs) para avaliar e comparar respostas geradas por IA ou por humanos com base em critérios específicos, simulando um júri que atribui pontuações e justifica suas avaliações. Nesse sentido, a ideia é que o ChatGPT 4.0 e o Gemini 2.0 avaliarão a qualidade dos planos de ensino gerados por eles mesmos, mas também de forma cruzada: isto é, o ChatGPT avaliará seu próprio plano de ensino e o plano de ensino gerado pelo Gemini 2.0, e vice-versa. Cada LLM atuará como um

avaliador imparcial e atribuirá uma pontuação de 1 a 5 para a resposta gerada pela IA com base em critérios específicos de avaliação que estão descritos a seguir:

- 1 - Os objetivos de aprendizagem estão claros e alinhados?
- 2 - Quão eficiente é o *Storytelling*?
- 3 - O quão bem conduz a pessoa ao estado de fluxo?
- 4 - Como está a coerência e estrutura?
- 5 - Como está a flexibilidade e adaptação?
- 6 - Quão inclusivo e acessível é?

Um plano de ensino/estudo bem estruturado e eficaz deve incluir **objetivos de aprendizagem** claros e alinhados com as necessidades do público-alvo, promovendo uma aprendizagem direcionada. É importante também utilizar *storytelling* para estruturar a informação de forma envolvente, conectando emocionalmente os usuários e facilitando a compreensão. A experiência deve conduzir os usuários ao **estado de fluxo**, mantendo-os imersos e engajados em desafios equilibrados. Além disso, a **coerência e estrutura** do conteúdo são essenciais para garantir uma organização lógica que facilite a navegação. A **flexibilidade e adaptação** permitem ajustes para diferentes perfis de usuários e ritmos de aprendizado, enquanto a **inclusão e acessibilidade** garantem que a experiência seja acessível a todos, independentemente de suas habilidades ou condições, utilizando práticas que promovam a equidade. O júri de *LLMs* deverá atribuir uma pontuação e escrever uma breve explicação justificando a decisão com base nestes critérios. Para gerar os planos de estudo e de ensino, foram criados vários prompts, a saber:

- **Prompt Simples:** “Forneça-me um plano de ensino de duas horas para ensinar (introdução à programação em *python*/introdução ao aprendizado de máquina) para meus estudantes.”. O objetivo deste *prompt* é permitir avaliar o que *prompts* que usam teorias pedagógicas agregam.
- **Prompt com uma teoria específica:** Forneça-me um plano de ensino de duas horas para ensinar (introdução à programação em *python*/introdução ao aprendizado de máquina) para meus estudantes. Use a abordagem de (*backward design framework* / transferência de aprendizado / *problem based learning* / associe o roteiro com cada item da taxonomia de Bloom.)
- **Prompt Completo (que usa todas as teorias simultaneamente):** “Sou professor de (programação/ciência de dados), preciso de um plano de ensino de duas horas para ensinar (introdução à programação em *Python*/introdução ao aprendizado de máquina) para meus estudantes. Use transferência de aprendizado, *backward design framework*, a abordagem de *problem-based learning* e associe cada etapa da taxonomia de Bloom.”

Além dos planos de ensino e estudo gerados com esses *prompts*, foi gerado um plano “danificado”: manualmente, alguns trechos foram deixados incompletos, a ordem

de algumas etapas foi trocada e parte das instruções estão trocadas. O objetivo é confirmar que planos gerados assim são penalizados pela avaliação.

4. Resultados

Nesta Seção, apresentamos e discutimos os principais resultados experimentais do trabalho. As Figuras 3, 4, 5 e 6 exibem a nota média obtida em cada critério. A nota média foi obtida a partir do plano gerado pelo ChatGPT 2.0 e Gemini 2.0, e avaliado também por ambos. Dessa maneira, retiramos possíveis vieses que um modelo possa ter.

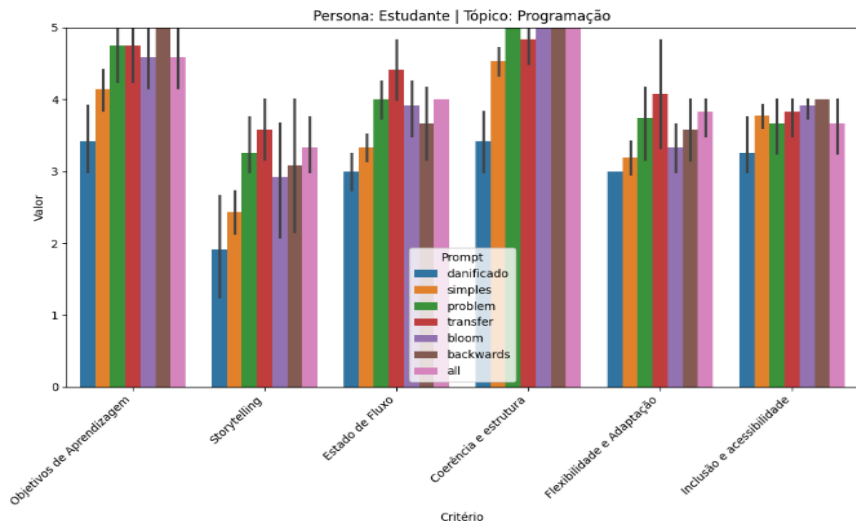


Figura 3: Avaliação Planos-Estudante Programação

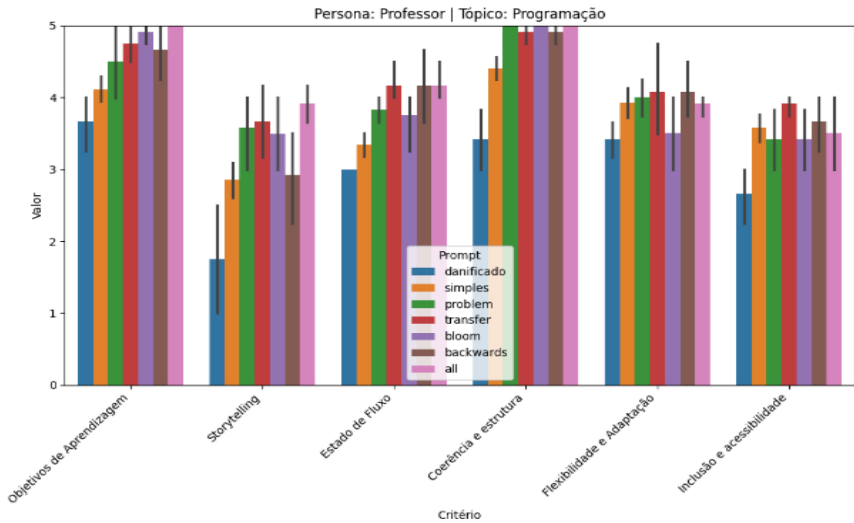


Figura 4: Avaliação Planos-Professor Programação

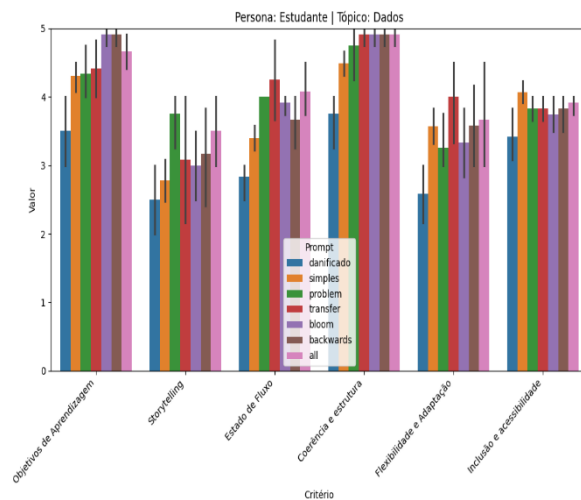


Figura 5: Avaliação Planos-Estudante Dados

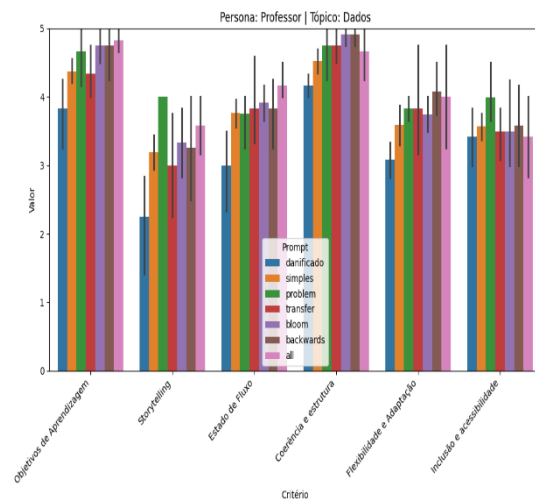


Figura 6: Avaliação Planos-Professor Dados

Os resultados das avaliações indicam que, no geral, os planos de ensino que incorporaram metodologias pedagógicas superaram os planos simples em termos de desempenho. Em todos os testes realizados, o plano de ensino danificado obteve notas inferiores ao plano simples, o que sinaliza a consistência do processo de avaliação. Os critérios “inclusão e acessibilidade” e “flexibilidade e adaptação” foram os critérios onde as notas do plano simples foram as mais próximas dos outros planos. Tanto em programação quanto em dados, independente da persona, todos os planos que mencionaram metodologias de ensino obtiveram, em média, notas mais altas que o *prompt* simples. A avaliação do critério de coerência e estrutura recebeu as melhores pontuações em ambos os tipos de *prompts*. Os planos gerados pelo Gemini 2.0 mostraram uma menor variação nas notas em comparação ao ChatGPT 4.0. Ambos os modelos de IA reconheceram que certos planos eram mais adequados para públicos específicos, o que favoreceu o uso de *prompts* simples em alguns casos. O Gemini 2.0, por sua vez, atribuiu notas equivalentes a *prompts* simples e específicos com maior frequência, sugerindo que ambos atendem a diferentes necessidades. Em resumo, os resultados indicam que os planos gerados com base em teorias de ensino obtiveram notas superiores ao simples, independentemente da *persona*, assunto ou IA como avaliadora utilizada.

5. Conclusões

As análises realizadas neste estudo destacam a relevância da incorporação de metodologias pedagógicas consagradas em inteligências artificiais voltadas para a educação. Os resultados indicam que os modelos ChatGPT 4.0 e Gemini 2.0 reconheceram que determinados planos de ensino e estudo, elaborados com base em teorias educacionais específicas, eram mais adequados para públicos específicos, como estudantes com maior proficiência em matemática. No entanto, em alguns cenários, os *prompts* mais simples, sem embutir teorias de aprendizagem, mostraram-se mais eficazes

para atender a um público mais amplo e heterogêneo.

Além disso, observamos que os planos gerados e avaliados pelo Gemini 2.0 apresentaram menor variância das pontuações atribuídas aos critérios avaliados em comparação aos do ChatGPT 4.0. Enquanto as notas dos planos de ensino e estudo, em média, subiram de 3,59 para 4,38 (Programação) e de 3,89 para 4,14 (Ciência de Dados) após a incorporação de teorias educacionais no ChatGPT 4.0, no Gemini 2.0, os escores passaram de 3,63 para 3,88 (Programação) e de 3,58 para 3,83 (Ciência de Dados). Para os planos de ensino e estudo categorizados como danificados, a média do resultado foi de 2,04 (Programação e Dados) quando avaliado pelo ChatGPT 4.0, enquanto pelo Gemini 2.0 foi de 1,84 (Programação) e 1,86 (Dados).

Este trabalho abre novas perspectivas para investigações futuras, incluindo a exploração de taxonomias alternativas à Taxonomia de Bloom, como *The New Taxonomy of Educational Objectives* (2007). A continuidade dessa pesquisa pode aprofundar ainda mais a compreensão sobre a interação entre IA e pedagogia. Conclui-se que a integração de diferentes abordagens teóricas e práticas é essencial para o desenvolvimento de soluções educacionais inovadoras e eficazes, impactando significativamente tanto a prática acadêmica quanto a profissional.

A partir das reflexões aqui apresentadas, espera-se que este estudo incentive novos questionamentos e fomenta um debate contínuo sobre o papel da inteligência artificial na educação, contribuindo para o aprimoramento de estratégias pedagógicas e metodológicas na era digital.

6. Referências

- Amalfitano, Domenico et al. (2023) *Artificial Intelligence Applied to Software Testing: A Tertiary Study*. *ACM Computing Surveys*, vol 56, Article No.: 58, p.1–38. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/3616372>
- Becker, Fernando. (2010) **O caminho da aprendizagem em Jean Piaget e Paulo Freire: Da ação à operação**. Petrópolis, RJ: Vozes, 2010.
- Bloom, Benjamin S. (1956) *Taxonomy of Educational Objectives*. Disponível em: https://eclass.uoa.gr/modules/document/file.php/PPP242/Benjamin%20S.%20Bloom%20%20Taxonomy%20of%20Educational%20Objectives%2C%20Handbook%201_%20Cognitive%20DomainAddison%20Wesley%20Publishing%20Company%20%281956%29.pdf
- Fiok, Krzysztof, Farzad V. Farahani, Waldemar Karwowski, and Tareq Z. Ahram. (2021). “Explainable Artificial Intelligence for Education and Training.” *Journal of Defense Modeling and Simulation* 19 (2): 133–144. <https://doi.org/10.1177/15485129211028651>.
- Gardner, H. (1993) *Multiple intelligences: The theory in practice*. Basic Books/Hachette Book Group.
- Lehmann, M., Conerlius, P. B., & Sting, F. J. (2024). AI Meets the Classroom: When Does ChatGPT Harm Learning? arXiv preprint arXiv:2409.09047.
- Kic-drzas, Joanna & Kılıçkaya, Ferit. (2024). *Using Artificial Intelligence (AI) to create language exam questions: A case study*. *XLinguae*. 17. 20-33.

10.18355/XL.2024.17.01.02.

- Kolodner, J. (2002). *Facilitating the learning of design practices: Lessons learned from an inquiry into science education*. *Journal of Industrial Teacher Education*, 39(3), 9–40
- Levy, Pérez Albertos, (2024) *Teaching Effectively with ChatGPT: A practical guide to creating better learning experiences for your students in less time*, Pag. 9.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence Unleashed. An argument for AI in education*. Pearson.
- Mardell et al. (2016) *Towards a Pedagogy of Play*. Disponível em: <https://pz.harvard.edu/sites/default/files/Towards%20a%20Pedagogy%20of%20Play.pdf>
- Mattalo, Brandon (2024) *Artificial Intelligence: The Future of Pedagogy*.
- Nguyen, Andy, Ha Ngan Ngo, Yvonne Hong, Belle Dang, and Bich-Phuong Thi Nguyen. (2022). “Ethical Principles for Artificial Intelligence in Education.” *Education and Information Technologies* 28 (4): 4221–4241. <https://doi.org/10.1007/s10639-022-11316-w>.
- Schmidt, H.G. (2012). A Brief History of Problem-based Learning. In: O'Grady, G., Yew, E., Goh, K.
- Steiner, G. (2001) *Transfer of Learning, Cognitive Psychology of International*. <https://www.sciencedirect.com/science/article/abs/pii/B0080430767014819>
- S. Z. Salas-Pilco, K. Xiao, X. Hu (2022) ***Artificial intelligence and learning analytics in teacher education: A systematic review***, *Education Sciences* 12 2022 569. doi:10.3390/educsci12080569. Disponível em: <https://www.mdpi.com/2227-7102/12/8/569>. Acesso em: set. 2024.
- Theobald, P. (2009) ***Education Now: How Re-thinking America's Past Can Change Its Future***. Colorado: Paradigm.
- Vanlehn, K. (2006). *The behavior of tutoring systems*. *International Journal of Artificial Intelligence in Education*, 16, 227–265
- Verga Pat et al. (2024) ***Replacing Judges with Juries: Evaluating LLM Generations with a Panel of Diverse Models***. Disponível em: <https://arxiv.org/abs/2404.18796>
- Woolf, Beverly Park. (2009). *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing e-Learning*. Burlington, MA: Morgan Kaufmann.
- Zawacki-Richter, OLAF, Victoria I. MARÍN, Melissa Bond, and Franziska Gouverneur. (2019). “Systematic Review of Research on Artificial Intelligence Applications in Higher Education—Where Are the Educators?” *International Journal of Educational Technology in Higher Education* 16:39. <https://doi.org/10.1186/s41239-019-0171-0>.