



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO

PRÓ-REITORIA DE GRADUAÇÃO

DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLOGIA DA INFORMAÇÃO

CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

ERIK DAVID DA SILVA LIMA

**Avaliação Da Experiência Do Usuário Em Chatbots Acadêmicos - Uma Análise
Comparativa Dos Métodos *Attrakdiff*, *Think Aloud* e *Max***

ANGICOS - RN

2025

ERIK DAVID DA SILVA LIMA

**Avaliação Da Experiência Do Usuário Em Chatbots Acadêmicos - Uma Análise
Comparativa Dos Métodos *Attrakdiff*, *Think Aloud* e *Max***

Trabalho de Conclusão de Curso
apresentado à Universidade Federal
Rural do Semi-Árido como requisito
para cumprimento do componente
curricular Trabalho de Conclusão de
Curso II.

Orientadora: Profa Dra Jarbele Cássia da
Silva Coutinho

ANGICOS - RN

2025

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

L732a Lima, Erik David da Silva.
Avaliação da experiência do usuário em chatbots acadêmicos: uma análise comparativa dos métodos AttrakDiff, Think Aloud e MAX / Erik David da Silva Lima. - 2025.
67 f. : il.

Orientadora: Jarbele Cássia da Silva Coutinho.
Monografia (graduação) - Universidade Federal Rural do Semi-árido, Curso de Sistemas de Informação, 2025.

1. Experiência do Usuário (UX). 2. Chatbots.
3. Avaliação de Usabilidade. 4. Triangulação Metodológica. 5. Métodos AttrakDiff, Think Aloud e MAX. I. Coutinho, Jarbele Cássia da Silva, orient. II. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade com AACR2 e os dados fornecidos pelo autor(a).
Biblioteca Campus Angicos
Bibliotecário: Helder Romero Maia Duarte
CRB: 15/673

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

ERIK DAVID DA SILVA LIMA

**Avaliação Da Experiência Do Usuário Em Chatbots Acadêmicos - Uma Análise
Comparativa Dos Métodos *Attrakdiff*, *Think Aloud* e *Max***

Trabalho de Conclusão de Curso
apresentado à Universidade Federal
Rural do Semi-Árido como requisito
para cumprimento do componente
curricular Trabalho de Conclusão de
Curso II.

Banca Examinadora:

Documento assinado digitalmente
 **JARBELE CASSIA DA SILVA COUTINHO**
Data: 19/12/2025 17:32:41-0300
Verifique em <https://validar.itl.gov.br>

Profª. Dra. Jarbele Cássia da Silva Coutinho - UFERSA
Presidente

Documento assinado digitalmente
 **JOAO BATISTA DE SOUZA NETO**
Data: 19/12/2025 21:26:41-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. João Batista de Souza Neto - UFERSA
Membro Examinador

Documento assinado digitalmente
 **BRUNO DE SOUSA MONTEIRO**
Data: 22/12/2025 15:25:36-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. Bruno de Sousa Monteiro - UFERSA
Membro Examinador

AGRADECIMENTOS

Agradeço, antes de tudo, a Deus que sempre guiou meus passos e me concedeu a resiliência e sabedoria necessária para superar todos os desafios e transformá-los em crescimento.

Aos meus pais, Edivan e Edineide, ao meus avós Edinalva, Geraldo e Benta, e as minhas irmãs Ellen e Emilly, minha eterna gratidão pelo apoio incondicional, pelos incentivos que me acompanharam em cada etapa da minha jornada e pelos valores que carrego junto a mim.

Aos meus amigos e colegas de curso, com quem compartilhei risadas, dúvidas e conquistas, deixando meu tempo de graduação mais leve e incrível. Por todos os momentos de colaboração e cumplicidade que fortaleceu bastante a nossa trajetória.

A minha orientadora Jarbele, onde sua paciência, expertise e dedicação foram fundamentais para a construção deste trabalho. Obrigado por todas as críticas construtivas, pelo olhar atento e por acreditar no potencial deste projeto, mesmo nos meus momentos de incerteza.

A Universidade Federal Rural do Semi-Árido (UFERSA), sendo essa a instituição que me acolheu e proporcionou não apenas conhecimento técnico, mas também as experiências que levarei comigo por toda minha vida.

Aos participantes Anderson Mateus, Ericathia Lopes, James Silva, Lizandra Ferreira, Sandro Felinto e Ywry Scheller. Sendo eles peças chaves para que esse experimento ocorresse da melhor forma.

Por fim, a todos aqueles que, de forma direta ou indireta, contribuíram para a minha formação acadêmica e pessoal. Cada conversa, conselho e gesto de apoio fez parte desta história, e é com muita humildade que reconheço: você ainda não consegue perceber o quão infinito você é.

RESUMO

Com o avanço das tecnologias de conversação, os *Chatbots* têm sido amplamente utilizados em diferentes contextos, incluindo o ambiente acadêmico. No entanto, a qualidade da experiência oferecida por esses sistemas ainda apresenta desafios, especialmente no que diz respeito à usabilidade e aceitação real pelos usuários. Este trabalho teve como objetivo avaliar a Experiência do Usuário (*User Experience* - UX) em *chatbots* acadêmicos, utilizando a triangulação de três técnicas específicas: *AttrakDiff*, *Think Aloud* e *MAX*. A pesquisa foi estruturada em três etapas: estudo bibliográfico, seleção e análise do *Chatbot* e realização de um experimento. O experimento contou com 6 participantes, que foram orientados a utilizar o chatbot *Study Skills Coach AI* e avaliá-lo por meio dos métodos *AttrakDiff*, *Think Aloud* e *MAX*. Os resultados revelaram uma dissonância entre a interface visual (UI) e a funcionalidade conversacional (UX). Enquanto os métodos *AttrakDiff* e *MAX* indicaram que a interface era considerada fácil de usar e visualmente agradável, o método *Think Aloud* diagnosticou que falhas técnicas (limite de caracteres) e respostas inadequadas geraram frustração e desejo de abandono pelo usuário. Conclui-se que a aplicação isolada de métodos quantitativos, como o *AttrakDiff* pode mascarar problemas graves de interação, tornando a triangulação com métodos qualitativos, como o *Think Aloud*, indispensável para o desenvolvimento de *Chatbots* eficazes.

Palavras-chave: Experiência do Usuário; *Chatbots*; Avaliação de Usabilidade.

ABSTRACT

With the advancement of conversational technologies, Chatbots have been widely used in different contexts, including the academic environment. However, the quality of the experience provide these systems still poses challenges, especially regarding usability and actual user acceptance. This work aimed to evaluate User Experience (UX) in academic Chatbots using a triangulation of three methods: *AttrakDiff*, *Think Aloud*, and *MAX*. The research was structured in three stages: bibliographic study, Chatbot selection and analysis, and the execution of an experiment. The experiment involved 6 participants who were instructed to use the Study Skills Coach AI Chatbot and evaluate it using the *AttrakDiff*, *Think Aloud*, and *MAX* methods. The results revealed a dissonance between the visual interface (UI) and the conversational functionality (UX). While the *AttrakDiff* and *MAX* methods indicated that the interface was considered easy to use and visually appealing, the *Think Aloud* method diagnosed that technical failures (character limit) and inadequate responses generated frustration and a desire to abandon usage. It is concluded that the isolated application of quantitative methods, such as *AttrakDiff*, can mask serious interaction problems, making triangulation with qualitative methods, such as *Think Aloud*, indispensable for the development of effective Chatbots.

Keywords: User Experience; Chatbots; Usability Evaluation.

LISTA DE FIGURAS

Figura 01 - Exemplo de aplicação do Tabuleiro <i>MAX</i>	8
Figura 02 - Fluxograma do Procedimento de Coleta de Dados.....	18
Figura 03 - Cartas do método <i>MAX</i> traduzidas e adaptadas.....	21
Figura 04 - Tabuleiro do método <i>MAX</i>	21
Figura 05 - Interface Padrão do <i>ScienceOS</i>	26
Figura 06 - Interface Padrão do <i>Study Skills Coach AI</i>	27
Figura 07 - Interface Padrão do <i>Perplexity</i>	28
Figura 08 - Exemplo de item do questionário <i>AttrakDiff</i>	31
Figura 09 - Modelo das Cartas <i>MAX</i> para Impressão.....	49
Figura 10 - Modelo das Cartas <i>MAX</i> para Impressão.....	50
Figura 11 - Seção 1: Qualidade Pragmática.....	51
Figura 12 - Seção 2: Qualidade Hedônica (Estimulação).....	52
Figura 13 - Seção 3: Qualidade Hedônica (Identidade).....	53
Figura 14 - Seção 4: Atratividade Global.....	54

LISTA DE TABELAS

Tabela 1: Aspectos metodológicos dos Trabalhos Relacionados discutidos no Capítulo 3 (Autor, 2025).....	14
Tabela 2: Resultados e perspectivas dos Trabalhos Relacionados discutidos no Capítulo 3 (Autor, 2025).....	15
Tabela 3: Perfil dos Participantes.....	18
Tabela 4: Tarefas do <i>Think Aloud</i>	20
Tabela 5: Itens do Questionário <i>AttrakDiff</i> Adaptado.....	22
Tabela 6: Comparação dos <i>Chatbots</i> selecionados (Autor, 2025).....	29
Tabela 7: Resultados das Médias do <i>AttrakDiff</i> (Autor, 2025).....	31
Tabela 8: Cartas mais frequentes no Método <i>MAX</i>	32

LISTA DE APÊNDICES

APÊNDICES.....	51
APÊNDICE A - TRABALHOS RESULTANTES DO ESTUDO BIBLIOGRÁFICO.....	51
APÊNDICE B - ROTEIRO DE TAREFAS PARA O <i>THINK ALOUD</i>	53
APÊNDICE C – ARTEFATOS ADAPTADOS DO MÉTODO <i>MAX</i> PARA IMPRESSÃO...	55
APÊNDICE D – QUESTIONÁRIO <i>ATTRAKDIFF</i> ADAPTADO.....	58

LISTA DE ABREVIATURAS E SIGLAS

AI	<i>Artificial Intelligence</i> (Inteligência Artificial)
IA	Inteligência Artificial
IHC	Interação Humano-Computador (disciplina/área de pesquisa)
ISO	<i>International Organization for Standardization</i> (Organização Internacional para Padronização)
LLMs	Grandes Modelos de Linguagem
MAX	<i>Method for the Assessment of eXperience</i> (Método de Avaliação da eXperiência)
PLN	Processamento de Linguagem Natural
RSL	Revisão Sistemática da Literatura
SUS	<i>System Usability Scale</i> (Escala de usabilidade do sistema)
TCC	Trabalho de Conclusão de Curso
UEQ	<i>User Experience Questionnaire</i> (Questionário de experiência do usuário)
UI	<i>User Interface</i> (Interface do Usuário)
UX	User Experience (Experiência do Usuário)

SUMÁRIO

RESUMO.....	6
ABSTRACT.....	6
1. INTRODUÇÃO.....	1
1.1 Problematização.....	1
1.2 Justificativa.....	2
1.3 Objetivos da pesquisa.....	3
1.4 Estrutura do trabalho.....	3
2. FUNDAMENTAÇÃO TEÓRICA.....	4
2.1 Experiência do Usuário (UX).....	4
2.2 Avaliação de Experiência do Usuário.....	5
2.2.1 Método Attrakdiff.....	5
2.2.2 Método Think Aloud.....	6
2.2.3 Método MAX.....	6
2.2.4 Outras Abordagens.....	8
2.3 Chatbots.....	9
3. TRABALHOS RELACIONADOS.....	11
3.1 Investigando Chatbots governamentais: um panorama sobre a Usabilidade dentro e fora do Brasil.....	11
3.2 Avaliação da Usabilidade de um Chatbot por meio das heurísticas de Nielsen adaptadas.....	11
3.3 Um estudo exploratório sobre métodos de avaliação de User Experience em Chatbots....	12
3.4 Human factors in the design of Chatbot interactions: conversational design practices.....	12
3.5 Ai-enhanced Chatbot for improving healthcare Usability and accessibility for older adults.....	13
3.6 Análise comparativa dos trabalhos.....	13
4. METODOLOGIA.....	17
4.1 Estudo bibliográfico.....	17
4.2 Busca e Seleção do Chatbot.....	17
4.3 Planejamento do Experimento.....	18
4.3.1 Método Think Aloud.....	19
4.3.2 Método MAX.....	20
4.3.3 Questionário AttrakDiff.....	22
4.4 Análise dos Dados.....	22
5. RESULTADOS.....	24
5.1 Resultado do estudo bibliográfico.....	24
5.2 Seleção de Chatbots.....	24
5.2.1 Chatbot 1 - ScienceOS.....	26
5.2.2 Chatbot 2 - Study Skills Coach Ai.....	27
5.2.3 Chatbot 3 - Perplexity (Modo Acadêmico).....	28
5.2.4 Análise comparativa dos Chatbots.....	29

5.3 Resultados do Experimento.....	30
5.3.1 Resultados do AttrakDiff.....	30
5.3.2 Resultados do MAX.....	32
5.3.3 Resultados do Think Aloud.....	33
6. DISCUSSÃO.....	34
6.1 Diferença entre a Interface e a Interação.....	34
6.2 A complementaridade dos métodos na avaliação de Chatbots.....	35
6.3 Análise Comparativa com Resultados da Literatura.....	36
6.4 Ameaças à Validade e Limitações do Estudo.....	37
7. CONCLUSÃO.....	39
REFERÊNCIAS.....	40
APÊNDICES.....	43
APÊNDICE A - TRABALHOS RESULTANTES DO ESTUDO BIBLIOGRÁFICO.....	43
APÊNDICE B - ROTEIRO DE TAREFAS PARA O THINK ALOUD.....	45
APÊNDICE C – ARTEFATOS ADAPTADOS DO MÉTODO MAX PARA IMPRESSÃO..	47
APÊNDICE D – QUESTIONÁRIO ATTRAKDIFF ADAPTADO.....	50

1. INTRODUÇÃO

O avanço das tecnologias de interação humano-máquina vem possibilitando o desenvolvimento de sistemas que estão cada vez mais presentes no cotidiano das pessoas, como os *Chatbots* e assistentes virtuais. Estes, por sua vez, têm ganhado destaque em várias áreas, como saúde, comércio eletrônico, serviços públicos e mais recentemente, no campo da educação e pesquisa acadêmica. Com a crescente na presença desses sistemas sendo usados em contextos informacionais e pedagógicos, tornou-se fundamental compreender como os usuários percebem e interagem com eles, especialmente no que se refere a Experiência do Usuário (UX). Em ambientes acadêmicos, onde clareza, funcionalidade e confiabilidade são essenciais, avaliar a experiência oferecida por *Chatbots* pode trazer *insights* valiosos para a melhoria de sua efetividade e aceitação.

Segundo a norma ISO 9241-210 (2019), a UX engloba aspectos subjetivos, como satisfação, prazer e motivação, e também critérios objetivos, como eficiência, usabilidade e desempenho. Norman e Nielsen (2020) reforçam que um bom *design* centrado no usuário deve ir além da funcionalidade básica, promovendo interações significativas e agradáveis. Logo, compreender como o usuário interpreta uma interface, reage às respostas do sistema e se sente ao utilizá-lo é essencial para garantir não apenas a adoção da tecnologia, mas também sua permanência na rotina diária dos usuários.

De acordo com Liu et al. (2025), Quando se trata de *Chatbots*, essa avaliação torna-se ainda mais complexa e relevante. Os *Chatbots* funcionam com base em linguagem natural e assumem uma função interativa que simula conversas humanas, o que exige não apenas precisão funcional, mas também empatia e coerência no diálogo. Em contextos acadêmicos, isso é de grande importância, já que os usuários esperam informações claras, apoio informativo e uma navegação sem obstáculos. Estudos como o de Følstad e Brandtzæg (2017) mostram que a confiança e o engajamento com assistentes conversacionais dependem diretamente da percepção de utilidade, da fluidez da conversa e da adequação ao contexto de uso. Avaliar essas interações sob a ótica da UX permite identificar falhas sutis, propor melhorias e orientar o desenvolvimento de sistemas mais eficazes.

1.1 Problemática

Apesar da popularização dos *Chatbots*, ainda existem desafios significativos quanto à qualidade da experiência oferecida ao usuário. Estudos como o de Rapp et al. (2021) apontam

que muitos *Chatbots* falham em oferecer interações empáticas, claras ou eficientes em sua função, o que impacta diretamente na aceitação dos usuários. No contexto acadêmico, essas limitações podem comprometer o propósito do *Chatbot* como ferramenta de apoio ao aprendizado, à pesquisa e à organização de estudos.

Desse modo, faz-se necessário entender como *Chatbots* são percebidos sob diferentes aspectos da interação, tais como a usabilidade percebida, a resposta emocional, a utilidade e a fluidez da conversa. A complexidade dessas dimensões sugere que métricas isoladas podem não ser suficientes para um diagnóstico completo. Com base nisso, foi definida a seguinte questão de pesquisa: "*De que maneira diferentes métodos de avaliação se complementam para diagnosticar a Experiência do Usuário(UX) em Chatbots acadêmicos?*"

1.2 Justificativa

A relevância deste estudo ancora-se na necessidade crítica de desenvolver *Chatbots* que transcendam a funcionalidade técnica básica e ofereçam interações fluidas e satisfatórias. Silva e Canedo (2022) destacam que, embora as ferramentas de desenvolvimento de *Chatbots* tenham evoluído e se popularizado, a interação com o usuário permanece um dos maiores desafios enfrentados pelos desenvolvedores. Decisões de *design* conversacional inadequadas podem impactar negativamente a percepção do usuário, independentemente da robustez tecnológica do sistema.

No contexto acadêmico, essa preocupação é amplificada. Ferramentas educacionais não exigem apenas precisão informacional, mas também a capacidade de engajar o estudante e transmitir confiança. Rapp et al. (2021) reforçam que fatores subjetivos, como empatia e percepção de utilidade, são decisivos para a aceitação dessas tecnologias. A simples implementação de um *Chatbot* não garante seu sucesso, é necessário assegurar que a "conversa" atenda às expectativas de um público que busca suporte para aprendizado e pesquisa.

Portanto, este trabalho se justifica pela aplicação de uma triangulação metodológica, abordagem defendida por Barbosa et al. (2022), que argumentam que a combinação de métodos qualitativos e quantitativos é fundamental para capturar a totalidade da experiência do usuário em *Chatbots*. Conforme Cavalcante et al. (2015), o uso de técnicas apropriadas revela falhas que passariam despercebidas em análises superficiais. Ao investigar a Experiência do Usuário (UX) sob diferentes óticas, como pragmática, emocional e cognitiva,

esta pesquisa busca contribuir não apenas para a melhoria técnica das ferramentas, mas para o aprimoramento do *design* conversacional, garantindo que *Chatbots* acadêmicos sejam efetivamente mais úteis e não apenas interfaces visualmente agradáveis, mas funcionalmente frustrantes.

1.3 Objetivos da pesquisa

O objetivo geral deste trabalho é avaliar a Experiência do Usuário (UX) com *Chatbots* acadêmicos, a fim de compreender as percepções cognitivas, funcionais e emocionais dos usuários durante a interação. Para isso, foram definidos os seguintes objetivos específicos de pesquisa:

- Identificar e selecionar *Chatbots* acadêmicos acessíveis e adequados para o estudo.
- Identificar e aplicar métodos de Avaliação de UX com usuários reais para avaliar a experiência de uso.
- Analisar comparativamente os resultados obtidos por cada técnica de UX, identificando suas contribuições e limitações para *Chatbots* acadêmicos.

1.4 Estrutura do trabalho

Este trabalho está organizado em sete capítulos. No Capítulo 1, apresenta-se a introdução, com a contextualização, problematização, justificativa e objetivos da pesquisa. O Capítulo 2 aborda a fundamentação teórica, apresentando os conceitos de Experiência do Usuário, os métodos de avaliação (*AttrakDiff*, *Think Aloud* e *MAX*) e o contexto dos *Chatbots*. O Capítulo 3 traz a revisão dos trabalhos relacionados, destacando metodologias e resultados de estudos anteriores. No Capítulo 4, é detalhada a metodologia adotada, que abrange desde o estudo bibliográfico e a seleção do chatbot objeto de estudo até os procedimentos de coleta de dados com os usuários. O Capítulo 5 apresenta os resultados obtidos na aplicação prática dos três métodos. O Capítulo 6 promove a discussão dos dados, realizando a triangulação entre as métricas quantitativas e os relatos qualitativos. Por fim, o Capítulo 7 apresenta a conclusão do estudo, retomando os objetivos, apontando as limitações e sugerindo trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais conceitos que fundamentam a pesquisa. Inicialmente, serão discutidos fundamentos relacionados à Experiência do Usuário (UX), mostrando sua definição e importância no desenvolvimento de sistemas, na Seção 2.1. Em seguida, serão descritos os métodos de avaliação da UX, como *Attrakdiff*, *Think Aloud* e *MAX*, destacando as suas características, aplicações e vantagens, na Seção 2.2. Por fim, é apresentado o conceito de *Chatbots*, com ênfase em sua evolução, uso e relação com a Experiência do Usuário, na Seção 2.3.

2.1 Experiência do Usuário (UX)

A Experiência do Usuário, ou UX (*User Experience*) se refere ao conjunto de percepções, emoções e reações que uma pessoa tem ao interagir com um determinado sistema ou produto. Segundo a ISO 9241-210 (2019), a UX resulta não apenas da interação direta, mas também das expectativas e do contexto de uso, sendo influenciada por fatores como usabilidade, estética e emoções associadas à experiência.

De acordo com Norman e Nielsen (2020), a Experiência do Usuário vai além da funcionalidade, abrangendo aspectos subjetivos, como sentimentos, valores e motivações. Logo, não basta que o sistema seja eficiente, ele deve fornecer uma interação positiva, intuitiva e também agradável, esses fatores impactam diretamente na aceitação e engajamento do usuário com a tecnologia.

No contexto de desenvolvimento de sistemas, adotar uma abordagem centrada no usuário é essencial para garantir a qualidade da experiência. Preece, Rogers e Sharp (2013) destacam que compreender as necessidades e limitações dos usuários, assim como observar seu comportamento em situações reais de uso, é fundamental para a criação de produtos mais eficientes, úteis e satisfatórios.

Por fim, em aplicações educacionais e acadêmicas, oferecer uma boa Experiência do Usuário (UX) é essencial para garantir o engajamento, compreensão e continuidade do uso. Desenvolver interfaces que promovam confiança, clareza e suporte ao usuário contribui significativamente para o sucesso de sistemas baseados em conversação, como *Chatbots*, especialmente quando inseridos em ambientes de estudo, pesquisa ou apoio pedagógico (RAPP et al., 2021).

2.2 Avaliação de Experiência do Usuário

A avaliação da Experiência do Usuário desempenha um papel fundamental no processo de desenvolvimento de sistemas, facilitando a identificação de pontos fortes e fracos a serem melhorados a partir do ponto de vista do usuário. Diversos métodos foram desenvolvidos com o intuito de coletar dados qualitativos e quantitativos sobre como as pessoas percebem e interagem com as tecnologias. Destacam-se, portanto, métodos como *Attrakdiff*, *Think Aloud* e *MAX*, descritos em mais detalhes a seguir.

2.2.1 Método Attrakdiff

O método *AttrakDiff* trata de um questionário para avaliar a qualidade pragmática e hedonística percebida em produtos interativos. Segundo Hassenzahl et al. (2003), a ideia do método é ir além de simplesmente medir a usabilidade, integrando os aspectos emocionais e subjetivos que influenciam a UX. O questionário combina pares de adjetivos opostos, que são apresentados em uma escala semântica diferencial de sete pontos, o que torna possível aos usuários indicar sua percepção em relação ao produto avaliado.

Qualidade pragmática se refere a capacidade de um produto apoiar o usuário na realização de suas tarefas de uma maneira mais eficiente e clara, ou seja, está relacionado a utilidade e facilidade de uso. Em contrapartida, a qualidade hedonística é associada a fatores como o estímulo, estética, originalidade e identidade, contribuindo para o prazer e satisfação durante o uso do produto (HASSENZAHN et al., 2003). Para Hassenzahl et al. (2003), esses aspectos estão ligados a necessidades humanas, como desenvolvimento pessoal e a auto expressão. Briva-Iglesias e O'Brien (2023) destacam que o *AttrakDiff* é especialmente eficaz para capturar essas dimensões subjetivas, oferecendo uma visão mais sensível sobre a percepção emocional do usuário, algo essencial em experiências baseadas em interação natural, como ocorre com sistemas inteligentes e conversacionais.

O método *AttrakDiff* foi proposto como uma alternativa às limitações de métodos mais tradicionais de usabilidade, focados em atributos objetivos de desempenho. Hassenzahl et al. (2003) argumentam que mesmo aqueles produtos tecnicamente eficientes podem falhar em atender as necessidades subjetivas dos usuários, o que compromete a aceitação e sucesso do mesmo. Assim, o questionário permite uma forma de análise mais holística da experiência, considerando tanto o desempenho funcional quanto os valores simbólicos associados ao uso do produto.

O *AttrakDiff* também se destaca por sua simplicidade de aplicação, podendo assim ser utilizado em diferentes estágios do desenvolvimento do produto, como nos protótipos ou até

sistemas finais. Nesse sentido, Briva-Iglesias e O'Brien (2023) ressaltam que, ao comparar o *AttrakDiff* com outros instrumentos como o *User Experience Questionnaire (UEQ)*, o *AttrakDiff* apresenta maior profundidade na análise hedônica, sendo mais adequado quando se pretende investigar aspectos subjetivos com mais nuance e fidelidade.

2.2.2 Método *Think Aloud*

O método *Think Aloud* é um método utilizado na engenharia cognitiva e no *design* de interfaces, considerado fundamental para a avaliação da Experiência do Usuário. De acordo com Jaspers et al. (2004), a abordagem consiste em pedir para os usuários verbalizarem, em tempo real, seus pensamentos enquanto realizam tarefas específicas com um sistema, fornecendo um entendimento direto dos processos cognitivos envolvidos.

A principal vantagem do método *Think Aloud* está na geração de dados verbais que refletem o raciocínio dos participantes no momento da tarefa. Segundo Charters (2003), esse método é eficaz para revelar os processos cognitivos durante a execução de atividades e permite que o pesquisador tenha acesso imediato aos pensamentos, motivações e estratégias utilizadas. Isso o torna relevante em contextos exploratórios e qualitativos, onde o objetivo é compreender o comportamento em profundidade.

Os resultados de um estudo desenvolvido por Jaspers et al. (2004) mostrou que a aplicação do *Think Aloud*, junto com a gravação do vídeo, permitiu compreender com mais profundidade as necessidades informacionais e os padrões de comportamento de oncologistas pediátricos ao interagirem com prontuários médicos. Neste caso, em específico, os dados coletados foram essenciais para o desenvolvimento de uma interface mais intuitiva, o que reduziu significativamente a carga cognitiva dos profissionais e melhorou a eficiência no preparo das consultas.

Portanto, quando aplicado corretamente, o *Think Aloud* é útil para detectar falhas precoces, orientar melhorias no *design* e apoiar a construção de sistemas mais alinhados às necessidades reais do usuário. Neste sentido, sua combinação com outras técnicas pode enriquecer ainda mais a avaliação de UX em contextos onde se trabalha com *Chatbots* e sistemas interativos.

2.2.3 Método *MAX*

O *Method for the Assessment of eXperience (MAX)* é um método criado para avaliar a Experiência do Usuário após o uso de sistemas interativos, utilizando um conjunto de cartas e

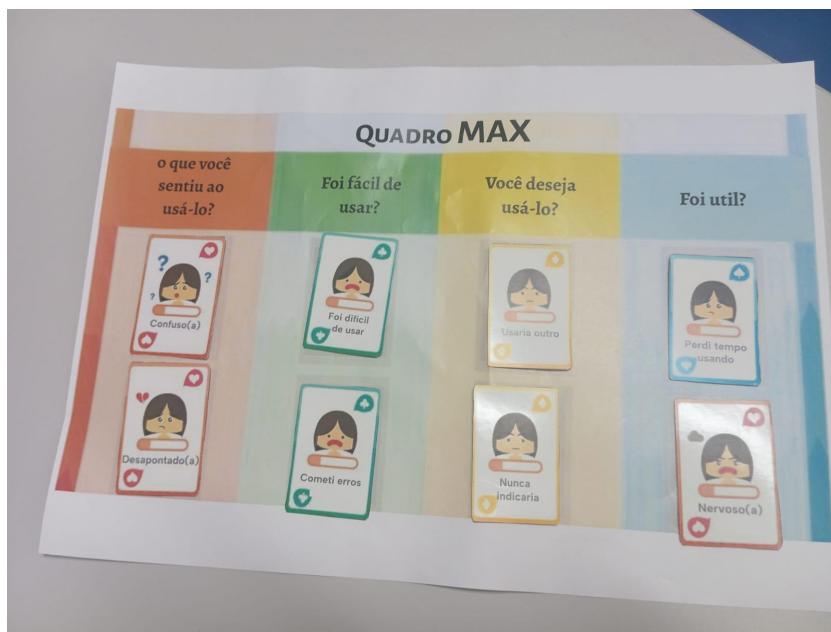
um tabuleiro. Segundo Cavalcante, Riveiro e Conte (2015), o objetivo principal do *MAX* é fazer o processo de avaliação mais confortável para os usuários, os fazendo relatar sua experiência de forma mais natural possível, sem se sentirem forçados ou constrangidos, como pode realmente acontecer em métodos tradicionais baseados em entrevistas diretas.

O método *MAX* utiliza um tabuleiro físico dividido em colunas que representam quatro dimensões centrais da experiência: emoção ("O que você sentiu?"), facilidade de uso ("Foi fácil de usar?"), utilidade ("Foi útil?") e intenção de uso ("Você deseja usá-lo?"). Para cada categoria, o participante dispõe de cartas ilustradas com avatares e escalas de intensidade, contendo frases curtas (ex: "Confuso", "Satisfeito", "Perdi tempo"). A dinâmica consiste em solicitar que o usuário selecione as cartas que melhor resumem sua interação e as posicione no tabuleiro, o que ajuda a materializar sentimentos abstratos. A Figura 01 ilustra esse processo, apresentando um exemplo prático de um tabuleiro preenchido durante uma sessão de avaliação. De acordo com Cavalcante et al. (2015), essa abordagem lúdica visa diminuir a resistência dos participantes, criando um ambiente mais descontraído e propício para coletar dados mais ricos e autênticos sobre a experiência vivida.

A aplicação do método *MAX* ocorre após a interação com o sistema, onde o usuário seleciona cartas que representam sua experiência emocional e pragmática. Segundo Barbosa et al. (2021), esse método se destaca por ter uma abordagem visual e intuitiva, permitindo que os participantes relatem seus sentimentos e percepções com um menor esforço cognitivo em comparação a questionários tradicionais.

Os resultados de um estudo piloto realizado por Cavalcante et al. (2015) indicaram que o *MAX* é uma ferramenta útil e de fácil aplicação, favorecendo tanto a obtenção de dados qualitativos sobre a UX quanto a participação ativa dos usuários. Além disso, o método sofreu melhorias em sua segunda versão, como a redução do número de cartas e ajustes no *design*, buscando torná-lo ainda mais acessível e eficaz para avaliações de experiência em diferentes contextos (CAVALCANTE et al., 2015).

Figura 01 - Exemplo de aplicação do Tabuleiro MAX



Fonte: Elaborado pelo autor (2025).

2.2.4 Outras Abordagens

A literatura da área de Interação Humano-Computador dispõe de diferentes métodos para avaliação de interfaces, como a avaliação heurística, e a técnica de inspeção por especialistas. Coelho et al. (2022) demonstraram sua eficácia em detectar inconsistências na interface de *chatbots* acadêmicos utilizando as heurísticas de Nielsen adaptadas. Entretanto, por depender do julgamento de especialistas, a avaliação heurística não captura diretamente a resposta emocional, às frustrações ou a satisfação subjetiva do usuário final durante a interação real.

No campo das métricas quantitativas, o *User Experience Questionnaire* (UEQ) se destaca como uma alternativa robusta ao método *AttrakDiff*. Briva-Iglesias e O'Brien (2023) compararam ambas as ferramentas, concluindo que o UEQ oferece escalas adicionais de eficiência, mas o *AttrakDiff* é superior na distinção clara entre qualidades pragmáticas e hedônicas. Essa clareza na separação entre a utilidade percebida (pragmática) e o estímulo emocional (hedônico) torna o *AttrakDiff* mais alinhado ao objetivo deste estudo de investigar o impacto emocional da interação.

Portanto, a escolha pela triangulação entre *AttrakDiff*, *Think Aloud* e *MAX* justifica-se pela necessidade de uma análise mais holística, conforme defendido por Barbosa et al. (2022). Enquanto o *AttrakDiff* mensura quantitativamente a qualidade percebida, o *Think Aloud* diagnostica as causas cognitivas das falhas e o Método *MAX* facilita a expressão de sentimentos negativos através de sua abordagem lúdica. Essa combinação permite cobrir as lacunas que o uso isolado de técnicas de inspeção ou questionários puramente métricos deixaria na avaliação de sistemas conversacionais.

2.3 Chatbots

Chatbots são sistemas computacionais capazes de interagir com seres humanos por meio de linguagem natural, escrita (geralmente), simulando uma conversa humana real. Segundo Rapp et al. (2021), nos últimos anos houve um crescimento significativo na utilização de *Chatbots* em diversas áreas, como no atendimento ao cliente, saúde, educação e entretenimento. Este fenômeno ficou conhecido como o “tsunami dos *Chatbots*”. Essa popularização se deve, principalmente, ao avanço das tecnologias de Inteligência Artificial (IA) e a facilidade de integração que esses sistemas possuem em plataformas digitais.

A interação entre humanos e *Chatbots* é influenciada por fatores que vão além da funcionalidade técnica. Segundo Adamopoulou e Moussiades (2020), a confiança do usuário depende de várias características do *chatbot*, como humanização (nomes humanos, linguagem natural) e profissionalismo, quanto do contexto de uso (privacidade, marca associada). Adamopoulou e Moussiades (2020) destacam que, embora os *Chatbots* ainda apresentem limitações em compreender as nuances emocionais, avanços na área de aprendizado de máquina têm permitido respostas mais empáticas, o que os aproxima das interações humanas.

Os *Chatbots* podem ser classificados de acordo com a sua finalidade principal, enquanto alguns são orientados para tarefas mais específicas, como fornecer informações ou executar ações, outros possuem um caráter mais social, buscando manter conversas e estabelecer vínculos com os usuários (RAPP et al., 2021). Essa é uma distinção fundamental, pois define o tipo de experiência proporcionada e influencia diretamente a expectativa e o nível de engajamento dos usuários com o sistema.

Adamopoulou e Moussiades (2020) destacam que o futuro dos *Chatbots* aponta para uma integração ainda mais invisível no cotidiano, onde a distinção entre o que é interação humana e artificial se tornará irrelevante; e ressaltam que, embora desafios como compreensão emocional persistam, o potencial de redução de custos em serviços e a escalabilidade mantêm os *Chatbots* como ferramentas estratégicas em negócios e saúde. Essa

evolução aliada a avanços com a IA generativa, mostra a necessidade de *designs* centrados não somente na eficiência, mas também em ética e adaptação contextual.

3. TRABALHOS RELACIONADOS

Este capítulo apresenta alguns trabalhos relacionados ao contexto desta pesquisa, advindos de um estudo bibliográfico sobre Usabilidade em *Chatbots*, Experiência do Usuário e *Design* de Interação. Desse modo, são discutidos os trabalhos de Batista, Monteiro e Salgado (2022), Coelho et al. (2022), Silva e Canedo (2024), e Santos et al. (2022).

3.1 Investigando Chatbots governamentais: um panorama sobre a Usabilidade dentro e fora do Brasil

Batista, Monteiro e Salgado (2022) analisam a Usabilidade de *Chatbots* governamentais no Brasil e em outros países (Argentina, Portugal, EUA, Reino Unido e Singapura), destacando os problemas que afetam a interação entre os cidadãos e os serviços públicos. A questão central da pesquisa foi: “*Quais são os potenciais impactos dos padrões de Usabilidade violados em Chatbots governamentais brasileiros e estrangeiros?*”. Tendo como objetivo avaliar a qualidade desses sistemas utilizando as heurísticas de Langevin et al. (2021) (sendo essa uma heurística adaptada de Nielsen). A metodologia incluiu a avaliação de dez *Chatbots* (sendo cinco brasileiros e cinco estrangeiros), por meio dos seguintes critérios: controle de usuário, prevenção de erros e confiabilidade. Os resultados da pesquisa mostraram que os *Chatbots* brasileiros apresentam uma quantidade significativa de problemas graves (como loops de conversa e falta de transparência com o usuário) em comparação com os estrangeiros. Por fim, a conclusão apontou a necessidade de melhorias no *design* e na funcionalidade para garantir uma maior eficiência e satisfação aos usuários, sugerindo a aplicação de avaliações contínuas de Usabilidade.

3.2 Avaliação da Usabilidade de um Chatbot por meio das heurísticas de Nielsen adaptadas

Coelho et al. (2022) avaliaram a Usabilidade do *chatbot InfoRussas*, que foi desenvolvido por acadêmicos da Universidade Federal do Ceará - Campus Russas, visando auxiliar os estudantes da instituição. O problema abordado foi a ausência de diretrizes específicas para se avaliar *Chatbots*. O objetivo foi aplicar as heurísticas adaptadas de Nielsen (2020) para identificar suas falhas e propor melhorias. A metodologia utilizada envolveu quatro inspetores, responsáveis por analisar o sistema para detectar violações como inconsistências na interface e dificuldades de navegação. Os resultados obtidos mostraram

que as heurísticas mais afetadas foram *visibilidade de status do sistema* e *reconhecimento ao invés de memorização*. Dentre as melhorias implementadas, foi incluída a adição de botões para facilitar a interação. Por fim, foi concluído que houve destaque na eficiência das heurísticas adaptadas, mas também foi ressaltado a necessidade de métodos mais específicos para os *Chatbots*.

3.3 Um estudo exploratório sobre métodos de avaliação de User Experience em Chatbots

Barbosa et al. (2022) investigaram a aplicabilidade dos métodos de avaliação de UX: *AttrakDiff*, *Think Aloud* e *MAX*, sendo eles utilizados no *chatbot* TeleCOVID. A questão central da pesquisa foi: *Como esses métodos capturaram a experiência do usuário em interações com Chatbots?*. Tendo como objetivo testar a viabilidade e aceitação dessas técnicas. A metodologia incluiu sete usuários que interagiram com o *chatbot* enquanto relataram as suas percepções sobre o processo. Os resultados do estudo indicaram que o *AttrakDiff* apresentou limitações em alguns processos; o *Think Aloud* foi o mais eficaz, mas exigiu maior esforço dos avaliadores; e, o *MAX* foi considerado o mais intuitivo de se realizar. Por fim, concluiu-se que a combinação dos três métodos proporciona uma avaliação mais completa da UX, recomendando adaptações para alguns contextos específicos.

3.4 Human factors in the design of Chatbot interactions: conversational design practices

Silva e Canedo (2024) investigaram fatores humanos no *design* de interação com *Chatbots*, focando nas práticas de *design* conversacional e o seu impacto na Experiência do Usuário. A questão de pesquisa investigou “*Como práticas textuais, visuais e interativas influenciaram percepções e sentimentos dos usuários, como satisfação, engajamento e confiança?*”. O objetivo principal foi propor e validar o *Guidelines for Chatbot Conversational Design (GCCD)*, um guia baseado em práticas extraídas de uma *Revisão Sistemática da Literatura (RSL)*, e validadas por meio de enquete e experimento. A metodologia baseou-se em uma combinação do uso de RSL, análise qualitativa e quantitativa, tendo participações de desenvolvedores com diferentes níveis de experiência. Os resultados indicaram que o guia é claro e funcional, sendo bastante útil e capaz de induzir a maior satisfação do engajamento dos usuários. Por fim, se concluiu que o GCCD pode ser uma ferramenta útil para os *designers* de *Chatbots*, permitindo interações mais naturais, transparentes e emocionalmente adequadas.

3.5 *Ai-enhanced Chatbot for improving healthcare Usability and accessibility for older adults*

Khamaj (2025) desenvolveu um *Chatbot* com foco em acessibilidade e usabilidade para adultos mais velhos em contexto de saúde. O objetivo do estudo foi explorar como um *Chatbot* com recursos de Inteligência Artificial (IA) pode melhorar a Experiência do Usuário para este público. A metodologia adotada combinou pesquisa aplicada com desenvolvimento experimental. Foram conduzidas entrevistas com 15 participantes idosos para compreender suas necessidades e dificuldades, seguidas de testes com um protótipo funcional do *Chatbot*. A avaliação da UX foi realizada por meio de questionários e análise qualitativa das interações, destacando aspectos como compreensão, engajamento, empatia e clareza na comunicação. Os resultados mostraram que o *Chatbot* contribuiu positivamente para a usabilidade e percepção hedônica, especialmente em relação a confiança e a sensação de apoio emocional. No entanto, Khamaj (2025) também apontou limitações, como a necessidade de adaptar a linguagem e a estrutura conversacional para diferentes níveis de alfabetização digital. A conclusão ressaltou o potencial de soluções baseadas em IA para ampliar a inclusão digital e recomendou o uso de abordagens centradas no usuário para o *design* de interfaces conversacionais.

3.6 Análise comparativa dos trabalhos

Os trabalhos apresentados e discutidos acima abordam diferentes focos no campo de avaliação de *Chatbots*, desde a análise de usabilidade em serviços públicos até a investigação de métodos específicos para mensurar a Experiência do Usuário. A Tabela 01 apresenta uma síntese dos aspectos metodológicos dos trabalhos analisados, destacando os objetivos de cada pesquisa, a questão investigada e a abordagem metodológica utilizada. A Tabela 02 complementa a análise ao sintetizar os principais resultados obtidos, os desafios ou limitações enfrentados por cada estudo e as sugestões de trabalhos futuros. Os estudos analisados possuem em comum a necessidade de validações mais amplas e específicas. Além disso, embora trabalhos como o de Barbosa et al. (2022) introduzem ferramentas para mensurar aspectos subjetivos e emocionais, observa-se no panorama geral uma oportunidade para consolidar e expandir a aplicação desses métodos em contextos específicos, como o de *Chatbots* acadêmicos, visando uma compreensão mais holística da UX.

Tabela 1: Aspectos metodológicos dos Trabalhos Relacionados discutidos no Capítulo 3
(Autor, 2025)

Trabalho	Objetivo da Pesquisa	Questão Investigada	Tipo de pesquisa (Metodologia)
Batista et al. (2022)	Avaliar a Usabilidade de <i>chatbots</i> governamentais nacionais e internacionais,	Quais são os potenciais impactos dos padrões de Usabilidade violados em <i>chatbots</i> governamentais?	Avaliação usando as heurísticas de Langevin (Adaptadas de Nielsen).
Coelho et al. (2022)	Avaliar o <i>chatbot</i> INFORussas para propor melhorias.	Como as heurísticas adaptadas de Nielsen identificam falhas na Usabilidade de <i>chatbots</i> ?	Avaliação usando as heurísticas de Nielsen
Barbosa et al. (2022)	Testar a viabilidade e aceitação de métodos de avaliação de UX em <i>chatbots</i> .	Como os métodos AttrakDiff, Think Aloud e MAX capturam a experiência do usuário com <i>chatbots</i> ?	Experimento com 7 usuários utilizando os três métodos de avaliação.
Silva e Canedo (2024)	Propor e validar um guia de práticas de <i>design</i> convencional de <i>chatbots</i> .	Como as práticas de <i>design</i> influenciam a satisfação, engajamento e confiança dos usuários?	Uso de revisão sistemática(RSL), Enquetes e experimentos.
Khamaj (2025)	Desenvolver um <i>Chatbot</i> com foco em acessibilidade e usabilidade para adultos mais velhos.	Como um <i>Chatbot</i> com IA pode melhorar a Experiência do Usuário (UX) para idosos no acesso à saúde?	Pesquisa aplicada com entrevistas, prototipação e avaliação empírica.
Este trabalho (2025)	Avaliar a UX em <i>chatbots</i> acadêmicos via triangulação de métodos (<i>AttrakDiff</i> , <i>Think Aloud</i> , <i>MAX</i>).	De que maneira diferentes métodos se complementam para diagnosticar a UX em <i>chatbots</i> acadêmicos?	Experimento com 6 participantes utilizando triangulação metodológica.

Fonte: Elaboração do autor (2025)

Tabela 2: Resultados e perspectivas dos Trabalhos Relacionados discutidos no Capítulo 3
(Autor, 2025)

Trabalho	Resultado Principal	Desafios/Limitações	Trabalhos Futuros
Batista et al. (2022)	<i>Chatbots</i> brasileiros apresentam mais problemas graves que estrangeiros.	Loops em conversas, falta de transparência pelo <i>chatbot</i> e ausência de avaliações contínuas.	Implementar avaliações periódicas, desenvolver padrões de <i>design</i> para os serviços públicos.
Coelho et al. (2022)	Falhas identificadas principalmente na visibilidade de status e reconhecimento ao invés de memorização.	Ausência de métodos específicos para <i>chatbots</i> .	Criar métodos próprios, comparar as heurísticas tradicionais e específicas.
Barbosa et al. (2022)	Think Aloud foi o mais eficaz, a combinação dos métodos é mais completa.	Validação inicial limitada, necessidade de aplicação em outros cenários.	Explorar os métodos menos custosos e ampliar as amostras.
Silva e Canedo (2024)	O guia proposto promove maior satisfação e engajamento.	Validação inicial limitada, necessidade de aplicação em outros cenários,	Expandir as validações e investigar impactos longitudinais.
Khamaj (2025)	O <i>Chatbot</i> melhorou a usabilidade e confiança entre usuários idosos.	Adaptação da linguagem e alfabetização digital foram desafios.	Investigar variações de linguagem e expandir o público alvo.
Este trabalho (2025)	Identificação de dissonância entre UI e UX; O método <i>Think Aloud</i> diagnosticou falhas (bugs) mascaradas pelos métodos quantitativos.	Amostra reduzida (6 participantes), barreira do idioma inglês.	Replicar a triangulação em modelos de linguagem mais robustos e investigar métricas de empatia.

Fonte: Elaboração do autor (2025)

Diferentemente dos trabalhos apresentados, este trabalho adota uma abordagem exploratória com foco na avaliação prática da UX em *Chatbots* acadêmicos, utilizando simultaneamente os métodos *Think Aloud*, *AttrakDiff* e *MAX*. A proposta busca compreender como diferentes dimensões da experiência são percebidas em interações reais, considerando tanto os aspectos funcionais quanto os subjetivos e emocionais.

4. METODOLOGIA

Neste capítulo, serão apresentados os procedimentos metodológicos a serem adotados para o desenvolvimento da pesquisa. O estudo foi estruturado em três etapas principais, sendo elas: o levantamento bibliográfico, apresentado na Seção 4.1; busca e análise de *chatbots* gratuitos e voltados ao suporte acadêmico, apresentado na Seção 4.2; e, a realização de um experimento com usuários, por meio de técnicas de avaliação de Experiência do Usuário.

4.1 Estudo bibliográfico

A primeira etapa consiste na realização de um estudo bibliográfico para compreender os conceitos da área de conhecimento relacionada à Experiência do Usuário (UX), e identificar métodos de avaliação aplicáveis a sistemas interativos, com o foco em *Chatbots*. Assim, foram utilizados como base de dados para pesquisa, as seguintes fontes de busca: *ACM Digital Library*¹, *ScienceDirect*² e o Google Acadêmico³. Para a busca foram usadas as seguintes palavras-chave: “*user experience*”, “*UX*”, “*Chatbots*”, “*Usabilidade*”, “*IA*”. Foram priorizados trabalhos publicados no período compreendido entre os anos de 2013 a 2025, no contexto de aplicação prática de técnicas de avaliação de *Chatbots*.

4.2 Busca e Seleção do Chatbot

A segunda etapa envolveu a identificação e seleção de um *Chatbot* acadêmico para ser usado no experimento. Foi realizada uma busca exploratória entre junho e julho de 2025, resultando em uma lista inicial de 8 *Chatbots*. Após a aplicação dos critérios de inclusão definidos para a análise geral dos *Chatbots encontrados* (disponibilidade pública gratuita, interação textual e interface funcional), três *Chatbots* foram previamente selecionados: *ScienceOS*⁴, *Study Skills Coach AI*⁵ e *Perplexity*⁶.

Em seguida, foram aplicados critérios mais específicos de análise, como clareza, empatia e adequação aos métodos de UX, cujos resultados detalhados estão apresentados na Seção 5.2. Por fim, foi definido como objeto de estudo do Experimento o *Chatbot Study Skills Coach AI*. A escolha justifica-se pelo seu modelo de interação orientado a aconselhamento

¹ ACM Digital Library: <https://dl.acm.org/>

² ScienceDirect: <https://www.sciencedirect.com/>

³ Google Acadêmico: <https://scholar.google.com/?hl=pt-BR>

⁴ ScienceOS: <https://www.scienceos.ai/science-ai-chat>

⁵ Study Skills Coach AI: <https://onlinechatbot.ai/chatbots/study-skills-coach/>

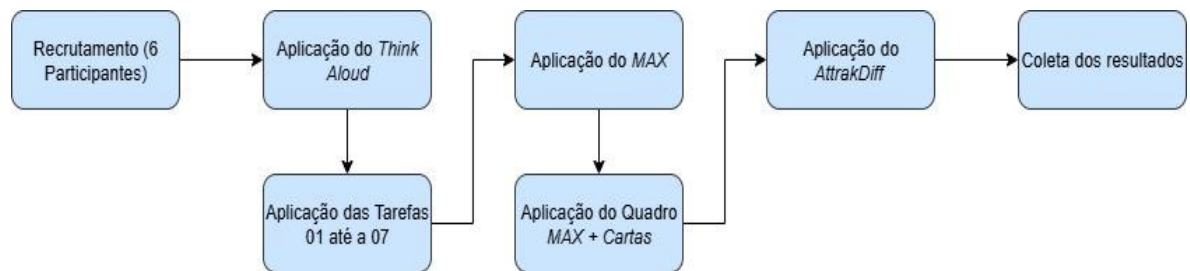
⁶ Perplexity: <https://www.perplexity.ai/>

acadêmico, que prioriza um diálogo personalizado e motivacional, ideal para estimular as reações emocionais e cognitivas necessárias para a aplicação dos métodos *MAX* e *AttrakDiff*.

4.3 Planejamento do Experimento

Para garantir a consistência na aplicação dos métodos e a integridade da triangulação proposta, o experimento seguiu um protocolo rígido de atividades. A Figura 02 ilustra o fluxo sequencial percorrido individualmente com cada usuário, dividindo o procedimento em três fases distintas, sendo elas: (1) Preparação, que envolveu o recrutamento e a garantia dos aspectos éticos; (2) Execução, onde os métodos *Think Aloud*, *MAX* e *AttrakDiff* foram aplicados nesta ordem específica; e (3) Finalização, com o encerramento da sessão.

Figura 02 - Fluxograma do Procedimento de Coleta de Dados



Fonte: Elaborado pelo autor (2025).

Esse protocolo foi aplicado com 6 participantes voluntários, todos estudantes de graduação da Universidade Federal Rural do Semi-Árido (UFERSA), matriculados nos cursos de Bacharelado em Sistemas de Informação (BSI) e Ciências e Tecnologia (C&T). A seleção da amostra buscou garantir diversidade em termos de familiaridade prévia com *chatbots* e variabilidade no nível de proficiência em língua inglesa, visto que a ferramenta objeto de estudo opera exclusivamente neste idioma. O detalhamento do perfil demográfico e técnico desses participantes é apresentado na Tabela 3.

Tabela 3: Perfil dos Participantes

Participante	Curso	Idade	Nível de Inglês (Autoavaliado)	Já utilizou algum Chatbot previamente?
Participante 1 (P1)	BSI	23	Mediano	Sim
Participante 2 (P2)	BSI	21	Baixo	Sim

Participante 3 (P3)	C&T	21	Mediano	Não
Participante 4 (P4)	C&T	22	Baixo	Sim
Participante 5 (P5)	C&T	20	Mediano	Sim
Participante 6 (P6)	C&T	26	Baixo	Sim

Fonte: Elaboração do autor (2025)

As avaliações foram planejadas para ocorrerem em sessões individuais com cada participante, seguindo uma ordem de aplicação desenhada para capturar a experiência desde a interação com o *Chatbot* (na qual o participante detinha autonomia para controlar o ritmo e a profundidade de exploração antes de avançar para a próxima etapa), até a avaliação estruturada. A triangulação metodológica ocorreu na seguinte sequência: primeiramente, aplicou-se o método *Think Aloud* para registrar os processos cognitivos em tempo real, utilizando o roteiro de tarefas detalhado no Apêndice B. Imediatamente após a interação, foi aplicado o método *MAX* para capturar a percepção emocional imediata, com o auxílio das cartas e do tabuleiro apresentados no Apêndice C. Por fim, a sessão foi concluída com a aplicação do questionário *AttrakDiff* para mensuração quantitativa das dimensões de qualidade, conforme estrutura disponível no Apêndice D.

É importante ressaltar que, embora a maioria dos voluntários (5 de 6) já possuísse experiência prévia com o uso de *chatbots* em geral, nenhum dos participantes teve contato anterior com a ferramenta específica selecionada para este estudo (*Study Skills Coach AI*). Essa condição foi assegurada para garantir a isenção na avaliação das primeiras impressões e a autenticidade das reações durante a fase de exploração inicial.

4.3.1 Método *Think Aloud*

Para viabilizar a coleta de dados durante a interação, os participantes receberam um roteiro estruturado contendo 7 tarefas (detalhadas na Tabela 4). A Tarefa 01 foi desenhada especificamente para motivar a exploração inicial livre da interface, permitindo que o usuário se familiarizasse com o *chatbot* antes de prosseguir para comandos mais complexos.

Os participantes foram instruídos a seguir o protocolo do método *Think Aloud*, verbalizando seus pensamentos, dúvidas e sentimentos em tempo real enquanto executavam cada uma das tarefas propostas. É importante ressaltar que este roteiro serviu como o condutor

da experiência: foi através da execução destas tarefas que os usuários geraram as percepções que seriam avaliadas em tempo real (pelo *Think Aloud*) e posteriormente consolidadas (pelos métodos *MAX* e *AttrakDiff*). As sessões foram gravadas em áudio para transcrição e análise.

Tabela 4: Tarefas do Think Aloud

Tarefa	Descrição	Objetivo
Tarefa 01	Observar a tela inicial e descrever impressões.	Avaliar a UI e a expectativa inicial.
Tarefa 02	Pedir dicas para organizar horários de estudo.	Testar a utilidade prática inicial.
Tarefa 03	Pedir uma técnica para memorizar conteúdos.	Avaliar clareza da instrução.
Tarefa 04	Pedir ajuda para um plano de estudos na rotina.	Testar personalização e empatia.
Tarefa 05	Pedir dicas para melhorar a concentração.	Verificar a relevância das respostas.
Tarefa 06	Pedir revisão de conteúdo técnico (Por exemplo: Algoritmos e programação 1).	Testar limites do conhecimento do bot.
Tarefa 07	Pedir um resumo das dicas da conversa.	Avaliar capacidade de síntese.

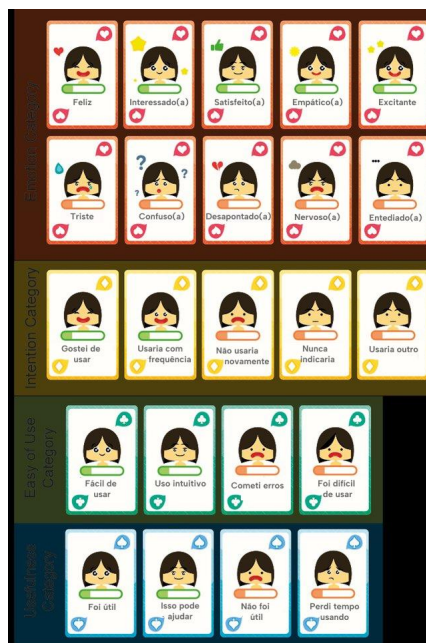
Fonte: Elaboração do autor (2025)

4.3.2 Método *MAX*

Imediatamente após a interação com o *Chatbot*, foi apresentado o Quadro *MAX* (devidamente traduzido para português) e um conjunto de cartas. Para facilitar a expressão dos sentimentos, foram utilizadas cartas traduzidas e adaptadas para o contexto da pesquisa, conforme ilustrado na Figura 03. A versão completa das cartas e do tabuleiro para reutilização em pesquisas futuras encontra-se disponível no Apêndice C. Além disso, disponibilizamos o endereço web onde os arquivos originais foram obtidos⁷.

⁷ Arquivos digitais do Método *MAX* : https://www.researchgate.net/publication/294580750_MAX_A_Method_for_Evaluating_the_Post-use_User_eXperience_through_Cards_and_a_Board

Figura 03 - Cartas do método *MAX* traduzidas e adaptadas



Fonte: Elaborado pelo autor (2025).

Além das cartas, os participantes utilizaram um tabuleiro físico, impresso em papel, para organizar suas escolhas presencialmente, permitindo uma visualização clara e tangível da experiência, conforme apresentado na Figura 04. Vale ressaltar que, embora o método original e esta aplicação tenham utilizado artefatos físicos, trabalhos correlatos como o de Barbosa et al. (2022) já demonstraram a viabilidade de adaptação do MAX para ambientes virtuais utilizando plataformas de colaboração online (como o Miro), o que não foi o caso nesta pesquisa.

Figura 04 - Tabuleiro do método *MAX*



Fonte: Elaborado pelo autor (2025).

4.3.3 Questionário *AttrakDiff*

Por fim, os participantes responderam ao questionário *AttrakDiff* via um formulário digital desenvolvido no *Google Forms*⁸. Foram utilizados pares de adjetivos opostos para mensurar quantitativamente as quatro dimensões de qualidade percebida: Qualidade Pragmática (PQ), Qualidade Hedônica de Estimulação (QH-E), Qualidade Hedônica de Identidade (QH-I) e Atratividade Global (ATT). A estrutura completa dos itens avaliados é apresentada na Tabela 5.

Tabela 5: Itens do Questionário *AttrakDiff* Adaptado

Dimensão	Pares de Adjetivos (Escala de 1 a 7)
Qualidade Pragmática (QP)	Difícil de usar – Fácil de usar; Confuso – Claro; Desorganizado – Organizado; Ineficiente – Eficiente; Complicado – Simples.
Qualidade Hedônica – Estimulação (QH-E)	Monótono – Estimulante; Sem graça – Interessante; Comum – Diferente; Previsível – Surpreendente; Pouco criativo – Criativo.
Qualidade Hedônica – Identidade (QH-I)	Distante – Acolhedor; Pouco confiável – Confiável; Frio – Próximo; Sem prestígio – Respeitável; Pouco profissional – Profissional.
Atratividade Global (ATT)	Feio – Bonito; Desagradável – Agradável; Ruim – Bom; Pouco atrativo – Atrativo; Desconfortável – Confortável.

Fonte: Elaborado pelo autor (2025).

4.4 Análise dos Dados

A análise seguiu uma abordagem de triangulação. Os dados quantitativos do *AttrakDiff* (médias) e do *MAX* (frequência de cartas) foram cruzados com os dados qualitativos das transcrições do *Think Aloud*. Essa estratégia permitiu não apenas identificar quais problemas ocorreram, mas também o motivo e como impactam a experiência do usuário.

A escolha e a combinação do método *Think Aloud* juntamente com os métodos *AttrakDiff* e *MAX* foram fundamentadas no estudo de Barbosa et al. (2022), que investigou a

⁸ *Google Forms*: <https://forms.gle/XgkEEfTZ7N3YxUWv9>

aplicabilidade, viabilidade e aceitação dessas três abordagens na avaliação da Experiência do Usuário em *Chatbots*. O estudo demonstrou que a utilização conjunta dessas técnicas e métodos é capaz de capturar, de forma complementar, diferentes aspectos da UX, desde percepções emocionais e cognitivas até elementos de atratividade e usabilidade. Assim, a adoção desse conjunto metodológico neste trabalho se justifica tanto pela sua robustez quanto pela sua adequação ao contexto de avaliação remota com *chatbots* de perfil acadêmico. Por fim, visando a transparência científica e a possibilidade de replicação deste estudo, todos os materiais utilizados e os dados anonimizados foram disponibilizados em um repositório digital público⁹.

⁹ Repositório de Dados do
<https://drive.google.com/drive/u/6/folders/1PFDtwqdz7sMLKiz5TwYXY4VWGiC6h75>

5. RESULTADOS

Este capítulo apresenta os resultados obtidos a partir da busca e análise feita sobre *Chatbots*, bem como os dados coletados durante a execução do estudo de caso. Inicialmente, a Seção 5.1 detalha os achados do levantamento bibliográfico. Na sequência, a Seção 5.2 descreve o processo de triagem e seleção das ferramentas, apresentando a análise comparativa que fundamentou a escolha do objeto de estudo. Por fim, a Seção 5.3 expõe os dados empíricos coletados junto aos participantes, organizados conforme os resultados específicos de cada método aplicado (*AttrakDiff*, *MAX* e *Think Aloud*).

5.1 Resultado do estudo bibliográfico

A busca por trabalhos relacionados foi direcionada à identificação de estudos sobre Experiência do Usuário (UX), métodos de avaliação e a aplicação de *Chatbots* em contextos digitais. Foram estabelecidos critérios de inclusão, como a presença de fundamentação teórica sólida, exemplos de aplicação prática (estudos de caso ou experimentos) e publicações compreendidas no período de 2013 a 2025.

Após a triagem inicial baseada na leitura de títulos e resumos, foram identificados 10 estudos que atendiam aos critérios de relevância, os quais foram catalogados no Apêndice A. Deste conjunto, foi realizada uma leitura completa e crítica, resultando na seleção de 5 trabalhos para a análise aprofundada apresentada no Capítulo 3.

Os demais estudos foram excluídos desta etapa final com base nos seguintes critérios: (1) foco estritamente técnico/arquitetural, abordando a construção do *software* sem ênfase na interação com o usuário; (2) ausência de validação empírica, tratando-se de propostas teóricas sem testes com usuários reais; ou (3) obsolescência do modelo de interação, referindo-se a *Chatbots* baseados em comandos rígidos que não refletem a fluidez das interfaces conversacionais atuais baseadas em Processamento de Linguagem Natural (PLN).

A síntese dos dados extraídos, incluindo objetivos e metodologias, serviu de base para o desenvolvimento da Fundamentação Teórica (Capítulo 2) e permitiu mapear as lacunas e oportunidades de pesquisa discutidas nos Trabalhos Relacionados (Capítulo 3), justificando a escolha pela triangulação metodológica adotada nesta pesquisa.

5.2 Seleção de Chatbots

Foi conduzida uma busca exploratória por *chatbots* disponíveis gratuitamente na web, com foco em aplicações acadêmicas. A busca foi realizada entre os meses de junho e julho de

2025, por meio de pesquisas em mecanismos como Google¹⁰, utilizando as seguintes palavras-chave: “*Chatbot para estudantes*”, “*Chatbot para estudantes*”, “*Chatbot acadêmico gratuito*” e “*AI study assistant*”.

No total, foram selecionados oito *Chatbots* que atendiam parcialmente os critérios definidos na Seção 4.2, foram eles: *ScienceOS*¹¹, *Study Skills Coach AI*¹², *Perplexity*¹³, *Scite*¹⁴, *Consensus*¹⁵, *Jenni AI*¹⁶, *Merlin AI*¹⁷ e *Virtual Writing Tutor*¹⁸. A triagem consistiu em interações simuladas com cada sistema, observando aspectos como acessibilidade, uso de linguagem natural, disponibilidade sem custo, funcionalidade de interface e compatibilidade com métodos e técnicas de avaliação da UX.

Dentre as opções, três *chatbots* foram selecionados para a análise aprofundada por apresentarem maior aderência aos critérios e representarem diferentes espectros de uso acadêmico:

- ***ScienceOS***: Atua no contexto de busca científica, especializado em artigos e referências.
- ***Study Skills Coach AI***: Atua no contexto de apoio ao aprendizado, focando em organização e técnicas de estudo.
- ***Perplexity (Modo Acadêmico)***: Atua no contexto de pesquisa informacional, servindo como um motor de respostas gerais com fontes citadas.

Essa diversidade possibilitou uma avaliação mais ampla da Experiência do Usuário (UX) em diferentes cenários de interação voltados à educação.

¹⁰ Google: <https://www.google.com/>

¹¹ *ScienceOS*: <https://www.scienceos.ai/science-ai-chat>

¹² *Study Skills Coach AI*: <https://onlinechatbot.ai/chatbots/study-skills-coach/>

¹³ *Perplexity*: <https://www.perplexity.ai/>

¹⁴ *Scite*: <https://scite.ai/>

¹⁵ *Consensus*: <https://consensus.app/>

¹⁶ *Jenni AI*: <https://jenni.ai/>

¹⁷ *Merlin AI*: <https://www.getmerlin.in>

¹⁸ *Virtual Writing Tutor*: <https://virtualwritingtutor.com/>

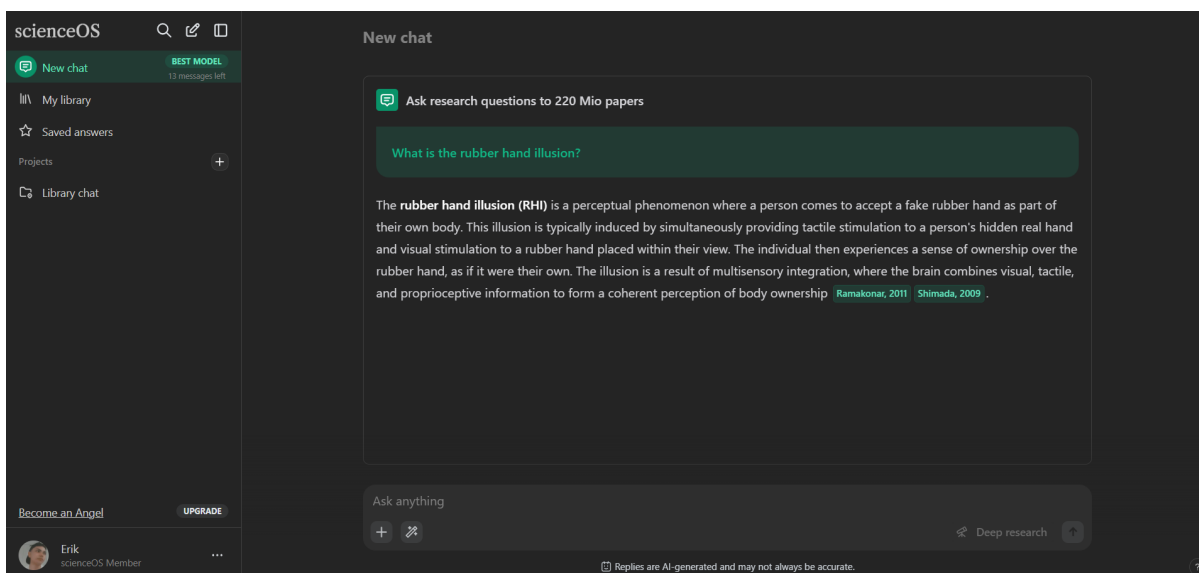
5.2.1 Chatbot 1 - *ScienceOS*

O *ScienceOS*¹⁹ é um *chatbot* gratuito focado em pesquisas científicas. Ele permite aos usuários realizar buscas com linguagem textual em um banco de dados que possui mais de 200 milhões de artigos científicos, além de conseguir gerar diagramas de relação, citações automáticas e redes de coautoria. Possui uma interface simples e direta.

Seu uso é mais direcionado para estudantes, pesquisadores e profissionais que precisam consultar literatura científica de forma rápida e estruturada. Ele se destaca por traduzir comandos simples fornecidos pelo usuário em resultados organizados e práticos, o que o torna uma ferramenta funcional e atraente.

A fluidez nas respostas e a capacidade de sintetizar os conteúdos faz com que o *ScienceOS* seja uma opção viável para a análise da UX. A ferramenta permite acompanhar a linha de raciocínio do usuário, avaliar facilmente a atratividade e funcionalidade do *chatbot*. As informações aqui apresentadas foram obtidas por meio de observações práticas realizadas durante o uso do *chatbot*, conforme os critérios estabelecidos na metodologia da pesquisa (consultar a Seção 4.2).

Figura 05 - Interface Padrão do *ScienceOS*



Fonte: Elaboração do autor (2025)

¹⁹ *ScienceOS*: <https://www.scienceos.ai/science-ai-chat>

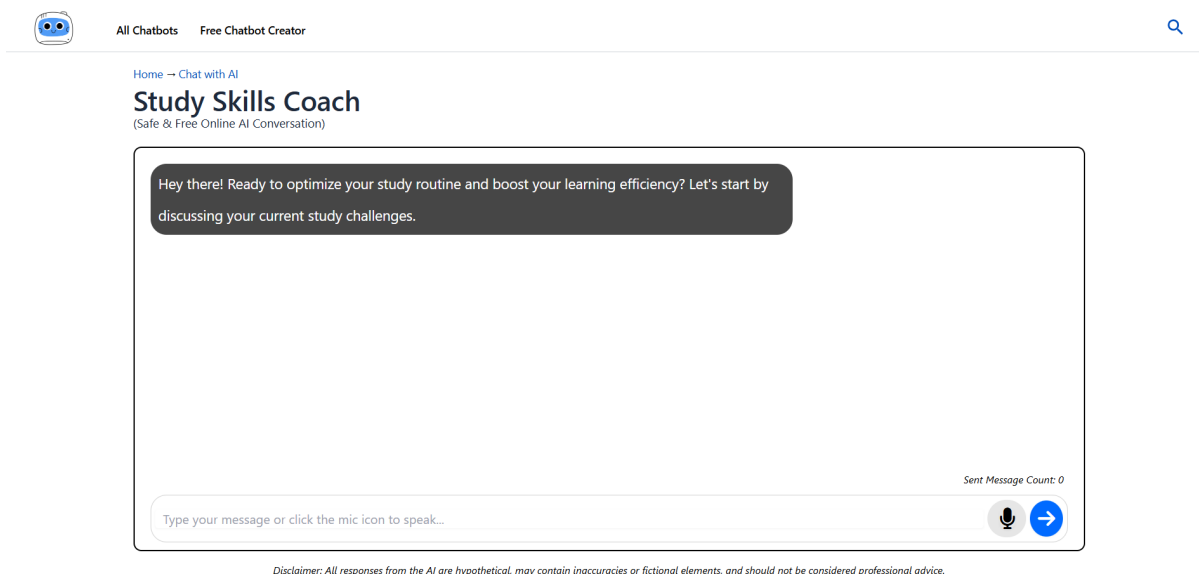
5.2.2 Chatbot 2 - *Study Skills Coach AI*

O *Study Skills Coach AI*²⁰ é um *chatbot* focado no desenvolvimento de habilidades acadêmicas, especialmente relacionadas à organização pessoal, cronograma de estudos, estratégias de estudo, técnicas de memorização e redação. A proposta da ferramenta é oferecer orientações que sejam práticas para estudantes em busca de melhorar seu desempenho acadêmico, fornecendo uma experiência mais interativa com dicas e sugestões em tempo real.

Diferente de um tutor tradicional, esse *chatbot* atua mais como um conselheiro digital, se adaptando às necessidades do usuário, usando respostas contextuais e orientações simples. A linguagem usada é amigável e direta, o que torna seu uso agradável e acessível. A estrutura da interação também torna favorável a criação de um caminho lógico de aprendizado, com foco em gestão e autonomia pessoal de estudos.

A aplicação desse *chatbot* pode permitir observar como o usuário compreende, assimila e aplica sugestões práticas ao longo da conversa. A clareza das instruções e artefatos gerados pela interação são elementos ideais para uma análise de UX. Além disso, é um *chatbot* acessível, funcionando diretamente em navegadores, o que facilita a aplicação das técnicas propostas. As informações aqui apresentadas foram obtidas por meio de observações práticas realizadas durante o uso do *chatbot*, conforme os critérios estabelecidos na metodologia da pesquisa (Seção 4.2).

Figura 06 - Interface Padrão do *Study Skills Coach AI*



Fonte: Elaboração do autor (2025)

²⁰ *Study Skills Coach AI*: <https://onlinechatbot.ai/chatbots/study-skills-coach/>

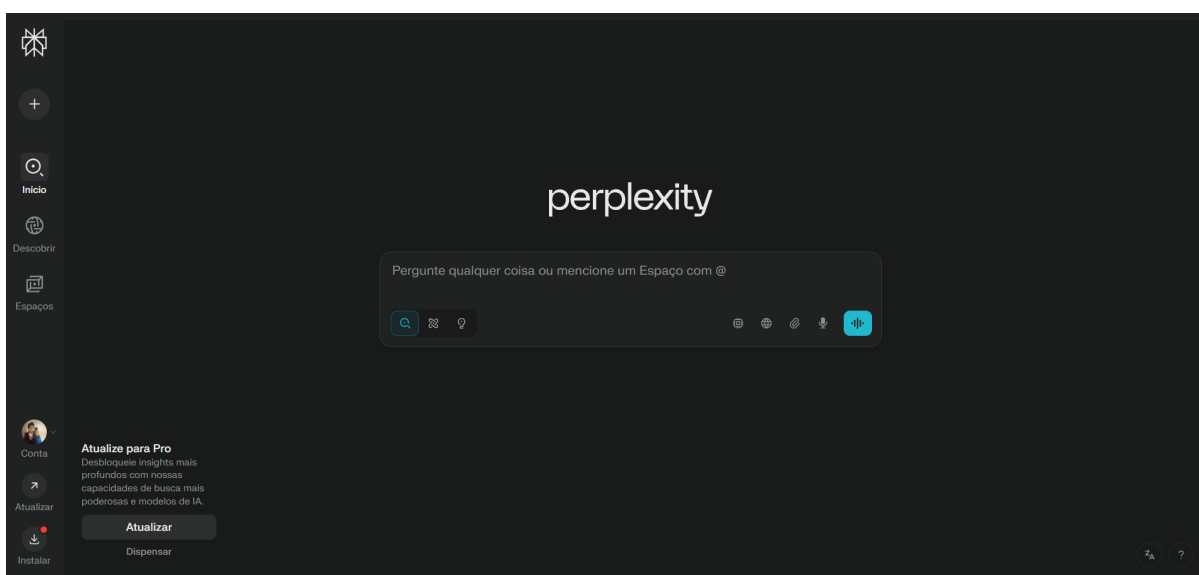
5.2.3 Chatbot 3 - *Perplexity* (Modo Acadêmico)

O *Perplexity*²¹ é uma plataforma conversacional baseada em Inteligência Artificial (IA) que oferece uma modalidade voltada para o ambiente acadêmico. Por meio do modo “*Academic*”, o sistema faz possível ao usuário pesquisar temas específicos com o suporte de fontes confiáveis, como artigos científicos, relatórios técnicos e bases de dados especializadas.

A ferramenta possui uma interface limpa e objetiva. Seu funcionamento é similar ao de buscadores acadêmicos, mas com a diferença sendo seu modo de organizar utilizando linguagem textual, o que torna mais dinâmico em alguns casos. O usuário pode fazer perguntas abertas, melhorar a busca com base em buscas anteriores e acessar de maneira direta as fontes citadas. Isso tudo o torna um *chatbot* útil para pesquisas exploratórias e para aprofundamento em temas específicos.

Por fim, esse *chatbot* é uma alternativa para se observar o comportamento do usuário em tarefas informacionais, avaliar o impacto emocional da clareza ou frustração na busca e investigar a percepção de confiabilidade nas respostas. As informações aqui apresentadas foram obtidas por meio de observações práticas realizadas durante o uso do *chatbot*, conforme os critérios estabelecidos na metodologia da pesquisa (consultar a Seção 4.2).

Figura 07 - Interface Padrão do *Perplexity*



Fonte: Elaboração do autor (2025)

²¹ *Perplexity*: <https://www.perplexity.ai/>

5.2.4 Análise comparativa dos Chatbots

A Tabela 6 apresenta uma análise comparativa entre os três *chatbots* selecionados: *ScienceOS*, *Study Skills Coach AI* e *Perplexity* (no modo acadêmico). A avaliação baseou-se em critérios técnicos e de interação definidos para garantir a viabilidade do estudo, sendo eles: contexto de uso, idioma, acessibilidade da interface, gratuidade, embasamento em fontes reais, clareza e fluidez do diálogo e nível de empatia percebida. O objetivo desta análise comparativa é fundamentar a escolha da ferramenta que melhor se alinha aos objetivos do experimento, priorizando aquela com maior potencial de interação conversacional.

Tabela 6: Comparação dos Chatbots selecionados (Autor, 2025)

Critério de Análise	<i>ScienceOS</i>	<i>Study Skills Coach AI</i>	<i>Perplexity (modo acadêmico)</i>
Contexto de uso	Pesquisa científica	Auxílio no estudo	Busca informacional
Idioma	Inglês	Inglês	Inglês
Interface acessível (Web/Mobile)	Sim	Sim	Sim
Gratuito e sem cadastro	Sim	Sim	Sim
Respostas com base em fontes reais	Sim	Parcial	Sim
Clareza e fluidez na conversa	Alta	Média	Alta
Empatia e engajamento	Baixo	Média	Média

Fonte: Elaboração do autor (2025)

Por fim, com base nos critérios analisados, os três *chatbots* demonstram o potencial para a aplicação dos métodos propostos. O *ScienceOS* apresenta uma clareza maior e foco funcional, o *Study Skills Coach AI* se destaca por possuir um caráter mais interativo e pessoal e o *Perplexity* combina a sua precisão na informação fornecida e a sua boa fluidez nas conversas, sendo uma opção equilibrada.

Apesar de todos os *chatbots* selecionados operarem no idioma inglês, essa condição não compromete a proposta deste estudo. A escolha foi motivada pela ausência de soluções em português que atendessem simultaneamente aos critérios de inclusão preliminares

(disponibilidade pública, interação textual e interface funcional). É importante ressaltar que esses critérios funcionaram como pré-requisitos para a seleção, logo, todos os três *Chatbots* finalistas os atendem integralmente. Por essa razão, a análise comparativa presente na Tabela 6 focou exclusivamente nos aspectos qualitativos que diferenciam as ferramentas, como clareza, fluidez, empatia e engajamento, permitindo uma avaliação mais aprofundada da experiência de uso.

Considerando a Tabela 06, observa-se que os três *Chatbots* atendem aos critérios técnicos definidos, porém diferem significativamente em seu estilo de conversação e abordagem ao usuário. O *ScienceOS* se destaca pela objetividade técnica voltada à pesquisa científica, adotando um tom estritamente formal e utilitário. Essa característica, embora eficiente para buscas de dados, inibe a criação de um vínculo empático, reduzindo a probabilidade de o usuário manifestar reações afetivas variadas. O *Perplexity*, por sua vez, equilibra fluidez e precisão informacional, mas mantém um perfil predominantemente impessoal. Já o *Study Skills Coach AI* diferencia-se pela proposta de atuar como um mentor (*coach*). Sua linguagem mais acolhedora e personalizada favorece o engajamento, tornando-o o candidato ideal para um estudo que visa avaliar não apenas a eficácia das respostas, mas também as dimensões emocionais e hedônicas da experiência.

Assim o *Study Skills Coach AI* foi selecionado como o *chatbot* principal para a aplicação dos métodos *AttrakDiff*, *Think Aloud* e *MAX*. Essa escolha se justifica pela sua capacidade de estimular interações mais humanas e motivacionais, favorecendo a coleta de dados significativos para a triangulação proposta nesta pesquisa.

5.3 Resultados do Experimento

Esta seção apresenta os resultados obtidos a partir da aplicação dos métodos *Think Aloud*, *MAX* e *AttrakDiff*, realizados com 6 participantes, entre os dias 21 a 28 de novembro de 2025.

5.3.1 Resultados do *AttrakDiff*

O questionário foi utilizado para mensurar quantitativamente a percepção dos usuários sobre o chatbot *Study Skills Coach AI*. A avaliação baseia-se em uma escala de diferencial semântico de 7 pontos, onde os extremos são definidos por adjetivos opostos (ex: "Confuso" vs. "Claro"). O usuário deve selecionar o valor que melhor representa sua opinião, onde 1 indica a adesão total ao adjetivo negativo, 7 ao positivo, e 4 representa a neutralidade. A Figura 08 exemplifica a estrutura de uma das perguntas aplicadas.

Figura 08 - Exemplo de item do questionário AttrakDiff

Seção 2 de 5

1. Facilidade de uso (Qualidade Pragmática)

Aqui avaliamos se o chatbot é fácil de usar e se ajuda nas tarefas.

O chatbot é

1234567

Diffícil de usarFácil de usar

Fonte: Elaboração do autor (2025)

As pontuações foram calculadas com base na média dos itens respondidos por cada participante. A Tabela 7 detalha os resultados individuais (P1 a P6) e a média global para cada dimensão avaliada.

Tabela 7: Resultados das Médias do AttrakDiff (Autor, 2025)

Dimensão	P1	P2	P3	P4	P5	P6	Média Geral
Qualidade Pragmática (PQ)	2.00	5.00	6.00	2.50	6.50	5.50	4.58
Qualidade Hedônica - Estimulação (QH-E)	3.00	5.00	5.50	3.00	6.00	5.50	4.67
Qualidade Hedônica - Identidade (QH-I)	3.25	5.25	6.25	3.00	5.00	4.25	4.50
Atratividade Global (ATT)	4.00	6.20	6.00	3.20	5.20	3.60	4.70

Fonte: Elaboração do autor (2025)

Os resultados mostram uma percepção geral ligeiramente acima do ponto neutro. O chatbot não foi avaliado de forma negativa, mas também não demonstrou qualidades excepcionais em nenhuma de suas dimensões. O destaque para a Atratividade Global (4.70) e a Estimulação (4.67) sugere que os participantes perceberam a ferramenta como visualmente agradável, inovadora e motivadora em um primeiro momento. Em contrapartida, as

pontuações comparativamente menores em Qualidade Pragmática (4.58) e Identidade (4.50) indicam que a eficiência, a previsibilidade e a confiança transmitida pelo sistema não acompanharam o mesmo nível de satisfação estética, sinalizando uma possível lacuna entre a apresentação da ferramenta e a sua utilidade prática.

5.3.2 Resultados do *MAX*

O método *MAX* capturou as percepções emocionais e pragmáticas dos participantes. Com o objetivo de sintetizar os dados e evidenciar os padrões de experiência predominantes no grupo, a Tabela 8 apresenta apenas as cartas que obtiveram recorrência (selecionadas por dois ou mais participantes). Esse critério de filtragem foi adotado para priorizar a análise dos consensos e das percepções compartilhadas, permitindo uma visualização mais clara das tendências de comportamento em detrimento de ocorrências isoladas.

Tabela 8: Cartas mais frequentes no Método *MAX*

Categoria	Carta (Tradução)	Votos
Emoção	Interessado(a)	4
Emoção	Confuso(a)	3
Emoção	Satisfeito(a)	2
Facilidade de Uso	Fácil de usar	5
Facilidade de Uso	Uso intuitivo	5
Intenção de Uso	Usaria outro	4
Intenção de Uso	Usaria com frequência	3
Intenção de Uso	Gostei de usar	3
Utilidade	Foi útil	4
Utilidade	Isso pode ajudar	4
Utilidade	Perdi tempo usando	2

Fonte: Elaboração do autor (2025)

A análise condensada dos dados revela um cenário de polarização na Experiência do Usuário. Nas categorias Facilidade de Uso e Utilidade, observa-se um consenso positivo, o

que indica que a interface foi percebida como acessível e o propósito da ferramenta como válido. No entanto, as categorias Emoção e Intenção de Uso expõem um conflito crítico, o interesse inicial (4 votos) foi confrontado pela confusão (3 votos), e a facilidade de navegação não foi suficiente para garantir a retenção, visto que a intenção de "usar outro" (4 votos) superou a de continuidade. A presença da carta "Perdi tempo usando" é um indicador qualitativo severo, sugerindo que, para uma parcela dos usuários, as falhas na interação anularam os benefícios da interface, resultando em uma experiência final frustrante e pouco confiável.

5.3.3 Resultados do *Think Aloud*

A aplicação do método *Think Aloud* foi importante para qualificar a experiência dos usuários, pois permitiu identificar não apenas o que os participantes sentiram, mas quais eventos específicos da interação foram responsáveis pela reação. A análise das transcrições revelou várias oscilações que interromperam a satisfação e geraram uma frustração nos usuários.

Inicialmente, a barreira do idioma foi o primeiro obstáculo identificado. Vale ressaltar que a seleção dos participantes considerou um nível mínimo de leitura na língua inglesa como pré-requisito, mas não excluiu aqueles com proficiência básica, visando representar a realidade do estudante médio que frequentemente utiliza ferramentas técnicas em inglês. Contudo, essa decisão evidenciou que, embora a tarefa fosse viável, a interface em língua estrangeira exigiu um esforço mental significativo. A necessidade de tradução constante gerou confusão e quebra de fluidez, transformando a língua em um fator de fricção. O P6, por exemplo, reagiu com surpresa negativa: “*Nossa, mas ele só funciona em inglês assim. [...] Só não estou entendendo o inglês, né?*”. Do mesmo modo, o P4 teve dificuldade em compreender a resposta inicial do sistema, o que gerou uma tentativa frustrada de forçar a mudança de idioma manualmente.

Apesar da barreira inicial, houve momentos positivos na interação. Durante as tarefas 2 a 5, o *Chatbot* desempenhou bem o seu papel de “conselheiro de estudos”. Os participantes elogiaram a interatividade do sistema. O P2 destacou positivamente o seu comportamento: “*Ele realmente pergunta sobre o que você quer [...] tenta dar uma resposta. [...] ele não tenta dar uma resposta rápida assim como outras IAs*”. A interface também foi elogiada, com o P3 mencionando a estética do avatar presente na tela inicial e a apresentação geral do site como sendo bonito e agradável.

A percepção de qualidade foi colocada à prova quando o *Chatbot* foi submetido à Tarefa 6, que exigia a revisão de tópicos técnicos da disciplina de Algoritmos e Programação I, componente curricular já cursado por todos os participantes. Essa limitação tornou-se evidente através das verbalizações dos usuários, que constataram que o *Chatbot*, embora treinado para habilidades de estudo, não mostrou competência para explicar códigos de forma didática. A resposta gerada foi percebida como confusa e insuficiente. O P6 verbalizou sua insatisfação de forma direta: “*Eu não entendi foi nada do que ela disse. Não foi útil*”. O P1 também relatou instabilidade técnica, indicando que o fluxo da interação foi quebrado naquele ponto específico.

O ponto crítico da avaliação foi a identificação de um erro grave de usabilidade relacionado ao limite de caracteres. O *Think Aloud* capturou o momento exato em que a experiência de vários usuários foi interrompida porque o sistema não aceitava mensagens curtas ou comandos simples, ocasionando o bloqueio do *chat*. O P5 descreveu o problema com precisão: “*Não está enviando por algum motivo [...] cheguei no limite de caracteres*”. Esse erro técnico teve um impacto severo na percepção de valor da ferramenta. O P4, ao se deparar com a interrupção, mostrou o sentimento de desperdício: “*Travou, que tristeza travar. [...] Eu sinto que estou perdendo tempo aqui*”. Esses relatos corroboram diretamente as cartas “*Perdi tempo usando*” e “*Usaria outro*” selecionadas no método *MAX*, confirmando que a falha técnica foi um fator determinante para a rejeição do sistema.

6. DISCUSSÃO

A presente pesquisa buscou investigar de que maneira a triangulação de métodos distintos contribui para a avaliação da Experiência do Usuário (UX) em *chatbots* acadêmicos, alinhando-se ao objetivo de compreender as percepções cognitivas, funcionais e emocionais dos usuários durante a interação. A análise integrada dos dados indicou que a utilização isolada de qualquer um dos métodos teria conduzido a conclusões parciais ou equivocadas sobre a qualidade do *chatbot Study Skills Coach AI*. Para detalhar esses achados, este capítulo discute, na Seção 6.1, a dissonância identificada entre a interface visual e a interação real e, na Seção 6.2, analisa a complementaridade diagnóstica dos métodos e o potencial de generalização dessa abordagem.

6.1 Diferença entre a Interface e a Interação

O cruzamento dos dados revelou uma contradição fundamental na experiência dos participantes: o *chatbot* demonstrou possuir uma interface visual (UI) percebida como intuitiva e de fácil navegação, mas falhou na qualidade da conversação (UX). Os resultados do método *MAX* demonstraram essa dualidade de forma clara. Por um lado, houve uma concordância expressiva entre 83% dos participantes (5 de 6) de que o sistema era "Fácil de usar" e de "Uso intuitivo". Essa percepção positiva foi corroborada pelas notas de Atratividade Global no *AttrakDiff*, que se mantiveram acima da média. Visualmente e operacionalmente, a ferramenta cumpriu seu papel inicial de apresentar uma interface acessível e esteticamente agradável.

No entanto, essa facilidade de navegação não se traduziu em satisfação geral ou em intenção de uso continuado (retenção). Paradoxalmente, a maioria dos participantes (67%, ou 4 de 6) indicou no método *MAX* que "Usaria outro" *Chatbot* ou que "Perdeu tempo usando". Essa rejeição, que contrasta com a alta avaliação de usabilidade da interface, encontra sua explicação nos dados qualitativos do método *Think Aloud*. As transcrições revelaram que a frustração não vinha da dificuldade em operar os comandos visuais, mas sim da incapacidade do sistema em manter uma conversa funcional em momentos críticos, seja por falhas técnicas (como bloqueios na interação) ou por insuficiência na qualidade das respostas (como observado na Tarefa 5).

6.2 A complementaridade dos métodos na avaliação de *Chatbots*

A comparação entre os resultados obtidos permite compreender o papel específico que cada método desempenha no diagnóstico de UX. O questionário *AttrakDiff*, por sua natureza quantitativa e baseada em médias, agiu como um termômetro de baixa sensibilidade. Ao apresentar resultados medianos, ele acabou por suavizar a polarização das opiniões, ocultando tanto a satisfação quanto a frustração intensa relatada individualmente pelos participantes. Se a avaliação dependesse apenas desse método, a conclusão seria de que o *Chatbot* é "aceitável", ignorando que usuários específicos tiveram experiências catastróficas que os levaram a abandonar a tarefa.

O método *MAX*, mostrou ser superior em capturar a dimensão emocional e das contradições da experiência. Ele desenhou um retrato mais fiel da jornada do usuário, mostrando que uma expectativa inicial alta pode se converter em frustração. O fato de um participante ter utilizado uma carta de emoção negativa (Nervoso) para classificar a utilidade

do sistema é um indicativo de como as falhas funcionais podem impactar diretamente o estado afetivo do usuário, algo que o *MAX* capturou com eficácia.

Por fim, o método *Think Aloud* se consolidou como o elemento diagnóstico indispensável. Enquanto o *MAX* apontou que os usuários estavam frustrados e queriam abandonar, apenas o *Think Aloud* explicou o porquê. Foi através da verbalização em tempo real que se identificou o problema do limite de caracteres como o problema central da rejeição. Sem esses relatos verbais, seria impossível distinguir se a insatisfação vinha do conteúdo, da interface ou de um erro técnico. Portanto, se conclui que para sistemas conversacionais, a métrica de satisfação (dada pelo *AttrakDiff* e *MAX*) é insuficiente sem a narrativa do processo (capturada pelo *Think Aloud*) que contextualiza as falhas durante a interação.

Embora aplicada neste estudo especificamente a um *chatbot*, a metodologia de triangulação demonstra um potencial teórico de replicabilidade para outros artefatos digitais, como *websites* e aplicativos móveis. Conforme destacado por Barbosa et al. (2022), os métodos como *AttrakDiff*, *Think Aloud* e *MAX* são instrumentos consolidados na literatura de Interação Humano-Computador (IHC) e, embora tenham sido aplicados a um *Chatbot acadêmico*, possuem raízes na avaliação de sistemas interativos gerais. A universalidade dessa abordagem está na capacidade de capturar dimensões transversais da experiência, enquanto o *Think Aloud* mapeia processos cognitivos de navegação e resolução de tarefas, o *MAX* e o *AttrakDiff* mensuram respostas emocionais e atratividade que independem da modalidade de interação. Logo, a combinação destas três perspectivas compõe a criação de um diagnóstico versátil, eficaz para identificar dissonâncias entre a interface visual e a funcionalidade real em qualquer produto digital onde a satisfação do usuário seja um critério determinante de qualidade.

6.3 Análise Comparativa com Resultados da Literatura

Ao confrontar os resultados desta pesquisa com o estudo de Barbosa et al. (2022), nota-se que a escolha da triangulação metodológica (*AttrakDiff*, *Think Aloud* e *MAX*) foi eficaz em ambos os casos, porém aplicada em contextos distintos. Enquanto Barbosa et al. (2022) validaram a viabilidade dessas técnicas avaliando um *chatbot* na área da saúde, o presente estudo utilizou a metodologia para um *chatbot* acadêmico. A principal distinção está no objetivo da aplicação, enquanto o trabalho de referência focou na aceitação das técnicas pelos avaliadores, esta pesquisa buscou investigar explicitamente a complementaridade

diagnóstica entre os métodos, testando se a combinação seria capaz de revelar falhas em um ambiente educacional, onde a uma menor tolerância ao erro.

Uma divergência fundamental foi observada no comportamento dos dados. No estudo de Barbosa et al. (2022), os métodos convergiram para uma avaliação positiva consistente, confirmando a qualidade do sistema analisado. Em contraste, nesta pesquisa, a triangulação expôs uma dissonância. O *AttrakDiff* apresentou médias moderadas (entre 4,58 e 4,70), sugerindo uma aceitação razoável impulsionada pela estética da interface. No entanto, essa métrica quantitativa falhou em detectar a frustração severa causada pelos bloqueios técnicos, que só foi plenamente revelada pelo método qualitativo *Think Aloud* e projetivo *MAX*.

Essa discrepância reforça a tese de complementaridade, mas sob uma nova ótica. No contexto acadêmico, a aparência agradável da interface capturada pelo *AttrakDiff*, mascarou problemas funcionais críticos. Foi o Método *MAX* que atuou como o diferencial decisivo, evidenciando a rejeição do usuário através das cartas de "Perdi tempo" e "Usaria outro". Portanto, conclui-se que a aplicação conjunta dos métodos é indispensável para evitar "falsos-positivos" de qualidade, provando que a satisfação estética medida por escalas hedônicas não sustenta a experiência quando a funcionalidade pedagógica falha.

6.4 Ameaças à Validade e Limitações do Estudo

A interpretação dos resultados apresentados deve considerar limitações inerentes ao desenho experimental, que podem ter influenciado as percepções coletadas. A primeira ameaça à validade refere-se à barreira do idioma. O *Chatbot* selecionado, *Study Skills Coach AI*, opera exclusivamente em língua inglesa. Embora a seleção dos participantes tenha considerado um nível mínimo de proficiência, foi observado que a necessidade de tradução e a insegurança linguística geraram uma carga cognitiva adicional. Essa situação pode ter sido interpretada por alguns usuários como uma falha de "facilidade de uso" do sistema, quando, na verdade, tratava-se de uma dificuldade de acesso ao conteúdo. Portanto, os resultados de usabilidade devem ser lidos sob a ótica de usuários não-nativos interagindo com interfaces estrangeiras.

Outra ameaça à validade diz respeito à ordem de aplicação dos métodos. O protocolo seguiu a sequência: *Think Aloud* (durante a interação) → *MAX* (pós-interação) → *AttrakDiff*. Existe a possibilidade do esforço mental exigido pela verbalização constante no *Think Aloud* ter influenciado negativamente a avaliação emocional subsequente no *MAX* e no *AttrakDiff*.

(devido ao efeito de fadiga). No entanto, essa ordem foi mantida para garantir que os dados do *MAX* fossem coletados enquanto a memória da interação estava fresca, priorizando a precisão emocional sobre o conforto.

7. CONCLUSÃO

Este trabalho analisou a avaliação da Experiência do Usuário em *chatbots* acadêmicos, confirmando que a triangulação dos métodos *AttrakDiff*, *Think Aloud* e *MAX* é fundamental para compreender a complexidade da interação em interfaces conversacionais voltadas ao contexto educacional. A avaliação experimental do *chatbot Study Skills Coach AI* revelou que o uso isolado de métodos quantitativos, como o *AttrakDiff*, resultaria em uma visão superficial da qualidade do sistema. Foi a combinação com métodos qualitativos que permitiu expor a contradição central da experiência: uma interface considerada visualmente agradável e fácil de navegar (UI), mas que falha em sustentar uma conversa funcional (UX), gerando frustração e desejo de abandono nos usuários.

O método *Think Aloud* se consolidou como o elemento diagnóstico indispensável da pesquisa, pois foi capaz de identificar as causas da insatisfação revelada pelo método *MAX*. O estudo mostrou que a rejeição ao sistema não veio da dificuldade de navegação, mas sim de uma soma de fatores críticos: o erro técnico do limite de caracteres não informado, a barreira do idioma que elevou o esforço cognitivo e a insuficiência do *chatbot* em responder a tópicos técnicos específicos. Esses problemas mostram que, em ferramentas acadêmicas, a tolerância a erros é mínima, quando a utilidade da ferramenta é interrompida por problemas técnicos ou respostas inadequadas, a percepção de valor severamente comprometida, fazendo um contraponto com os benefícios que foram observados, como a boa interface gráfica, por exemplo.

Para trabalhos futuros, sugere-se primeiramente a replicação desta metodologia de triangulação em *chatbots* baseados em Grandes Modelos de Linguagem (LLMs) mais robustos, buscando verificar se a dissonância entre a qualidade da interface e a precisão da resposta persiste em sistemas mais avançados. Adicionalmente, recomenda-se comparar os métodos aqui utilizados (*AttrakDiff*, *Think Aloud* e *MAX*) com abordagens clássicas, como a Avaliação Heurística e o *System Usability Scale* (SUS), para validar se a discrepância diagnóstica encontrada no experimento se mantém frente a métricas tradicionais. Por fim, sugere-se investigar o impacto da barreira linguística como uma variável isolada (comparando *chatbots* nativos no idioma inglês versus *chatbots* traduzidos para o idioma português brasileiro) e aprofundar o estudo de métricas específicas para empatia em agentes educacionais, visto que a personalização inicial foi um dos poucos aspectos positivos observados antes das falhas técnicas interromperem a interação.

REFERÊNCIAS

- ADAMOPOULOU, Eleni; MOUSSIADES, Lefteris. An overview of chatbot technology. In: Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5-7, 2020, Proceedings, Part II 16. Springer International Publishing, 2020. p. 373-383.
- BARBOSA, Marcus et al. Um Estudo Exploratória sobre Métodos de Avaliação de User Experience em Chatbots. Universidade Federal do Pampa (UNIPAMPA). 2022. Disponível em: <https://sol.sbc.org.br/index.php/eres/article/view/22369>. Acesso em: 6 jun. 2025.
- BARBOSA, Marcus et al. UX of Chatbots: An Exploratory Study on Acceptance of User Experience Evaluation Methods. In: ICEIS (2). 2022. p. 355-363.
- BATISTA, Gabriely Oliveira da Silva; MONTEIRO, Mateus de Souza; SALGADO, Luciana Cardoso de Castro. Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do Brasil. Instituto de Computação - Universidade Federal Fluminense (UFF). 2022. Disponível em: <https://sol.sbc.org.br/index.php/wcge/article/view/20712>. Acesso em: 6 jun. 2025.
- BRIVA-IGLESIAS, Vicent; O'BRIEN, Sharon. Measuring Machine Translation User Experience (MTUX): A Comparison between AttrakDiff and User Experience Questionnaire. In: Proceedings of the 24th Annual Conference of the European Association for Machine Translation. 2023. p. 335-344.
- CAVALCANTE, Emanuelle; RIVERO, Luis; CONTE, Tayana. MAX: A Method for Evaluating the Post-use User Experience through Cards and a Board. In: Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering (SEKE). 2015. DOI: 10.18293/SEKE2015-136.
- CHARTERS, Elizabeth. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Education Journal, v. 12, n. 2, 2003. Disponível em: <https://journals.library.brocku.ca/brocked/index.php/home/article/view/38>. Acesso em: 6 jun. 2025.
- COELHO, Natália et al. Avaliação da usabilidade de um chatbot por meio das heurísticas de Nielsen adaptadas. Universidade Federal do Ceará (UFC) - Campus Russas. 2022. Disponível em: https://sol.sbc.org.br/index.php/ihc_estendido/article/view/22052. Acesso em: 6 jun. 2025.
- FØLSTAD, Asbjørn; BRANDTZÆG, Petter Bae. Chatbots and the new world of HCI. **interactions**, v. 24, n. 4, p. 38-42, 2017.
- HASSENZAHN, Marc; BURMESTER, Michael; KOLLER, Franz. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: SZWILLUS, G.; ZIEGLER, J. (org.). Mensch & Computer 2003: Interaktion in Bewegung. Stuttgart: B. G. Teubner, 2003. p. 187-196.
- ISO 9241-210:2019. Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. International Organization for Standardization, 2019.

JASPERS, Monique W. M. et al. The think aloud method: a guide to user interface design. *International Journal of Medical Informatics*, v. 73, p. 781-795, 2004. DOI: 10.1016/j.ijmedinf.2004.08.003.

KHAMAJ, Abdulrahman. AI-enhanced chatbot for improving healthcare usability and accessibility for older adults. *Alexandria Engineering Journal*, v. 116, p. 202-213, 2025.

LANGVIN, Raina et al. Heuristic evaluation of conversational agents. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021. p. 1-15.

LIU, Tingting et al. The illusion of empathy: How ai chatbots shape conversation perception. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2025. p. 14327-14335.

NORMAN, Donald; NIELSEN, Jakob. The definition of user experience (UX). Nielsen Norman Group, 2020. Disponível em: <https://www.nngroup.com/articles/definition-user-experience/>. Acesso em: 6 jun. 2025.

PREECE, Jenny; ROGERS, Yvonne; SHARP, Helen. *Interaction Design: Beyond Human-Computer Interaction*. 3. ed. New York: John Wiley & Sons, 2013.

RAPP, Amon; CURTI, Lorenzo; BOLDI, Arianna. The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, v. 151, 102630, 2021. DOI: 10.1016/j.ijhcs.2021.102630.

SILVA, Geovana Ramos Sousa; CANEDO, Edna Dias. Human Factors in the Design of Chatbot Interactions: Conversational Design Practices. In: *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*. 2024. p. 1-12.

SILVA, Geovana Ramos Sousa; CANEDO, Edna Dias. Towards user-centric guidelines for chatbot conversational design. *International Journal of Human-Computer Interaction*, v. 40, n. 2, p. 98-120, 2024.

APÊNDICES

APÊNDICE A - TRABALHOS RESULTANTES DO ESTUDO BIBLIOGRÁFICO

Referências Bibliográficas dos Estudos Seleccionados
BATISTA, Gabriely Oliveira da Silva; MONTEIRO, Mateus de Souza; SALGADO, Luciana Cardoso de Castro. Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do Brasil. In: LATIN AMERICAN SYMPOSIUM ON DIGITAL GOVERNMENT (LASDIGOV), 10. , 2022, Niterói. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022 . p. 73-84. ISSN 2763-8723. DOI: https://doi.org/10.5753/wcge.2022.223318 .
COELHO, Natália; SANTOS, Alex Alan; DA SILVA, Naum; DA SILVA, Pedro; LIMA, Samuel; ARRUDA, Alexandre; MARQUES, Anna Beatriz. Avaliação da usabilidade de um chatbot por meio das heurísticas de Nielsen adaptadas. In: PÔSTERES E DEMONSTRAÇÕES - SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 21. , 2022, Diamantina. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022 . p. 124-127. DOI: https://doi.org/10.5753/ihc_estendido.2022.226185 .
PIRES, Isabella; CASELI, Helena; NERIS, Vânia. Design de um chatbot para o diálogo com universitários com possível perfil depressivo. In: CONCURSO DE TRABALHOS DE INICIAÇÃO CIENTÍFICA - SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE (SBCAS), 23. , 2023, São Paulo/SP. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023 . p. 7-12. ISSN 2763-8987. DOI: https://doi.org/10.5753/sbcas_estendido.2023.229543 .
BARBOSA, Marcus; VALLE, Pedro; NAKAMURA, Walter; GUERINO, Guilherme; FINGER, Alice; LUNARDI, Gabriel; SILVA, Williamson. Um Estudo Exploratório sobre Métodos de Avaliação de User Experience em Chatbots. In: ESCOLA REGIONAL DE ENGENHARIA DE SOFTWARE (ERES), 6. , 2022, Blumenau. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022 . p. 21-30. DOI: https://doi.org/10.5753/eres.2022.227723 .
FREIRE, André Pimenta; CARDOSO, Paula Christina Figueira; SALGADO, André de Lima. May We Consult ChatGPT in Our Human-Computer Interaction Written Exam? An Experience Report After a Professor Answered Yes. In: SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 22. , 2023, Maceió/AL. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023 .
SILVA, Geovana Ramos Sousa; CANEDO, Edna Dias. Human Factors in the Design of Chatbot Interactions: Conversational Design Practices. In: SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 23. , 2024, Brasília/DF. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2024 . p. 487-498.

DE MELO, Diogo Nazareno Almeida; MONTEIRO, Ingrid Teixeira. Communication and Personality: how COVID-19 government chatbots express themselves. In: SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 20. , 2021, Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021 .

DE SOUZA, Paula Maia; PIRES, Isabella da Costa; MOTTI, Vivian Genaro; CASELI, Helena Medeiros; BARBOSA NETO, Jair; MARTINI, Larissa C.; NERIS, Vânia Paula de Almeida. Design recommendations for chatbots to support people with depression. In: SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 21. , 2022, Diamantina. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022.

DE SOUZA, Ana Carolina Rossi; MARIANO, Pamella Augusta De Lima; GUERINO, Guilherme Corredato; CHAVES, Ana Paula; VALENTIM, Natasha Malveira Costa. Technologies for Hedonic Aspects Evaluation in Text-based Chatbots: A Systematic Mapping Study. In: SIMPÓSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC), 22. , 2023, Maceió/AL. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023

APÊNDICE B - ROTEIRO DE TAREFAS PARA O *THINK ALOUD*

Este roteiro foi utilizado para guiar os participantes durante a interação com o *Chatbot Study Skills Coach AI*, seguindo o protocolo do método *Think Aloud*.

Tarefa 01

- Acesse o *Chatbot Study Skills Coach AI*.
- Observe a tela inicial e descreva suas primeiras impressões (o que chama sua atenção, o que você entende sobre o que ele faz).

Tarefa 02

- Peça ao *Chatbot* dicas para organizar seus horários de estudo para uma prova próxima.
- Leia a resposta e diga o que você achou dela: foi útil? clara? esperava algo diferente?

Tarefa 03

- Peça ao *Chatbot* uma técnica para memorizar conteúdos difíceis.
- Comente em voz alta o que achou da explicação e se ela faz sentido para você.

Tarefa 04

- Peça ajuda ao *Chatbot* para montar um plano de estudos que caiba na sua rotina diária.
- Fale o que achou das sugestões, parecem realistas? aplicáveis?

Tarefa 05

- Peça dicas para melhorar sua concentração enquanto estuda.
- Diga se as respostas são motivadoras, práticas ou genéricas.

Tarefa 06

- Peça ao *Chatbot* ajuda para revisar conteúdos da disciplina Algoritmos e Programação 1.
- Escolha um dos tópicos abaixo e faça uma pergunta sobre ele:
 - Estruturas de repetição (como *for*, *while* e *do...while* em C ou Python).
 - Estruturas condicionais (*if*, *else*, *elif*).
 - Declaração e uso de variáveis (tipos de dados, entrada e saída).

- Comente se a resposta foi clara, útil ou se faltou algo que você esperava.

Tarefa 07

- Peça ao *Chatbot* um resumo das dicas que ele deu durante a conversa.
- Diga se esse resumo ajuda a lembrar das informações.
- Por fim, dê sua opinião geral: O chatbot foi fácil de usar? Você usaria novamente?

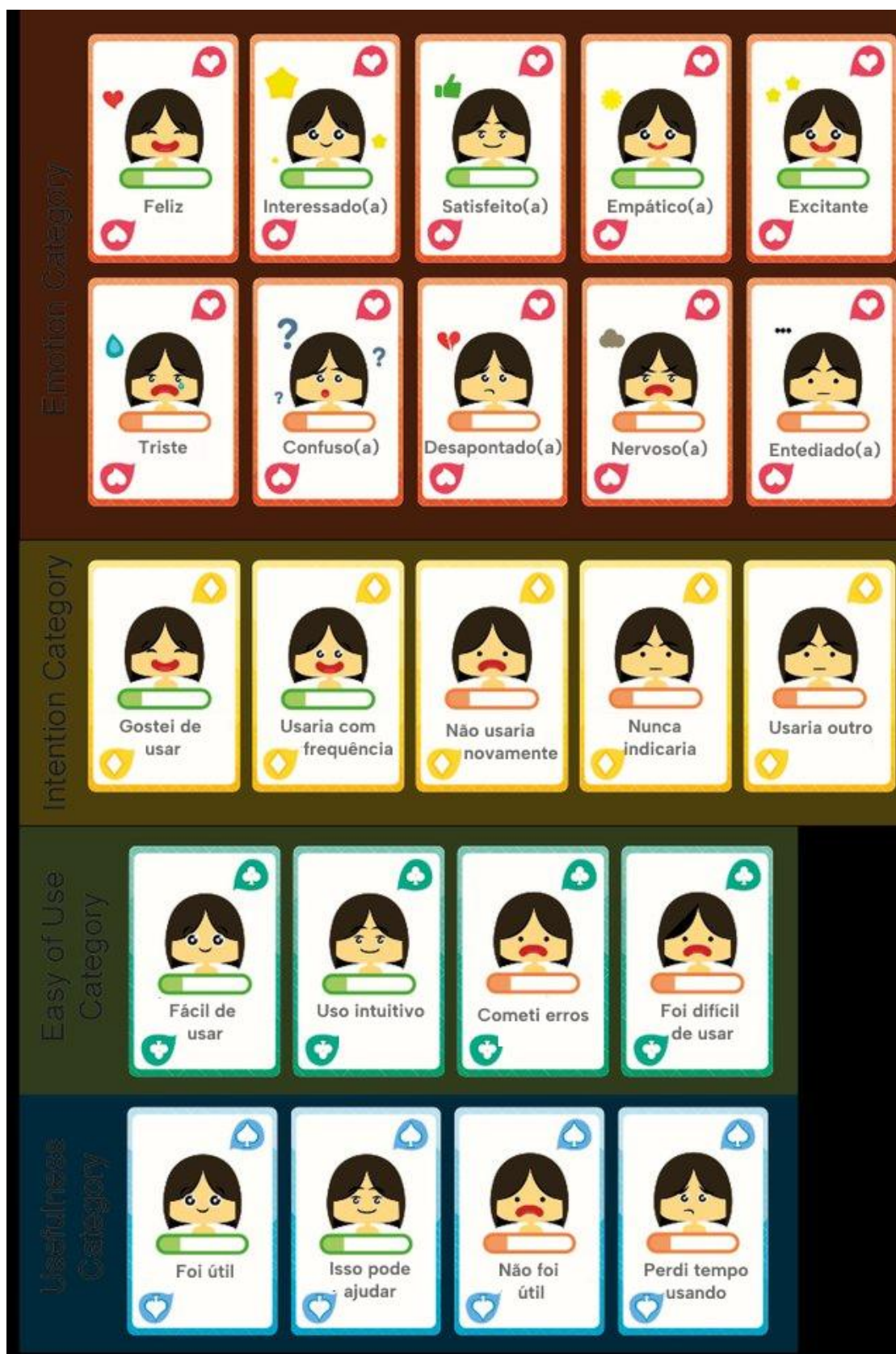
APÊNDICE C – ARTEFATOS ADAPTADOS DO MÉTODO *MAX* PARA IMPRESSÃO

Este apêndice disponibiliza os materiais visuais do método *MAX* (Tabuleiro e Cartas), traduzidos e adaptados para o português brasileiro pelo autor, visando facilitar a replicação da metodologia em estudos futuros.

1. Cartas de Avaliação

Recomenda-se a impressão colorida e o recorte individual dos itens para aplicação dinâmica durante as sessões de avaliação. O conjunto completo de cartas traduzidas e adaptadas é apresentado na Figura 09.

Figura 09 - Modelo das Cartas MAX para Impressão



Fonte: Elaborado pelo autor (2025).

2. Tabuleiro de Aplicação

O tabuleiro deve ser impresso preferencialmente em formato A4 ou A3 para permitir a disposição confortável das cartas pelos participantes. O modelo para impressão é apresentado na Figura 10.

Figura 10 - Modelo das Cartas *MAX* para Impressão

QUADRO MAX			
o que você sentiu ao usá-lo?	Foi fácil de usar?	Você deseja usá-lo?	Foi útil?
			
			

Fonte: Elaborado pelo autor (2025).

APÊNDICE D – QUESTIONÁRIO *ATTRAKDIFF* ADAPTADO

Este apêndice apresenta a visualização do instrumento de coleta de dados utilizado para a aplicação do método *AttrakDiff*. O questionário foi estruturado na plataforma *Google Forms*, utilizando escalas de diferencial semântico de 7 pontos. O formulário completo encontra-se disponível para consulta através do link: <https://forms.gle/FcwKUdDtVHn8k3Au7>.

Abaixo, são apresentadas as capturas de tela das quatro seções avaliativas.

Figura 11 - Seção 1: Qualidade Pragmática

1. Facilidade de uso (Qualidade Pragmática)

Aqui avaliamos se o chatbot é fácil de usar e se ajuda nas tarefas.

O chatbot é

	1	2	3	4	5	6	7	
Difícil de usar	☐	☐	☐	☐	☐	☐	☐	Fácil de usar

O chatbot é

	1	2	3	4	5	6	7	
Confuso	☐	☐	☐	☐	☐	☐	☐	Claro

O chatbot é

	1	2	3	4	5	6	7	
Desorganizado	☐	☐	☐	☐	☐	☐	☐	Organizado

O chatbot é

	1	2	3	4	5	6	7	
Ineficiente	☐	☐	☐	☐	☐	☐	☐	Eficiente

O chatbot é

	1	2	3	4	5	6	7	
Complicado	☐	☐	☐	☐	☐	☐	☐	Simples

Fonte: Elaborado pelo autor (2025).

Figura 12 - Seção 2: Qualidade Hedônica (Estimulação)

2. Criatividade e novidade (Qualidade Hedônica – Estimulação)

Aqui avaliamos se o chatbot é interessante e traz novidades.

O chatbot é

1234567

Monótono☐☐☐☐☐☐☐

Estimulante

O chatbot é

1234567

Sem graça☐☐☐☐☐☐☐

Interessante

O chatbot é

1234567

Comum☐☐☐☐☐☐☐

Diferente

O chatbot é

1234567

Previsível☐☐☐☐☐☐☐

Surpreendente

O chatbot é

1234567

Pouco criativo☐☐☐☐☐☐☐

Criativo

Fonte: Elaborado pelo autor (2025).

Figura 13 - Seção 3: Qualidade Hedônica (Identidade)

3. Conexão e confiança (Qualidade Hedônica – Identidade)

Aqui avaliamos se o chatbot passa confiança e cria proximidade.

O chatbot é

1234567

Distante

Acolhedor

O chatbot é

1234567

Pouco confiável

Confiável

O chatbot é

1234567

Frio

Próximo

O chatbot é

1234567

Sem prestígio

Respeitável

O chatbot é

1234567

Pouco profissional

Profissional

Fonte: Elaborado pelo autor (2025).

Figura 14 - Seção 4: Atratividade Global

4. Impressão geral (Atratividade Global)

Aqui avaliamos a impressão geral sobre o chatbot.

O chatbot é

1234567

Feio☐☐☐☐☐☐☐

Bonito

O chatbot é

1234567

Desagradável☐☐☐☐☐☐☐

Agradável

O chatbot é

1234567

Ruim☐☐☐☐☐☐☐

Bom

O chatbot é

1234567

Pouco atrativo☐☐☐☐☐☐☐

Atrativo

O chatbot é

1234567

Desconfortável☐☐☐☐☐☐☐

Confortável

Fonte: Elaborado pelo autor (2025).