RIO

PUC

Dissertação de Mestrado

Persuasão Personalizada Mediada por Modelos de Linguagem de Grande Escala:

evidências teóricas e empíricas em saúde mental digital

Bruno Grossman

Pontifícia Universidade Católica do Rio de Janeiro

Centro de Teologia e Ciências Humanas

Departamento de Psicologia

Rio de Janeiro, 17 de dezembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Dissertação de Mestrado

Persuasão Personalizada Mediada por Modelos de Linguagem de Grande Escala:

evidências teóricas e empíricas sobre sua aplicação em saúde mental digital

Bruno Grossman

Orientação: Professor Leonardo Fernandes Martins
Coorientação: Professor Jairo Francisco de Souza
Professor Heder Soares Bernardino

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Psicologia pelo programa de Pós-Graduação em Psicologia Clínica, no Departamento de Psicologia

Rio de Janeiro, 17 de dezembro de 2025.

Pontifícia
Universidade
Católica do
Rio de Janeiro

Persuasão Personalizada Mediada por Modelos de Linguagem de Grande Escala:

evidências teóricas e empíricas sobre sua aplicação em saúde mental digital

Bruno Grossman

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Psicologia Clínica Aprovada pela Comissão examinadora abaixo:

Professor Leonardo Fernandes Martins
Orientador
Departamento de Psicologia PUC-Rio

Professor Heder Soares Bernardino
Co-Orientador
Universidade Federal de Juiz de Fora

Professor Jairo Francisco de Souza
Co-Orientador
Universidade Federal de Juiz de Fora

Professor Waldir Monteiro Sampaio
Departamento de Psicologia PUC-Rio

Professor Ricardo Primi
Universidade de São Francisco

Rio de Janeiro, 17 de dezembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Bruno Grossman

Graduou-se em Psicologia pela PUC-Rio e é especialista em Marketing pelo IAG PUC-Rio. Atua com inovação e transformação digital no Instituto ECOA PUC-Rio e integrou o Grupo de Trabalho em Inteligência Artificial na Psicologia do Conselho Federal de Psicologia. Seus interesses acadêmicos concentram-se na interseção entre Psicologia e tecnologia, com foco nos impactos, potenciais benefícios e riscos do uso da inteligência artificial na sociedade, especialmente em processos de influência mediados por IA, tomada de decisão humana e intervenções digitais em saúde mental.

Ficha Catalográfica

Grossman, Bruno

Persuasão personalizada mediada por modelos de linguagem de grande escala : evidências teóricas e empíricas em saúde mental digital / Bruno Grossman ; orientação: Leonardo Fernandes Martins ; coorientação: Jairo Francisco de Souza, Heder Soares Bernardino. – 2025.

91 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Psicologia, 2025.

Inclui bibliografia

1. Psicologia – Teses. 2. Persuasão personalizada. 3. Inteligência artificial. 4. Personalidade. 5. Saúde mental digital. I. Martins, Leonardo Fernandes. II. Souza, Jairo Francisco de. III. Bernardino, Heder Soares. IV. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Psicologia. V. Título.

CDD: 150

À minha família.

Agradecimentos

À minha esposa, Deborah, pelos papos leves, pelos papos profundos, pelas reflexões que me enriqueceram, por todo amor, o apoio incondicional, a compreensão dos inúmeros sacrifícios, pela tranquilidade que me trouxe ao longo desse ciclo e por sobreviver aos meus ensaios de apresentações.

Aos meus pais e irmãos, que desde cedo estimularam em mim a curiosidade, a educação, o humanismo e, não menos importante, o humor.

Ao meu orientador, Leonardo Martins, pela confiança, generosidade, amizade, os ensinamentos constantes, a inspiração e a serenidade ao longo do caminho.

Aos meus coorientadores, Heder Soares Bernardino e Jairo Francisco de Souza, pela atenção generosa e por todas as contribuições.

Aos professores Ricardo Primi e Waldir Sampaio, por aceitarem compor a Comissão Examinadora e por dedicarem seu tempo para a discussão desta dissertação.

Ao Rafael Nasser, pelo incentivo decisivo e o apoio para essa formação.

Aos colegas do ECOA PUC-Rio, pelo apoio, troca constante e compreensão nos momentos mais exigentes do mestrado.

Aos colegas do Grupo de Supervisão da UFJF e da PUC-Rio, pelo aprendizado compartilhado.

Aos colegas do Grupo de Trabalho em Psicologia e IA do Conselho Federal de Psicologia, pelo diálogo e colaboração.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de financiamento 001.

Resumo

Grossman, Bruno; Martins, Leonardo Fernandes. **Persuasão personalizada mediada por modelos de linguagem de grande escala: evidências teóricas e empíricas em saúde mental digital.** Rio de Janeiro, 2025. 91p. Dissertação de Mestrado – Departamento de Psicologia, Pontifícia Universidade Católica do Rio de Janeiro.

A alta prevalência de transtornos mentais e as desigualdades no acesso ao cuidado têm estimulado o interesse por intervenções digitais em saúde mental, incluindo agentes conversacionais baseados em inteligência artificial (IA). Nesse cenário, a persuasão personalizada mediada por IA tem sido explorada como estratégia para favorecer atitudes e adesão a esse tipo de intervenção. Esta dissertação teve como objetivo examinar a eficácia e as limitações da persuasão personalizada mediada por modelos de linguagem de grande escala na promoção da adoção de intervenções de saúde mental digital, com ênfase em um *chatbot* para manejo de estresse. Para atingir esse objetivo, a dissertação foi estruturada em dois estudos complementares. No Estudo I, foi conduzida uma revisão semi-sistemática da literatura recente, que reuniu oito artigos empíricos entre 2023 e 2025 (n= 26.584) nos domínios de política, consumo e saúde. Do total, 66% dos testes analisados relataram efeitos significativos a favor das mensagens personalizadas, com maior consistência quando a personalização se baseou em variáveis psicológicas, especialmente Abertura à Experiência e Extroversão. O domínio da saúde apresentou a maior proporção de efeitos significativos, apesar da escassez de estudos em saúde mental digital. No Estudo II, foi realizado um experimento com 451 participantes universitários brasileiros, comparando mensagens de um LLM personalizadas pelo traço Abertura à Experiência com uma mensagem genérica sobre o uso de *chatbots* para manejo do estresse diário. As análises não encontraram efeito significativo do tipo de mensagem nem interação com Abertura à Experiência, enquanto diferenças individuais, como maior Abertura à Tecnologia, menor Abertura à Experiência e área de formação, mostraram associação mais clara com atitudes favoráveis ou cautelosas em relação aos *chatbots*. De modo geral, os resultados indicam que a personalização mediada por IA pode produzir efeitos reais, porém moderados e sensíveis ao contexto e que, em aplicações voltadas à saúde mental, características individuais e a forma como a personalização é percebida podem pesar mais do que a manipulação automatizada de linguagem.

Palavras-chave

Persuasão personalizada; inteligência artificial; personalidade; saúde mental digital

Abstract

Grossman, Bruno; Martins, Leonardo Fernandes (Advisor). **Large language model mediated personalized persuasion: theoretical and empirical evidence in digital mental health**. Rio de Janeiro, 2025. 91pp. Master's thesis – Department of Psychology, Pontifical Catholic University of Rio de Janeiro.

The high prevalence of mental disorders and inequalities in access to care have fueled interest in digital mental health interventions, including conversational agents based on artificial intelligence (AI). In this context, AI mediated personalized persuasion has been explored as a strategy to promote favorable attitudes toward and adherence to such interventions. This dissertation aimed to examine the effectiveness and limitations of personalized persuasion mediated by large language models in promoting the adoption of digital mental health interventions, with a focus on a stress management chatbot. To achieve this objective, the dissertation was structured into two complementary studies. In Study I, a semi systematic review of the recent literature was conducted, encompassing eight empirical articles published between 2023 and 2025 (n= 26,584) across the domains of politics, consumption, and health. Overall, 66% of the analyzed tests reported significant effects in favor of personalized messages, with greater consistency when personalization was based on psychological variables, particularly Openness to Experience and Extraversion. The health domain showed the highest proportion of significant effects, despite the scarcity of studies specifically focused on digital mental health. In Study II, an experiment was conducted with 451 Brazilian university students, comparing LLM generated messages personalized according to Openness to Experience with a generic message about the use of chatbots for daily stress management. The analyses found no significant effect of message type nor an interaction with Openness to Experience. In contrast, individual differences such as higher Openness to Technology, lower Openness to Experience, and field of study showed clearer associations with favorable or more cautious attitudes toward chatbots. Overall, the findings suggest that AI mediated personalization can produce real but modest and context sensitive effects and that, in mental health applications, individual characteristics and the way personalization is perceived may weigh more heavily than automated language manipulation.

Keywords

Personalized persuasion; artificial intelligence; personality; digital mental health

Sumário

Lista de figuras

Lista de tabelas

Tem sido um equívoco comum em programas de saúde pública supor que mensagens e formas de comunicação podem servir igualmente para todas as populações, condições médicas ou circunstâncias.

Um conjunto crescente de pesquisas mostra que apresentar informações gerais de saúde sem considerar as necessidades individuais ou a relevância pessoal pode limitar de maneira substancial o alcance das mudanças de comportamento em saúde.
Lustria et al., 2013

Introdução

Segundo a OMS (2025), mais de um bilhão de pessoas vivem com algum transtorno mental, o que corresponde a cerca de 14% da população mundial. No Brasil e em outros países da América Latina, esse quadro é agravado por desigualdades regionais, instabilidade socioeconômica e limitações das políticas de atenção psicossocial, o que contribui para a manutenção de altas taxas de sofrimento psíquico e para a sobrecarga dos serviços disponíveis (Araújo; Torrenté, 2023).

Além disso, o acesso ao cuidado em saúde mental é limitado. Menos de um em cada dez indivíduos com depressão recebe tratamento adequado e a maioria das pessoas com psicose permanece sem acompanhamento formal. Em muitos países em desenvolvimento há escassez de profissionais e apenas uma parte do orçamento em saúde é destinada à área, o que aumenta o subtratamento e a insuficiência de serviços especializados (OMS, 2025).

Essas barreiras têm aumentado o interesse por agentes conversacionais autônomos baseados em inteligência artificial (IA), entendida aqui como sistemas computacionais capazes de aprender a partir de dados e de desempenhar tarefas que envolvem compreensão e tomada de decisão (IBM, 2024). Esses agentes, frequentemente denominados *chatbots* de saúde mental (Casu et al., 2024), interagem com usuários em linguagem natural para oferecer psicoeducação, apoio emocional e exercícios de manejo de estresse e de emoções, com alta disponibilidade e baixo custo (Chaudhry; Debi, 2024; Durden et al., 2023).

O avanço recente dos modelos de linguagem de grande escala (LLMs) tem permitido que esses agentes sustentem diálogos mais longos e contextualizados. Esses modelos são sistemas de IA treinados em grandes volumes de texto para processar e gerar textos em linguagem natural em diferentes tarefas (IBM, 2023). Isso ocorre em um contexto de alta conectividade, em que cerca de 68% da população mundial e 84% da população brasileira têm acesso à internet (World Bank Group, 2024).

Ainda assim, a disponibilidade tecnológica não garante adoção. Estudos indicam que atitudes em relação à IA são frequentemente ambivalentes e combinam interesse por inovação com dúvidas sobre privacidade, autenticidade das interações e adequação do apoio oferecido (Kaya et al., 2024; Varghese et al., 2024). Além disso, diferenças individuais, como traços de personalidade e familiaridade com

tecnologias digitais, influenciam a avaliação e a aceitação de agentes conversacionais. Entender esses fatores é essencial para desenvolver estratégias de comunicação mais eficazes e favorecer atitudes positivas e informadas em relação ao uso de *chatbots* voltados ao manejo do estresse e ao bem-estar psicológico.

A persuasão, definida como o processo de influenciar crenças, atitudes ou comportamentos por meio da comunicação (Luttrell et al., 2025), é um dos fenômenos centrais na Psicologia Social, que busca entender em que condições mensagens produzem mudanças mais duradouras em atitudes e respostas comportamentais (Druckman, 2022). Embora grande parte das mensagens persuasivas ainda siga formatos genéricos, nas últimas décadas consolidou-se na literatura teórica e experimental uma estratégia eficaz para tornar as mensagens mais convincentes: a persuasão personalizada. Ela é definida como o alinhamento deliberado de um ou mais aspectos da comunicação (conteúdo e forma da mensagem, fonte ou contexto) a uma ou mais características do destinatário, incluindo traços psicológicos, motivações, valores, identidades ou outros atributos individuais (Matz et al., 2023; Teeny et al., 2021).

O Modelo da Probabilidade de Elaboração (ELM) ajuda a entender por que a personalização tende a favorecer a eficácia persuasiva. O modelo propõe que o impacto de uma mensagem depende do nível de elaboração que o indivíduo dedica ao conteúdo. Em situações de baixo envolvimento, a personalização pode funcionar como pista de afinidade, facilitando a aceitação. Em contextos de maior reflexão, pode atuar como argumento relevante, influenciando a avaliação do conteúdo e influenciando a qualidade dos pensamentos gerados (Petty e Cacioppo, 1986; Teeny et al., 2021). Contudo, a personalização não garante um efeito positivo, pois seus resultados dependem de como o destinatário interpreta esse ajuste, podendo levar à aceitação, indiferença ou resistência.

A expansão dos algoritmos de aprendizado de máquina, aliada à digitalização das interações cotidianas, permitiu que diferenças psicológicas fossem inferidas a partir de rastros deixados em textos, plataformas *online* e padrões de uso de dispositivos (Park et al., 2015). Com a chegada dos modelos de linguagem de grande escala, as inferências psicológicas a partir de dados digitais tornaram-se mais acessíveis. Esses mesmos sistemas também passaram a gerar mensagens em linguagem natural ajustadas a tais perfis, o que a literatura descreve como persuasão mediada por IA ou por LLM (Matz et al., 2024; Rogiers et al., 2024).

Nesse contexto, a dissertação foi desenvolvida a partir de dois estudos. O Estudo I apresentou uma revisão semi-sistemática da literatura recente sobre persuasão personalizada mediada por inteligência artificial. A revisão reuniu oito artigos empíricos publicados entre 2023 e 2025, abrangendo principalmente os domínios de política, consumo e saúde, com um total de 26.584 participantes. Os resultados indicaram que 66% dos testes analisados mostraram efeitos significativos, com maior consistência quando a personalização se baseou em variáveis psicológicas. Entre os traços de personalidade, Abertura à Experiência e Extroversão apresentaram efeitos pequenos a moderados e estáveis entre estudos. Apesar desse padrão, as pesquisas permanecem concentradas em países de língua inglesa, com amostras recrutadas *online* e poucas aplicações em saúde, e nenhuma voltada especificamente à saúde mental ou à adoção de *chatbots* como intervenção digital.

O Estudo II buscou avançar sobre essa lacuna por meio de um experimento conduzido com estudantes universitários brasileiros, em língua portuguesa, no contexto específico do uso de *chatbots* para gerenciamento de estresse diário. O estudo avaliou se mensagens persuasivas geradas por um LLM, ajustadas ao traço Abertura à Experiência, produziriam atitudes mais favoráveis em relação ao uso desses sistemas em comparação com uma mensagem genérica não personalizada. Também foram examinados o papel da Abertura à Tecnologia e da área de formação, particularmente a diferença entre estudantes de Psicologia e de outros cursos, na aceitação dos *chatbots*.

Em conjunto, os dois estudos ajudam a compreender a persuasão personalizada mediada por IA ao articular a revisão das evidências recentes com um teste empírico conduzido no contexto da saúde mental digital.

Artigo I

Persuasão Personalizada Mediada por Inteligência Artificial: uma Revisão Semi-Sistemática das Evidências Empíricas

AI-Mediated Personalized Persuasion: A Semi-Systematic Review of Empirical Evidence

Persuasión Personalizada Mediada por Inteligencia Artificial: una Revisión Semisistemática de la Evidencia Empírica

Bruno Grossman

Heder Soares Bernardino

Jairo Francisco de Souza

Leonardo Fernandes Martins

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil

Resumo

A persuasão personalizada tem se mostrado uma estratégia eficaz para promover mudanças de atitude e comportamento em diferentes domínios, com resultados consistentes em áreas como saúde, política e consumo. O avanço recente da inteligência artificial, especialmente dos modelos de linguagem de grande escala, ampliou o alcance dessa abordagem ao permitir a geração automatizada de mensagens adaptadas ao perfil psicológico de cada indivíduo. Isso despertou interesse crescente pelo potencial de favorecer o engajamento e a adesão em aplicações de bem-estar psicológico e intervenções digitais em saúde mental. Os estudos recentes sobre o tema estão dispersos entre diferentes domínios. Esta revisão semi-sistemática reuniu evidências empíricas sobre a persuasão personalizada mediada por inteligência artificial, buscando identificar seus contextos de aplicação, variáveis psicológicas mais relevantes e principais limitações. Foram analisados oito artigos empíricos (n= 26.584), publicados entre 2023 e 2025. Os resultados mostraram que 66% dos testes de eficácia apresentaram efeitos significativos, em sua maioria na direção esperada. As variáveis psicológicas apresentaram maior eficácia (70%) que as demográficas (58%), com destaque para Abertura à Experiência e Extroversão, cujos efeitos se mantiveram mais estáveis entre domínios e, em geral, foram de magnitude pequena a moderada. No domínio da saúde, 86% dos testes indicaram efeitos positivos significativos. Na área da política, mensagens alinhadas a valores morais e à ideologia política também mostraram efeitos significativos, embora pequenos (β= 0,17 a 0,25). De forma geral, considerando os três domínios, as magnitudes dos efeitos foram pequenas a moderadas (β= 0,16 a 0,40). As limitações mais frequentes incluíram amostras geograficamente concentradas, predomínio de medidas de autorrelato, variação nas técnicas de engenharia de prompt e escassez de estudos no contexto de saúde mental. Em conjunto, as evidências indicam que a personalização mediada por IA é mais eficaz quando se baseia em atributos psicologicamente relevantes, configurando um campo em expansão ainda carente de estudos em saúde mental.

Palavras-chave

Persuasão personalizada; inteligência artificial; personalidade; saúde mental digital

Abstract

Personalized persuasion has proven to be an effective strategy for promoting attitudinal and behavioral change across different domains, with consistent findings in areas such as health, politics, and consumption. Recent advances in artificial intelligence, particularly large language models, have expanded the scope of this approach by enabling the automated generation of messages tailored to individuals' psychological profiles. This has sparked growing interest in its potential to enhance engagement and adherence in applications related to psychological well-being and digital mental health interventions. Recent studies on the topic remain dispersed across different domains.

This semi-systematic review synthesized empirical evidence on AI-mediated personalized persuasion, aiming to identify its application contexts, the most relevant psychological variables, and key limitations. Eight empirical articles were analyzed (n= 26,584), published between 2023 and 2025. The results showed that 66% of the efficacy tests reported significant effects, mostly in the expected direction. Psychological variables demonstrated higher effectiveness (70%) than demographic variables (58%), with Openness to Experience and Extraversion standing out for their more stable effects across domains, generally of small to moderate magnitude.

In the health domain, 86% of the tests indicated significant positive effects. In politics, messages aligned with moral values and political ideology also yielded significant, albeit small, effects (β= 0.17 to 0.25). Overall, across the three domains, effect sizes ranged from small to moderate (β= 0.16 to 0.40). The most frequent limitations included geographically concentrated samples, a predominance of self-report measures, variation in prompt-engineering techniques, and a scarcity of studies in digital mental health contexts. Taken together, the evidence suggests that AI-mediated personalization is more effective when grounded in psychologically relevant attributes, pointing to a growing field that still lacks sufficient research in mental health.

Keywords

Personalized persuasion; artificial intelligence; personality; digital mental health

Segundo a OMS (2025), mais de um bilhão de pessoas vivem com algum transtorno mental, o que representa aproximadamente 14% da população mundial. Entre eles, os transtornos de ansiedade e depressivos são os mais prevalentes, respondendo juntos por mais de dois terços dos casos. A dimensão desse cenário estimulou o interesse por abordagens comunicacionais capazes de favorecer atitudes e comportamentos voltados ao cuidado psicológico, à adesão a intervenções digitais e ao engajamento em práticas de bem-estar (Oakley-Girvan et al., 2021).

A persuasão, definida como o processo de influenciar crenças, atitudes ou comportamentos por meio da comunicação (Luttrell et al., 2025), constitui um dos fenômenos mais estudados para compreender como promover mudanças sustentadas (Druckman, 2022). Nesse campo, consolidou-se o conceito de persuasão personalizada, entendido como a adaptação deliberada de elementos comunicacionais, como a mensagem, a fonte ou o contexto, a uma ou mais características do destinatário, incluindo traços psicológicos, motivações, valores, identidades ou outros atributos individuais, com o objetivo de aumentar a relevância e a eficácia do apelo (Matz et al., 2017; Teeny et al., 2021).

Na literatura, essa estratégia aparece de formas distintas, como correspondência personalizada (*personalized matching*), adaptação à personalidade (*personality tailoring*), segmentação psicológica (*psychological targeting*), comunicação direcionada (*targeting, microtargeting*) e correspondência motivacional (*motivational matching*). Apesar de haver nuances entre esses termos, todos remetem ao mesmo princípio: explorar como conteúdos tornam-se mais eficazes quando se alinham diretamente com características individuais ou grupais do público, como destacam Matz et al. (2023) e Teeny et al. (2021).

O avanço recente da inteligência artificial, especialmente dos modelos de linguagem de grande escala (*large language models,* LLMs), ampliou essa possibilidade ao permitir a geração automatizada de conteúdos ajustados a perfis psicológicos individuais (Matz et al., 2024), o que abriu possibilidades para testar empiricamente a eficácia e as limitações dessa estratégia em diferentes domínios de aplicação. Modelos de linguagem de grande escala são sistemas de aprendizado profundo treinados em volumes massivos de texto, capazes de processar e gerar linguagem natural em múltiplas tarefas (IBM, 2023).

Na Retórica (Aristotle; Reeve, 2018), Aristóteles já observava, no século IV a.C., que a persuasão era mais eficaz quando ajustava seu apelo ao que era relevante para o ouvinte. Essa intuição antiga recebeu respaldo teórico (Petty et al., 2024) e empírico nas últimas décadas (Teeny et al., 2021), à medida que estudos demonstraram que mensagens cujo conteúdo é ajustado ao perfil psicológico do destinatário tendem a ser mais eficazes do que comunicações não personalizadas (Matz et al., 2024). A revisão de Teeny e colegas sintetizou um corpo de evidências, mostrando que a comunicação personalizada pode aumentar significativamente a persuasão, refletida em mudanças de atitudes, intenções e comportamentos, em diferentes domínios como política, consumo e saúde, a partir do uso de variáveis psicológicas, como traços de personalidade, valores morais e motivações regulatórias.

De modo semelhante, pesquisas aplicadas mostraram que mensagens congruentes a características motivacionais e disposicionais favorecem comportamentos desejáveis, como a prática de atividade física (Latimer et al., 2008), a adoção de hábitos alimentares saudáveis (Latimer et al., 2005) e o engajamento em ações de proteção à saúde (Williams-Piehota et al., 2003). Esses achados consolidaram a persuasão personalizada como uma estratégia promissora para promover mudanças positivas em atitudes e comportamentos.

O Modelo da Probabilidade de Elaboração (Elaboration Likelihood Model, ELM; Petty; Cacioppo, 1986) é uma referência central para entender como mensagens persuasivas levam à mudança de atitude. O modelo descreve que o processamento pode ocorrer com maior ou menor profundidade, variando entre uma rota central, marcada pela análise atenta dos argumentos, e uma rota periférica, baseada em pistas simples ou associações afetivas. O quanto o indivíduo elabora a mensagem depende de sua motivação e de sua capacidade de processar o conteúdo. Atitudes formadas predominantemente pela rota central tendem a ser mais estáveis, mais resistentes a contra-argumentos e mais relacionadas ao comportamento do que aquelas originadas a partir de sinais periféricos.

A revisão de Teeny et al. (2021) mostrou que os efeitos de congruência entre a mensagem e as características psicológicas do receptor podem ser explicados à luz desse modelo. A personalização pode atuar como pista de afinidade em contextos de baixa elaboração ou como argumento relevante quando o processamento é mais profundo. Em níveis elevados de reflexão, pode ainda

influenciar processos metacognitivos de validação e julgamento. Contudo, a personalização não garante maior eficácia: seu impacto depende de como o indivíduo interpreta a adequação entre a mensagem e suas próprias disposições, podendo gerar aceitação, indiferença ou resistência. Assim, o ELM oferece um quadro teórico útil para entender quando e por que a personalização de mensagens aumenta ou reduz a persuasão.

A personalização de mensagens é entendida como um contínuo, no qual os conteúdos variam do alinhamento positivo (*active match*) ao desalinhamento ativo (*active mismatch*), isto é, situações em que a mensagem toca diretamente um aspecto motivacional do receptor de forma contrária ao seu perfil (por exemplo, enfatizar disciplina rígida para alguém que valoriza autonomia em saúde mental). Entre esses extremos situam-se as mensagens neutras (*non-match* ou *passive mismatch*), que não apoiam nem contradizem essas disposições. Como mensagens alinhadas e desalinhadas ativam processos motivacionais distintos, comparações diretas entre as duas podem misturar efeitos diferentes e dificultar a interpretação. A noção de contínuo ajuda a organizar essas diferenças, pois qualquer mensagem persuasiva ocupa algum ponto entre o desalinhamento ativo e o alinhamento, permitindo avaliar a personalização de modo mais preciso (Joyal-Desmarais et al., 2025).

Teeny e colegas (2021) classificaram as variáveis de correspondência em cinco grupos: estados afetivos e cognitivos, objetivos e orientações motivacionais, bases e funções de atitudes, identidades e personalidades, e orientação cultural. Dentro desses grupos, os autores descreveram vinte e duas variáveis específicas empregadas para personalizar mensagens com eficácia. Entre as mais recorrentes, destacaram-se os traços do Modelo dos Cinco Grandes Fatores (*Big Five*; McCrae; John, 1992), em especial Abertura à Experiência e Extroversão, que apresentaram efeitos estáveis na eficácia da personalização (Matz et al., 2017; Zarouali et al., 2022).

A persuasão personalizada é descrita como um processo composto por duas etapas principais: o perfilamento psicológico e a adaptação da intervenção. Na primeira, busca-se identificar características individuais relevantes ao processamento persuasivo; na segunda, essas informações orientam a construção de comunicações ajustadas às motivações, valores ou disposições do público, aumentando sua relevância e impacto (Matz et al., 2020).

Por muito tempo, essa estratégia permaneceu limitada por barreiras metodológicas, uma vez que a mensuração de traços psicológicos dependia de métodos tradicionais, como questionários e entrevistas, lentos, caros e sujeitos a vieses de resposta (Podsakoff; Organ, 1986). O caso Cambridge Analytica evidenciou o poder e os riscos associados ao uso de dados pessoais para moldar mensagens direcionadas (Simchon et al., 2024), reforçando a importância de critérios éticos e transparentes no uso de informações psicológicas para fins persuasivos.

Nos últimos anos, modelos de aprendizado de máquina e o uso crescente de plataformas digitais ampliaram as formas de analisar rastros deixados por usuários. Pesquisas mostram que informações simples, como *likes* em redes sociais, podem ser usadas para estimar traços de personalidade com boa precisão (Youyou et al., 2015), além de prever características como orientação política e etnia (Kosinski et al., 2013). Também há evidências de que imagens faciais *online* permitem inferir orientação sexual com precisão superior à de julgamentos humanos (Wang; Kosinski, 2018). Esses resultados revelam o alcance técnico dessas ferramentas e evidenciam os riscos envolvidos quando lidam com dados sensíveis.

Nesse contexto, o campo da psicometria computacional ganhou força ao aproximar a ciência de dados da avaliação psicológica, permitindo estimar características individuais de forma automatizada (Franco; Primi, 2022). Como observaram Teeny et al. (2021) e Petty et al. (2024) a expansão digital e o avanço dos algoritmos fizeram da personalização de mensagens um tema cada vez mais presente na pesquisa atual. Com o surgimento dos modelos de linguagem de grande escala, esse movimento deu um passo adiante, abrindo espaço para o que Rogiers et al. (2024) denominam de Persuasão Mediada por LLM, uma subárea emergente dedicada a entender os efeitos persuasivos mediados por sistemas de IA generativa.

Diante desse avanço conceitual e tecnológico, torna-se essencial compreender o que as evidências empíricas indicam sobre a eficácia e as limitações da persuasão personalizada mediada por inteligência artificial. Apesar do interesse crescente, os estudos disponíveis ainda são recentes e heterogêneos, conduzidos em diferentes contextos e com métodos variados. Reunir essas evidências é fundamental para avaliar o alcance real da persuasão personalizada e seu potencial de aplicação em domínios de interesse social, como a promoção de comportamentos saudáveis e o apoio à saúde mental. Esta revisão teve, portanto, o objetivo de examinar

criticamente as pesquisas empíricas recentes sobre o tema, identificando seus contextos de uso, variáveis analisadas, resultados e lacunas que ainda limitam o avanço do campo.

1.
Objetivos

1.1
Objetivo Geral

O objetivo geral foi revisar e sintetizar as evidências empíricas disponíveis sobre a persuasão personalizada mediada por inteligência artificial, a fim de compreender em que contextos essa abordagem havia sido aplicada, quais variáveis psicológicas sustentavam sua eficácia e quais limitações permaneciam nas pesquisas existentes.

1.2
Objetivos Específicos

De modo específico, o estudo buscou:

a) Mapear os estudos empíricos que investigaram mensagens persuasivas personalizadas por inteligência artificial, descrevendo seus contextos de aplicação e o perfil das amostras analisadas, com o propósito de descrever o panorama atual da pesquisa sobre o tema.

b) Sintetizar as variáveis psicológicas utilizadas como base de personalização e os resultados obtidos, identificando quais características individuais se associaram com maior consistência aos efeitos persuasivos observados.

c) Comparar mensagens personalizadas e não personalizadas (neutras, incongruentes ou outras condições) quanto à eficácia em promover mudanças de atitude, intenção ou comportamento, destacando os padrões de resultado observados entre estudos com diferentes delineamentos experimentais.

2.
Método

A revisão foi conduzida no formato semi-sistemático, conforme as orientações metodológicas de Snyder (2019). Esse tipo de revisão é indicado para

temas emergentes e interdisciplinares, nos quais ainda não há padronização metodológica nem volume suficiente de estudos para uma abordagem sistemática. Esse formato foi adequado ao campo da persuasão personalizada mediada por inteligência artificial, que é recente, reúne pesquisas de áreas distintas e apresenta grande diversidade de desenhos empíricos. A revisão reuniu estudos experimentais publicados entre 2020 e 2025 e buscou sintetizar as evidências disponíveis. O processo seguiu etapas de busca, triagem e análise baseadas nas recomendações do protocolo PRISMA (Liberati et al., 2009), adaptadas ao caráter descritivo e interpretativo do estudo.

2.1 Estratégia de busca

a) Bases de dados de artigos científicos e periódicos acadêmicos

Para localizar estudos revisados por pares, foram consultadas bases de dados internacionais nas áreas de Psicologia, Comunicação e Ciências da Computação. A busca também incluiu periódicos de alto impacto e repositórios abertos que reúnem *preprints* e manuscritos recentes sobre inteligência artificial, utilizados apenas para acompanhar versões atualizadas de trabalhos em processo de publicação. As fontes utilizadas estão apresentadas na Tabela 1.

Tabela 1 – Bases de Dados, Periódicos, Repositórios e Fontes Complementares Utilizadas.

Categorias	**Fontes**
Bases de Dados de Artigos Científicos	APA PsycINFO, Scopus, Web of Science, PubMed, ACM Digital Library, IEEE Xplore, ScienceDirect, SpringerLink
Periódicos Acessados	Nature, Nature Human Behaviour, Journal of Experimental Political Science, Journal of Experimental Social Psychology, PNAS, Annual Review, Journal of Personality and Social Psychology, Computers in Human Behavior, Frontiers in Psychology, Frontiers in Artificial Intelligence
Repositórios de Pré-publicação	PsyArXiv, ArXiv
Portais Multidisciplinares e Institucionais	Portal CAPES

b) Expressões de busca

As estratégias de busca foram elaboradas para abranger os principais termos relacionados à persuasão personalizada e ao uso de inteligência artificial, incluindo modelos de linguagem de grande escala (LLMs) e agentes conversacionais.

Inicialmente, foram testadas buscas em português, mas não foram localizados estudos relevantes. Por essa razão, a busca principal foi realizada em inglês, garantindo maior abrangência de publicações recentes e representativas do campo.

Os termos foram combinados com operadores booleanos e ajustados para incluir expressões comuns em diferentes abordagens de personalização psicológica e tecnológica, como *matching*, *fit* e *congruence*. As expressões finais utilizadas estão apresentadas na Tabela 2.

Tabela 2 – Expressões de busca.

Termos gerais	"persuasion", "AI", "GPT", "LLM", "tailored", "personalized"
Busca ampliada – Personalização e IA	("personalized" or "tailored" or "matching" or "message fit" or "congruence" or "alignment") and ("persuasion" or "persuasiveness" or "motivational" or "regulatory focus") and ("artificial intelligence" or "AI" or "AI agent" or "conversational agent" or "chatbot" or "virtual assistant" or "large language model" or "LLM" or "Generative AI") and ("attitude" or "behavior" or "mental health" or "well-being" or "health communication")
Busca complementar	tailored or personalized or persuasion or persuasiveness and artificial intelligence or AI or AI agent or chatbot or large language model or LLM or Generative AI or mental health or well-being

c) Critérios de inclusão e exclusão

Os critérios de inclusão e exclusão foram definidos para garantir a adequação dos estudos ao tema da revisão e a consistência metodológica da amostra final. Foram considerados apenas trabalhos empíricos que testaram, de forma experimental, a eficácia persuasiva de mensagens personalizadas mediadas por inteligência artificial.

Tabela 3 – Critérios de inclusão e exclusão.

Critérios de inclusão	1. Publicações empíricas, revisadas por pares, disponíveis na íntegra e publicadas entre janeiro de 2020 e maio 2025, em inglês ou português. 2. Estudos que utilizaram sistemas de inteligência artificial na geração, adaptação ou mediação de mensagens persuasivas. 3. Medidas de eficácia persuasiva: os estudos incluídos avaliaram a eficácia persuasiva com base em pelo menos uma das seguintes variáveis dependentes: a) mudanças de atitude ou opinião; b) intenções comportamentais (por exemplo, disposição para doar, adotar ou utilizar um serviço); c) comportamentos observáveis, como escolhas simuladas,

	cliques ou outras respostas indicativas de persuasão; d) percepção de persuasão ou eficácia da mensagem. 4. Delineamentos experimentais conduzidos com participantes humanos. 5. Qualidade metodológica: os estudos deveriam apresentar: a) presença de grupo controle; b) randomização adequada; c) tamanho de amostra adequado ao delineamento e à análise proposta; d) metodologia claramente descrita e passível de replicação; e) análise estatística apropriada; f) discussão das limitações do estudo.
Critérios de exclusão	1. Revisões de literatura, ensaios teóricos e comentários conceituais. 2. Estudos que não empregaram inteligência artificial no processo de personalização das mensagens. 3. Estudos que avaliaram a eficácia persuasiva de mensagens geradas por IA, mas sem qualquer processo de personalização, mediação ou adaptação ao perfil dos receptores. 4. Estudos que analisaram personalização por meios tradicionais ou tecnologias não baseadas em IA, como vídeos editados manualmente ou segmentação estática. 5. Trabalhos que não apresentaram medidas de eficácia persuasiva. 6. Relatos técnicos, dissertações, teses e outros materiais não revisados por pares. 7. Publicações duplicadas ou incompletas.

Esses critérios foram aplicados de forma sequencial, garantindo que apenas estudos alinhados aos objetivos da revisão fossem incluídos na análise final.

d) Seleção, extração e avaliação da qualidade dos estudos

A seleção dos estudos seguiu as etapas gerais do protocolo PRISMA, adaptadas ao formato semi-sistemático da revisão. Após a busca nas bases, as referências foram organizadas e as duplicatas removidas. Os registros foram examinados por título e resumo, e os textos potencialmente relevantes foram lidos na íntegra. A elegibilidade foi verificada conforme os critérios previamente estabelecidos.

Os estudos que atenderam a esses critérios tiveram seus dados extraídos e organizados em planilhas, registrando informações sobre domínio de aplicação, amostra, delineamento, modelo de IA, variáveis psicológicas utilizadas e principais resultados. A análise foi descritiva e comparativa, buscando identificar padrões,

convergências e lacunas entre os trabalhos. Os achados foram sintetizados de forma narrativa e agrupados em categorias temáticas.

A qualidade metodológica foi avaliada com os mesmos itens definidos nos critérios (p.ex., grupo controle, randomização, tamanho amostral, clareza metodológica, análise estatística adequada e discussão de limitações) e registrada para qualificar a interpretação dos resultados, sem aplicar filtros adicionais nesta etapa.

3. Resultados

3.1 Panorama geral dos estudos incluídos

A revisão reuniu oito artigos empíricos, totalizando vinte e nove subestudos publicados entre 2023 e 2025, com 26.584 participantes. As amostras foram organizadas por domínio de aplicação (política, consumo e saúde). Um subestudo apresentou contexto combinado de consumo e política (n= 320) e foi mantido separado para evitar dupla contagem. Os totais finais, sem sobreposição entre categorias, foram: política (n= 22.965), consumo (n= 1.672), saúde (n= 1.627) e consumo-política (n= 320).

A produção é recente, com dois estudos em 2023, quatro em 2024 e dois em 2025. Todos foram conduzidos nos Estados Unidos ou no Reino Unido e adotaram desenhos experimentais voltados a avaliar a eficácia de mensagens personalizadas mediadas por inteligência artificial.

O número total de subestudos corresponde ao total de análises experimentais independentes relatadas nos oito artigos incluídos. Cada subestudo foi contado separadamente sempre que apresentou delineamento e comparação próprios entre condições personalizadas e não personalizadas, ainda que compartilhasse a mesma amostra principal. Essa opção metodológica foi adotada para garantir que todos os testes de eficácia relatados fossem representados na síntese, mesmo quando diferentes manipulações experimentais estavam reunidas em um único artigo.

Foram contabilizados: Matz et al. (2024), com dez subestudos (1a, 1b, 2a–2c, 3a–3c e 4a–4b); Simchon et al. (2024), com quatro subestudos (1a, 1b, 2a e 2b); Hackenburg; Margetts (2024), com quatro subestudos; Hackenburg et al. (2023) com quatro subestudos; Tappin et al. (2023), com três subestudos; Meguellati et al.

(2024), com dois subestudos; e Lu (2025) e Salvi et al. (2025), com um subestudo cada, totalizando 29.

Os subestudos se distribuíram em três domínios principais. Dezoito trataram de comunicação política, abordando temas como mudança climática, imigração, participação eleitoral, privacidade digital e polarização ideológica. Oito examinaram o consumo, incluindo mensagens sobre produtos como tênis esportivos e *smartphones*. Dois investigaram saúde, um voltado à vacinação e desinformação (Lu, 2025) e outro à promoção de comportamentos saudáveis e atividade física (Matz et al., 2024).

Tabela 4 – Artigos e Subestudos que Compõem a Amostra da Revisão.

Autores	Subestudos	Amostra	Domínio	Tipo de Personalização	Variáveis de Personalização	Técnica/ Modelo de IA
Tappin et al., 2023	3	5284	Política	Demográfica e/ou psicológica	Demográficas: Partidarismo, idade, renda, gênero, educação, etnia Psicológicas: Valores morais, religiosidade, conhecimento político, reflexão cognitiva, ideologia política	Aprendizado de máquina supervisionado
Hackenburg et al., 2023	4	4955	Política	Demográfica	Partidarismo	GPT-4
Matz et al., 2024	10	2172	Consumo, política e saúde	Psicológica	Big Five, Foco regulatório (promoção), Fundações Morais, Ideologia política	GPT-3 e ChatGPT 3.5 Turbo
Hackenburg & Margetts, 2024	4	8587	Política	Demográfica	Idade, gênero, escolaridade, etnia, ocupação, religiosidade, residência, orientação ideológica, engajamento político	GPT-4
Simchon et al. 2024	4	440	Política	Psicológica	Abertura à Experiência	Inferência preditiva, GPT-3 e ChatGPT 3.5 Turbo
Meguellati et al., 2024	2	400	Consumo	Psicológica	Abertura à Experiência e Neuroticismo	GPT-3.5
Lu, 2025	1	1435	Saúde	Psicológica	Extroversão e Crenças pseudocientíficas	GPT-4o
Salvi et al., 2025	1	900	Política	Demográfica	Afiliação política	GPT-4

As amostras foram formadas, em geral, por adultos recrutados *online*, principalmente em plataformas pagas como Lucid Theorem, Prolific e Mechanical Turk. A maioria dos delineamentos comparou condições personalizadas e não personalizadas (neutras, incongruentes ou de autoria humana). O ChatGPT-3.5-turbo e o GPT-4/4o foram empregados em 66% dos subestudos, enquanto o GPT-3 apareceu apenas nos estudos mais antigos. Em todos os trabalhos, a personalização foi conduzida de modo totalmente automatizado pela inteligência artificial, sem qualquer forma de revisão ou curadoria humana. Dois estudos utilizaram abordagens preditivas em vez de modelos generativos: Tappin et al. (2023) que identificou, por meio de algoritmos de aprendizado de máquina, quais

mensagens seriam mais eficazes para cada perfil, e Simchon et al. (2024), que estimou traços psicológicos dos participantes para calcular a correspondência entre perfis e mensagens. Nesse estudo, os autores também validaram os estímulos com o modelo aberto Llama 2 70B, usado apenas para verificar a consistência dos padrões gerados pelo ChatGPT.

No conjunto, as aplicações de persuasão personalizada mediada por IA ainda se concentram em contextos políticos e comerciais, enquanto os estudos em saúde são pontuais. Mesmo assim, o uso crescente de modelos generativos e algoritmos preditivos mostra que a área vem avançando rapidamente em sofisticação técnica, criando condições para que abordagens semelhantes sejam testadas em contextos de saúde mental digital.

3.2 – Variáveis psicológicas e estratégias de personalização

Nos estudos revisados, a personalização baseada em variáveis psicológicas foi a abordagem predominante, presente em 59% (17) dos subestudos. Nove utilizaram variáveis demográficas, e três combinaram as duas naturezas de dados (Tappin et al., 2023). Em geral, os trabalhos se concentraram em entender como diferenças psicológicas individuais influenciam a recepção de mensagens geradas por inteligência artificial.

Entre as variáveis psicológicas, os traços de personalidade foram os mais recorrentes, empregados em dez subestudos (34%), todos baseados no *Big Five*. Os traços mais explorados foram Extroversão (10 ocorrências) e Abertura à Experiência (9), seguidos de Amabilidade (3), Conscienciosidade (3), Neuroticismo (2) e traços dominantes do Big Five (3). Esses construtos orientaram o ajuste de enquadramento, o tom e o conteúdo das mensagens, buscando maior congruência entre o perfil do público e a estratégia persuasiva.

Além dos traços de personalidade, outros construtos psicológicos também foram empregados, ainda que com menor frequência. Destacaram-se, no contexto da política, o uso da Ideologia política (6) Fundações Morais (5) e da Reflexão cognitiva (1); e, no domínio da saúde, das Crenças pseudocientíficas (2) e do Foco regulatório, com ênfase na promoção (1).

A maior parte das variáveis psicológicas foi tratada como contínua, a partir de escores obtidos em escalas padronizadas; em alguns subestudos, porém, esses

escores foram categorizados e integrados a modelos de regressão ou a algoritmos preditivos de eficácia persuasiva. O número total de variáveis analisadas ultrapassou o número de estudos revisados, já que alguns trabalhos combinaram mais de uma variável no mesmo delineamento, como personalidade e valores morais.

As escalas empregadas foram breves e validadas, adequadas às exigências de experimentos *online*. Em Matz et al. (2024), os traços do Big Five foram avaliados pelo BFI-2-S (30 itens), enquanto outros construtos foram mensurados pelo MFQ-30 (Questionário de Fundações Morais), por uma versão adaptada da escala de Foco Regulatório e por uma escala de ideologia política de 7 pontos. Meguellati et al. (2024) utilizaram o SAPA Personality Inventory (20 itens). Simchon et al. (2024) adotaram o BFI-2 conforme Soto e John (2017). Tappin et al. (2023) empregaram o MFQ, além de testes de reflexão cognitiva e de conhecimento político. Lu (2025) avaliou Extroversão pelo BFI-2 e crenças pseudocientíficas pela Revised Pseudoscientific Belief Scale.

Em três estudos, porém, a personalização não se baseou em traços psicológicos: Salvi et al. (2025) não mediram qualquer variável psicológica e Hackenburg e Margetts (2024) utilizaram apenas atributos sociodemográficos. No estudo de Hackenburg et al. (2023), a filiação partidária foi tratada como variável demográfica, por ser apenas uma autodeclaração de identificação usada para balancear a amostra e definir a congruência das mensagens.

A eficácia persuasiva foi avaliada principalmente por medidas de autorrelato, incluindo atitudes, apoio às mensagens, intenções comportamentais e percepção de persuasão, além de alguns indicadores indiretos de comportamento. Predominaram escalas Likert de 5 ou 7 pontos, escalas bipolares de 0 a 10 ou 11 e, em menor número, formatos contínuos de 0 a 100. Em Matz et al. (2024), as avaliações incluíram interesse pelo produto, efetividade percebida e medidas monetárias de disposição a pagar ou doar. Em Meguellati et al. (2024), atitudes e intenção de compra foram complementadas por escolhas entre anúncios em um ambiente de navegação simulado, um dos poucos casos que se aproximaram de um desfecho comportamental.

A percepção de persuasão também apareceu como desfecho em parte dos estudos, especialmente em Matz et al. (2024) e Simchon et al. (2024), que utilizaram escalas para captar a impressão de impacto das mensagens. Os momentos

de aplicação foram bastante homogêneos, com avaliações concentradas no pós-estímulo e apenas um trabalho em formato pré e pós-teste (Salvi et al., 2025). Nenhum estudo incluiu acompanhamento posterior.

Alguns estudos verificaram a percepção da autoria das mensagens. Em Salvi et al. (2025), após os debates, a identificação de IA foi correta em cerca de 75%, enquanto a identificação de humanos permaneceu próxima ao acaso. Em Hackenburg e Margetts (2024), 17% dos participantes na condição de microdirecionamento classificaram a mensagem como produzida por IA, em comparação com 14% na condição de melhor mensagem, indicando um aumento pequeno na detecção da autoria artificial. Em Hackenburg et al. (2023), cerca de 23% dos participantes atribuíram as mensagens à IA mesmo quando eram produzidas por consultores humanos, mostrando que parte dos respondentes concluiu que qualquer conteúdo político poderia ter origem artificial. Em Matz et al. (2024), apenas o Estudo 1b informou previamente que os anúncios haviam sido gerados por IA e poderiam estar personalizados, mas essa informação não alterou as avaliações de eficácia. Os demais estudos não relataram testes formais de suspeita.

Em geral, os resultados indicaram que a personalização psicológica foi o foco das pesquisas revisadas, com destaque para os traços de personalidade do modelo *Big Five* e para variáveis morais e ideológicas. O uso combinado de modelos generativos e preditivos representou um avanço técnico importante e abre novas possibilidades para automatizar o ajuste de mensagens ao perfil psicológico de cada indivíduo.

3.3 Comparação da eficácia das mensagens personalizadas e genéricas

A avaliação reuniu um total de 56 testes de eficácia da personalização, conduzidos nos 29 subestudos incluídos. Esse número é superior ao total de subestudos porque vários experimentos avaliaram mais de uma variável psicológica ou demográfica dentro do mesmo delineamento, testando separadamente o efeito de diferentes traços sobre as medidas de persuasão. Em alguns casos, esses testes foram gerados a partir da mesma amostra.

Do total de 56 testes, 37 apresentaram resultados estatisticamente significativos ($p \leq 0,05$), o que corresponde a 66% das análises realizadas. Os

demais 19 testes (34%) não mostraram diferença significativa entre condições personalizadas e não personalizadas (neutras, incongruentes, de controle ou humanas) em diferentes medidas de eficácia, incluindo atitudes, intenções e respostas comportamentais.

Os padrões de significância foram semelhantes entre os domínios de consumo e política, com 70% e 68% de resultados significativos, respectivamente, enquanto o domínio da saúde apresentou 86% de efeitos estatisticamente significativos, ainda que sustentados por amostras menores (n= 1.627). Esse padrão se mostrou persistente nos dois estudos de saúde incluídos, embora o número reduzido de participantes impeça generalizações.

Entre os 37 testes com significância estatística ($p \leq 0,05$), 34 (92%) mostraram efeitos na direção esperada, indicando maior eficácia das mensagens personalizadas em comparação às condições de controle. Três testes mostraram efeito inverso, em contextos específicos. Em Lu (2025), o traço de crenças pseudocientíficas gerou um efeito de retrocesso ($\beta= 0,22$; $p= 0,010$), em que participantes com crenças mais elevadas reagiram negativamente às mensagens personalizadas sobre vacinação. Em Hackenburg et al. (2023), mensagens de GPT-4 com simulação (*role-play*) alinhado produziram efeito significativo na direção contrária ao argumento pró-mandato (–5,94 p.p.; $p< 0,001$). Já em Hackenburg e Margetts (2024), a condição de personalização incorreta (*false targeting*), que utilizou atributos trocados de forma intencional, apresentou efeito inverso, sendo menos eficaz que a mensagem genérica ($p= 0,016$). Esse resultado, esperado pela manipulação, indicou que personalizações incongruentes podem reduzir a eficácia persuasiva e gerar reações negativas.

Os resultados por tipo de personalização mostraram maior consistência para variáveis psicológicas, com 70% dos testes apresentando significância estatística (30 de 43). A personalização demográfica obteve 58% de resultados significativos (7 de 12), enquanto a híbrida, que combinou informações psicológicas e demográficas, não apresentou efeitos significativos. Apesar das diferenças numéricas, o padrão geral sugere que abordagens baseadas em características psicológicas tendem a gerar efeitos mais claros do que aquelas apoiadas apenas em dados demográficos.

Extroversão apresentou efeitos em consumo (5 testes) e saúde (4), enquanto Abertura mostrou resultados em consumo (5) e política (3). Os dois traços

demonstraram consistência entre diferentes domínios, embora com padrões de generalização distintos. Ideologia política também apresentou quatro efeitos significativos e dois nulos, enquanto Conscienciosidade teve dois efeitos significativos e um nulo. Crenças pseudocientíficas registraram dois resultados significativos, e Justiça, Lealdade, Autoridade e o traço dominante do Big Five mostraram um efeito significativo cada. Por outro lado, Amabilidade (0/3), Neuroticismo (0/2), Cuidado (0/1), Foco regulatório (0/1) e Pureza (0/1) não apresentaram significância estatística.

O uso de LLMs apresentou 29 resultados significativos, frente a 8 das abordagens preditivas e algorítmicas. Essa diferença reflete maior variedade de contextos avaliados e sugere que os modelos generativos mantiveram desempenho consistente em diferentes aplicações, com resultados mais representativos mesmo com percentuais próximos (69% e 57%).

Quanto aos tamanhos de efeito reportados, 34 testes apresentaram medidas padronizadas, o que permitiu avaliar a magnitude das diferenças observadas entre condições. Em Matz et al. (2024), os coeficientes β variaram entre 0,16 e 0,40, correspondendo a efeitos pequenos a moderados segundo os parâmetros de Cohen (1992), que considera valores de 0,10, 0,30 e 0,50 como limiares aproximados para efeitos pequeno, médio e grande, respectivamente. O estudo de Lu (2025) apresentou β entre 0,22 e 0,30 e η^2p= 0,075, valores que também indicam efeitos de magnitude média, próximos ao ponto de referência de 0,06 estabelecido por Cohen para esse índice. Assim como em Matz, as magnitudes observadas sugerem que os efeitos da personalização são estatisticamente convergentes, mas de amplitude moderada, refletindo mudanças sutis nas atitudes ou percepções dos participantes.

Entre os traços psicológicos analisados, Abertura à Experiência e Extroversão se destacaram por apresentarem os maiores coeficientes e maior estabilidade entre contextos. Abertura mostrou efeitos moderados nos domínios de consumo e política, enquanto Extroversão exibiu o maior valor individual (β= 0,40), especialmente em experimentos voltados ao consumo (anúncios de *iPhone*). Outros construtos, como Conscienciosidade (β= 0,20–0,29) e Justiça (β= 0,25), apresentaram efeitos intermediários, enquanto Ideologia política (β= 0,08–0,23) e Foco regulatório (β= 0,12) mostraram magnitudes pequenas.

Ao comparar os domínios, os estudos de consumo apresentaram a maior amplitude de efeitos (β= 0,16-0,40), com destaque para Extroversão (β= 0,40) e Abertura à Experiência (β= 0,36). Nos estudos políticos, os efeitos foram pequenos a moderados (β entre 0,17 e 0,25), o que pode refletir a influência de crenças prévias mais resistentes à mudança. Já no domínio da saúde, os efeitos foram moderados e recorrentes (β= 0,22–0,30; η^2p= 0,075), sugerindo que a personalização também é eficaz em temas informativos quando associada a traços relevantes. Em conjunto, os resultados indicam que a magnitude dos efeitos varia conforme o envolvimento com o tema, sendo mais alta em consumo, intermediária em política e estável em saúde.

4. Discussão

Esta revisão semi-sistemática teve como objetivo reunir e analisar as evidências empíricas sobre a persuasão personalizada mediada por inteligência artificial, em um campo ainda recente e em rápida expansão. O estudo abrangeu diferentes estratégias de personalização automatizada, com uso de modelos generativos e preditivos aplicados a variados domínios, como consumo, política e saúde. Foram considerados delineamentos experimentais distintos e medidas de eficácia que incluíram atitudes, intenções e comportamentos, baseadas em variáveis de natureza psicológica e demográfica. Essa abordagem ampla permitiu identificar padrões regulares entre diferentes contextos e apontar lacunas, especialmente a ausência de estudos que avaliaram essa estratégia no contexto de aplicações voltadas ao bem-estar psicológico e a intervenções digitais de saúde mental.

Os resultados desta revisão indicaram um padrão geral próximo ao descrito na literatura mais ampla sobre persuasão personalizada, embora com diferenças entre domínios. Em média, os efeitos da personalização mediada por IA foram significativos e de magnitude moderada, com variação considerável entre estudos. A meta-análise recente de Joyal-Desmarais et al. (2022; k= 702) havia encontrado efeito semelhante (r= 0,20), menor em pesquisas de saúde (r= 0,12) e apontado cerca de 25% de resultados nulos ou negativos.

Já a meta-análise de Noar et al. (2007; k= 57), conduzida em intervenções de saúde com materiais impressos, mostrou que mensagens personalizadas superaram tanto as versões genéricas ou segmentadas a grupos gerais (r= 0,06, k= 40 estudos)

quanto as condições sem intervenção (r= 0,11; k= 17 estudos), o que confirma efeitos pequenos, mas consistentes. De modo semelhante, esta revisão encontrou 66% de resultados estatisticamente significativos entre 56 testes analisados, com predominância de efeitos positivos. A principal diferença, contudo, apareceu no domínio da saúde, que apresentou 86% de efeitos significativos, sugerindo um padrão possivelmente mais estável da personalização mediada por IA em temas ligados ao bem-estar e à promoção de comportamentos saudáveis.

No total, 43 testes envolveram variáveis psicológicas, dos quais 30 apresentaram significância estatística (70%), enquanto 13 não foram significativos (30%). Esse padrão é compatível com a literatura que destaca maior eficácia de personalizações baseadas em características psicológicas. Meta-análises em saúde também relatam efeitos pequenos quando a personalização se baseia apenas em variáveis demográficas, com coeficientes na faixa de 0,06 a 0,08, o que reforça a diferença entre abordagens sustentadas por dados simples e aquelas orientadas por atributos psicológicos (Rothman et al., 2025).

Entre os traços analisados, Abertura à Experiência e Extroversão foram os mais consistentes, com efeitos replicáveis em diferentes domínios e amostras agregadas de 3.423 e 1.355 participantes, respectivamente. No caso de Abertura, o único estudo que inicialmente não havia alcançado significância estatística, conduzido com o modelo GPT-3 (Simchon et al., 2024), tornou-se significativo ($p< 0,001$) quando replicado com um modelo mais recente ChatGPT (gpt-3.5-turbo), o que refletiu o ganho técnico entre gerações de modelos e sua maior capacidade de produzir mensagens alinhadas ao perfil psicológico dos participantes.

A maior presença de efeitos significativos para Abertura à Experiência e Extroversão pode estar relacionada ao fato de que esses traços apresentam padrões linguísticos mais distintos, como mostram análises em mídias sociais conduzidas por Park et al. (2015), que identificaram Abertura (r= 0,43) e Extroversão (r= 0,42) entre os traços mais inferíveis a partir da linguagem. Abertura tende a aparecer no tipo de conteúdo e vocabulário usados, enquanto Extroversão se manifesta em expressões sociais e afetivas.

No domínio político, os efeitos mais claros apareceram para fundamentos morais, como Justiça, Autoridade e Lealdade, além da identificação política. Esse padrão é compatível com estudos que mostram que mensagens políticas são mais

eficazes quando apelam aos valores morais do público-alvo (Voelkel; Feinberg, 2018).

Apesar dos resultados promissores discutidos anteriormente, parte dos estudos revisados apresentou efeitos nulos ou inconsistentes. Esses casos envolveram, em parte, o uso de variáveis pouco relacionadas ao objeto atitudinal e diferenças nas técnicas de personalização. Como discutem Teeny et al. (2025), diferentes domínios favorecem tipos distintos de correspondência entre mensagem e receptor, e certas variáveis só produzem efeito quando tem relação direta com o conteúdo persuasivo. Quando essa ligação é fraca ou inexistente, a personalização tende a gerar pouco ganho adicional.

Em Hackenburg et al. (2023), o *microtargeting* político não superou a melhor mensagem genérica, mesmo quando o número de atributos usados na personalização foi ampliado, chegando em alguns casos a incluir até nove variáveis de segmentação. Os autores sugerem que esse padrão pode ter ocorrido porque o GPT-4 não refletia bem as posições reais de certos grupos, porque o treinamento por reforço acabou reduzindo a diversidade interna do modelo e porque alguns temas, como energia renovável, já tinham apoio muito alto entre os participantes, deixando pouco espaço para detectar aumentos adicionais de persuasão.

Tappin et al. (2023) relataram padrão semelhante: não houve ganhos ao aumentar a complexidade do *microtargeting*, e os efeitos significativos apareceram apenas em um tema isolado. Em Hackenburg e Margetts (2024), o uso de simulação partidária também não aumentou a eficácia das mensagens, o que sugere que o efeito persuasivo dos LLMs está mais ligado à qualidade do texto produzido do que à simulação de uma identidade política.

Parte dos resultados nulos ocorreu em variáveis psicológicas específicas, como Amabilidade, Neuroticismo, Foco regulatório, Pureza e Cuidado, mas a maior concentração apareceu em estudos baseados em variáveis demográficas ou em combinações híbridas de dados psicológicos e sociodemográficos. Esse padrão indicou que traços puramente psicológicos tendem a ser mais sensíveis à personalização. Como discutem Teeny e Matz (2024), a personalização tende a ser mais eficaz quando se apoia em características psicologicamente significativas para o público, o que ajuda a explicar a limitação observada em estudos baseados apenas em atributos sociodemográficos. Salvi et al. (2025) sugerem, em sua discussão, que efeitos mais fortes poderiam ser alcançados caso a personalização se apoiasse em

atributos psicológicos mais ricos, como traços de personalidade e fundamentos morais, combinados a técnicas de ajuste mais refinadas, como *prompt engineering* (engenharia de prompts) e *fine-tuning* (treinamento adicional do LLM).

Os estudos revisados mostraram grande variação na forma de *prompting*. Meguellati et al. (2024) usaram instruções curtas, como "escreva uma propaganda de uma linha para um telefone chamado Xphone direcionado a pessoas com o traço de Abertura sem mencionar o traço explicitamente". Já Simchon et al. (2024) e Hackenburg e Margetts (2024) adotaram prompts mais detalhados e estruturados. Essa diversidade sugere que o modo como o *prompt* é formulado afeta a qualidade das mensagens geradas. Em um estudo voltado à saúde, Karinshak et al. (2023) observaram que *prompts few-shot*, com exemplos breves, produzem textos mais completos e persuasivos do que instruções simples em formato *zero-shot* (sem exemplo), indicando que essa etapa tem papel importante na personalização mediada por IA.

Outro fator possível envolve a percepção limitada da personalização. No estudo de Hackenburg e Margetts (2024), o teste de manipulação indicou que os participantes perceberam apenas de forma parcial o alinhamento partidário do modelo, o que sugere baixo reconhecimento da manipulação. Como discutem Teeny e Matz (2024), a personalização precisa ser percebida o suficiente para gerar relevância pessoal, mas não a ponto de se tornar excessivamente evidente. Quando o ajuste é pouco notado, tende a perder o efeito de relevância que sustenta a mudança persuasiva, o que pode ter contribuído para alguns resultados nulos.

É importante destacar que nenhum dos estudos avaliou se os participantes perceberam o ajuste ao traço psicológico utilizado. Em Matz et al. (2024) e Simchon et al. (2024), as escalas mediram apenas persuasão percebida, expressa por itens como "acho este anúncio persuasivo" ou "este anúncio é eficaz", que captam apenas como o anúncio foi percebido, não o reconhecimento da personalização. Nos demais trabalhos, as verificações se limitaram à autoria das mensagens ou ao alinhamento político. Como argumentou Li (2016), personalização efetiva e personalização percebida são construtos distintos, e a ausência de medidas diretas sobre essa percepção reduziu o que se pode concluir sobre seu impacto nos resultados.

Diferenças na saliência dos temas também podem ter influenciado os resultados. Druckman (2022) destacou que assuntos mais salientes, como política e

saúde, tendem a gerar maior engajamento cognitivo, mas também reações mais resistentes por envolverem crenças e identidades já formadas. Em contrapartida, temas menos salientes, como consumo, costumam despertar menor envolvimento e atitudes mais flexíveis, facilitando a influência de mensagens personalizadas. Essa variação ajuda a explicar por que alguns estudos obtiveram efeitos significativos, enquanto outros resultaram em achados nulos ou inconsistentes.

Considerações Finais

Em síntese, os achados analisados nesta revisão indicaram que a personalização mediada por IA apresentou um padrão regular de resultados, ainda que sujeito a variações entre domínios, tipos de variável e modos de personalização pelos sistemas de IA. As variáveis psicológicas apresentaram resultados mais uniformes que as demográficas, com destaque para Abertura à Experiência e Extroversão.

Entre os domínios, saúde exibiu os efeitos mais estáveis, enquanto política e consumo mostraram maior variação e presença de resultados nulos. Durante o período de elaboração desta revisão, foi lançado o modelo GPT-5 (OpenAI, 2025), o que evidencia a rapidez com que os modelos de linguagem evoluem. Paralelamente, o campo passou a contar com alternativas abertas e gratuitas de desempenho semelhante. Essa diversificação amplia o acesso à pesquisa e permite comparar modelos de geração de texto, além de avaliar variações na qualidade das mensagens. A chegada de modelos multimodais também abre espaço para testar persuasão personalizada que combine texto, imagem ou vídeo.

Os achados desta revisão reforçaram que a eficácia da personalização mediada por IA depende menos do número de variáveis utilizadas e mais da relevância psicológica dos atributos utilizados. Entre as variáveis psicológicas, traços de personalidade foram os mais consistentes, enquanto valores morais e ideologia exibiram efeitos pontuais, concentrados sobretudo no domínio político. Um estudo recente de grande escala, no campo da persuasão de IA na política, conduzido por Hackenburg et al. (2025; n= 76.977) em 19 LLMs, também indicou que fatores técnicos, como o pós-treinamento e o tipo de instrução utilizada, podem influenciar a força persuasiva das mensagens e que interações dialogadas são substancialmente mais eficazes do que mensagens únicas geradas por IA.

Ainda assim, a personalização permanece um componente central, cuja efetividade depende do modo como é percebida e integrada ao conteúdo. Quando o ajuste é sutil demais ou se baseia em atributos pouco significativos, tende a ser ignorado pelo público. Esses aspectos apontam para a importância de integrar princípios da psicologia da personalidade e da persuasão ao desenvolvimento de intervenções mediadas por IA, tornando-as mais precisas, éticas e fundamentadas.

A maioria dos estudos utilizou medidas de autorrelato e avaliações de curta duração, o que pode não capturar integralmente o impacto real e prolongado das mensagens personalizadas. Essa limitação metodológica limita o entendimento sobre os efeitos comportamentais de longo prazo, especialmente em intervenções voltadas à saúde mental, nas quais o engajamento contínuo é essencial.

Além das restrições inerentes aos estudos incluídos, esta revisão também apresenta limitações próprias. O número reduzido de pesquisas disponíveis, a heterogeneidade de delineamentos e medidas e a concentração geográfica das amostras dificultam comparações diretas e sínteses quantitativas mais precisas. Embora tenha seguido princípios do PRISMA e de revisões semi-sistemáticas, é possível que publicações recentes, ainda em *preprint* ou fora das bases indexadas, não tenham sido contempladas. Por fim, a análise adotou uma abordagem descritiva, voltada a identificar padrões e lacunas nos estudos revisados. Esse tipo de síntese envolve algum grau de subjetividade, ainda que as decisões tenham sido guiadas por critérios explícitos definidos previamente.

Apesar da maioria dos efeitos observados tenha sido de magnitude pequena ou moderada, seu potencial em escala pode ser expressivo. Como destacam Simchon et al. (2024), efeitos modestos podem influenciar milhares de pessoas quando aplicados amplamente, chegando a alterar comportamentos coletivos e até resultados eleitorais. De forma complementar, Matz et al. (2023) argumentam que o valor de um efeito depende do contexto: um acréscimo mínimo na probabilidade de evitar um suicídio, por exemplo, pode ter importância muito maior do que o mesmo ganho em comportamentos de baixo impacto social. Assim, a relevância prática da personalização mediada por IA vai além de seus efeitos imediatos e reside no potencial de apoiar mudanças relevantes quando utilizada de maneira ética. A ausência de estudos em saúde mental na revisão reforça a necessidade de evidências específicas nesse campo.

Artigo II

Persuasão Personalizada Mediada por Inteligência Artificial: um Experimento com Chatbots para Manejo do Estresse

Personalized Persuasion Mediated by Artificial Intelligence: An Experiment with Chatbots for Stress Management

Persuasión Personalizada Mediada por Inteligencia Artificial: Un Experimento con Chatbots para el Manejo del Estrés

Bruno Grossman

Heder Soares Bernardino

Jairo Francisco de Souza

Leonardo Fernandes Martins

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil

Resumo

A persuasão personalizada tem se mostrado uma estratégia eficaz para promover mudanças de atitude e comportamento em diferentes domínios, variando de campanhas de saúde pública a engajamento político e decisões de consumo. Com o avanço técnico dos modelos de linguagem de grande escala (LLMs), passou a ser explorada a possibilidade de automatizar essa personalização em larga escala, permitindo que agentes de inteligência artificial adaptem mensagens ao perfil psicológico dos indivíduos. No entanto, a aplicação dessa estratégia em contextos de saúde mental ainda é incipiente, apesar de seu potencial para ampliar o acesso a intervenções e reduzir desigualdades no cuidado. O presente estudo examinou a eficácia persuasiva de mensagens geradas por IA, personalizadas de acordo com o nível de Abertura à Experiência, no contexto do uso de *chatbots* para gerenciamento de estresse diário. Participaram 451 estudantes universitários brasileiros (60% mulheres; M= 25,9), distribuídos aleatoriamente entre três condições experimentais: mensagem personalizada para alta Abertura, personalizada para baixa Abertura e mensagem genérica (não personalizada). As variáveis foram avaliadas por meio de instrumentos validados, abrangendo Abertura à Experiência, Atitudes frente aos Chatbots e Abertura à Tecnologia. A análise de covariância indicou que o tipo de mensagem não exerceu efeito significativo sobre as atitudes, tampouco houve interação com o nível de Abertura à Experiência. Por outro lado, níveis mais altos de Abertura à Tecnologia e níveis mais baixos de Abertura à Experiência associaram-se a atitudes mais favoráveis ao uso de *chatbots*. Observou-se também que estudantes de Psicologia apresentaram atitudes ligeiramente mais cautelosas em relação aos *chatbots* do que estudantes de outras áreas, independentemente do tipo de mensagem. No geral, os resultados indicam que diferenças individuais exerceram influência mais consistente do que a personalização automatizada na formação de atitudes frente a agentes de IA voltados ao bem-estar psicológico.

Palavras-chave

Persuasão personalizada; inteligência artificial; personalidade; saúde mental digital

Abstract

Personalized persuasion has been shown to be an effective strategy for promoting attitude and behavior change across different domains, ranging from public health campaigns to political engagement and consumer decision-making. With recent advances in large language models (LLMs), researchers have begun exploring the possibility of automating personalization at scale, allowing artificial intelligence agents to tailor messages to individuals' psychological profiles. However, the application of this approach in mental health contexts remains incipient, despite its potential to expand access to interventions and reduce inequalities in care. The present study examined the persuasive effectiveness of AI-generated messages personalized according to individuals' levels of Openness to Experience in the context of using chatbots for daily stress management. A total of 451 Brazilian university students (60% women; M= 25.9) were randomly assigned to one of three experimental conditions: a message tailored for high Openness, a message tailored for low Openness, or a generic (non-personalized) message. Variables were assessed using validated instruments covering Openness to Experience, Attitudes toward Chatbots, and Openness to Technology. The analysis of covariance indicated that message type did not have a significant effect on attitudes, nor did it interact with Openness to Experience. On the other hand, higher levels of Openness to Technology and lower levels of Openness to Experience were associated with more favorable attitudes toward chatbots. Psychology students also showed slightly more cautious attitudes compared to students from other fields, regardless of message type.

Taken together, the results suggest that individual differences exerted a more consistent influence than automated personalization on the formation of attitudes toward AI-based agents designed to support psychological well-being.

Keywords

Personalized persuasion; artificial intelligence; personality; digital mental health

Como persuadir alguém a aderir a um tratamento de cessação do tabagismo, a reduzir o consumo de álcool, a buscar apoio psicológico para lidar com a ansiedade ou a retomar hábitos de autocuidado? Entender como a comunicação pode ajudar as pessoas a reduzir o sofrimento e a promover mudanças de comportamento tem sido um tema central na Psicologia, que historicamente investiga os mecanismos pelos quais certas mensagens se tornam mais convincentes do que outras (Matz et al., 2023; Petty et al., 1997). O avanço recente da inteligência artificial (IA) tem expandido as possibilidades de estratégias persuasivas, permitindo investigar como sistemas automatizados podem favorecer a adoção de práticas de autocuidado e bem-estar psicológico.

A persuasão, entendida como o processo de modificar crenças, atitudes ou comportamentos por meio da comunicação (Luttrell et al., 2025), ocupa papel central nas interações humanas. Grande parte do comportamento social envolve, de forma explícita ou sutil, tentativas de influenciar ou ser influenciado por outros (Cialdini, 2001).

Os esforços persuasivos podem assumir finalidades diversas, como promover o bem-estar coletivo, estimular comportamentos altruístas, orientar escolhas de consumo ou mobilizar decisões políticas. Embora muitos desses esforços adotem abordagens generalistas, nas últimas décadas ganhou um corpo substancial na literatura e na pesquisa experimental uma estratégia particularmente eficaz: a persuasão personalizada. Conceitualmente, é definida como o alinhamento deliberado de um ou mais aspectos da comunicação (conteúdo e forma da mensagem, fonte ou contexto) a uma ou mais características do destinatário, incluindo traços psicológicos, motivações, valores, identidades ou outros atributos individuais (Matz et al., 2023; Teeny et al., 2021).

Essa correspondência tende a aumentar a relevância pessoal e a fluidez cognitiva da comunicação, favorecendo sua aceitação e potencializando os efeitos atitudinais (Teeny et al., 2021). Estudos experimentais e revisões recentes indicam que essa estratégia tem produzido resultados consistentes em múltiplos domínios e variáveis psicológicas. A meta-análise conduzida por Joyal-Desmarais et al. (2022; k= 702) identificou um efeito consistente da correspondência motivacional sobre atitudes, intenções e comportamentos (r= 0,20) em comparação a mensagens neutras ou não correspondentes, e um efeito maior (r= 0,30) quando comparada

especificamente a mensagens incongruentes. A maior parte (30%) das pesquisas da revisão concentrou-se nas áreas da saúde (r= 0,125) e do consumo.

O Modelo da Probabilidade de Elaboração (Elaboration Likelihood Model, ELM; Petty e Cacioppo, 1986) oferece uma base teórica para compreender os mecanismos psicológicos pelos quais mensagens persuasivas produzem mudança atitudinal. O modelo propõe que a persuasão pode ocorrer ao longo de um contínuo de processamento cognitivo, que varia entre duas rotas principais: uma rota central, em que o indivíduo processa cuidadosamente os argumentos da mensagem, e uma rota periférica, baseada em pistas simples ou afetivas. O nível de elaboração depende da motivação e da capacidade do indivíduo para refletir sobre o conteúdo, variando conforme fatores situacionais e disposicionais. Mudanças de atitude que resultam predominantemente do processamento pela rota central tendem a ser mais persistentes ao longo do tempo, mais resistentes à contra-persuasão e mais preditivas do comportamento do que aquelas produzidas principalmente por pistas periféricas.

A revisão de Teeny et al. (2021) demonstrou que os efeitos de congruência, o alinhamento entre características da mensagem e do receptor, podem ser compreendidos a partir dos processos descritos por esse modelo. Assim, a personalização de mensagens pode atuar como pista periférica de afinidade em contextos de baixa elaboração e, em níveis mais altos de processamento, como argumento relevante que orienta a direção dos pensamentos.

Além disso, pode influenciar processos metacognitivos de validação ou correção de julgamentos. Contudo, em níveis elevados de elaboração, o processamento mais crítico pode reduzir a eficácia persuasiva quando os argumentos são percebidos como fracos ou pouco convincentes. Importante observar que a personalização não garante maior impacto persuasivo: seus efeitos dependem do significado atribuído pelo indivíduo ao *match*, que pode ser interpretado como relevante e autêntico ou, ao contrário, como intrusivo e manipulativo. Desse modo, o ELM fornece um modelo processual útil para compreender quando e por que a personalização de mensagens resulta em maior aceitação, resistência ou reatância.

A personalização de mensagens é entendida como um contínuo, no qual os conteúdos variam do alinhamento positivo (*active match*) ao desalinhamento ativo (*active mismatch*), isto é, situações em que a mensagem toca diretamente um

aspecto motivacional do receptor de forma contrária ao seu perfil (por exemplo, enfatizar disciplina rígida para alguém que valoriza autonomia em saúde mental). Entre esses extremos situam-se as mensagens neutras (*non-match* ou *passive mismatch*), que não apoiam nem contradizem essas disposições. Como mensagens alinhadas e desalinhadas ativam processos motivacionais distintos, comparações diretas entre as duas podem misturar efeitos diferentes e dificultar a interpretação. A noção de contínuo ajuda a organizar essas diferenças, pois qualquer mensagem persuasiva ocupa algum ponto entre o desalinhamento ativo e o alinhamento, permitindo avaliar a personalização de modo mais preciso (Joyal-Desmarais et al., 2025).

No contexto de saúde, o estudo de Williams-Piehota et al. (2003) ilustrou a aplicação do ELM ao investigar mensagens destinadas a incentivar a realização do exame de mamografia. As participantes receberam comunicações ajustadas ao seu nível de necessidade de cognição, e o alinhamento entre esse perfil e o formato da mensagem aumentou em cerca de 34% a probabilidade de realização do exame. O trabalho mostrou, na prática, o potencial do que a literatura tem chamado de correspondência personalizada (*personalized matching*), termo também usado em diferentes campos para designar a adaptação à personalidade (*personality tailoring*), correspondência motivacional (*motivational matching*) e a comunicação direcionada (*targeting*, *microtargeting*). Embora cada expressão tenha nuances específicas, todas remetem ao mesmo princípio: conteúdos tendem a ser mais eficazes quando se alinham diretamente a características individuais ou grupais do público, como ressaltam Matz et al. (2023) e Teeny et al. (2021).

A persuasão personalizada envolve duas etapas complementares: o *perfilamento psicológico*, que estima características individuais a partir de rastros digitais e outras pistas comportamentais, e a *intervenção personalizada*, na qual essas informações orientam a adaptação de mensagens às motivações psicológicas do público, aumentando sua eficácia persuasiva (Matz et al., 2020). Embora apresente grande potencial, sua aplicação em larga escala foi por muito tempo limitada por restrições metodológicas e éticas. Métodos tradicionais de coleta, como questionários e entrevistas, são lentos, caros e sujeitos a vieses (Podsakoff; Organ, 1986), o que dificulta obter perfis psicológicos de grandes grupos em contextos aplicados.

A expansão das mídias sociais e das plataformas digitais tornou a personalização de conteúdos mais acessível e precisa, estimulando o interesse científico e público por estratégias de comunicação ajustadas a características individuais. O caso Cambridge Analytica evidenciou o poder e os riscos do uso de dados pessoais para moldar mensagens direcionadas (Simchon et al., 2023). Desde então, o avanço na coleta e análise de grandes volumes de informação sobre usuários conectados, aliado às técnicas de aprendizado de máquina e mais recentemente à inteligência artificial generativa, tem ampliado a possibilidade da persuasão em larga escala por meio de mecanismos de personalização cada vez mais sofisticados (Matz et al., 2017, 2024).

Nos últimos anos, modelos de aprendizado de máquina e o uso intenso de plataformas digitais ampliaram as formas de analisar os rastros deixados por usuários. Estudos mostram que dados aparentemente simples, como *likes* em redes sociais, podem ser usados para estimar traços de personalidade e prever características como orientação política e etnia, e que imagens faciais on-line permitem inferir orientação sexual com precisão maior que a de julgamentos humanos (Kosinski et al., 2013; Youyou et al., 2015; Wang; Kosinski, 2018). Esses resultados ajudam a ilustrar tanto o alcance técnico dessas ferramentas quanto os riscos envolvidos no uso de dados sensíveis. Em paralelo a esse cenário, consolidou-se o campo da psicometria computacional, que reúne ciência de dados e teoria psicométrica para desenvolver formas automatizadas de mensurar variáveis psicológicas em contextos de pesquisa e avaliação (Franco; Primi, 2022).

O surgimento dos modelos de linguagem de grande escala, sistemas de aprendizagem profunda treinados em volumes massivos de texto, capazes de compreender e gerar linguagem natural em múltiplas tarefas (IBM, 2023), como o GPT-4, inaugurou uma nova etapa da persuasão mediada por inteligência artificial.

Uma meta-análise (k= 121) mostrou que agentes de IA influenciam percepções, atitudes e comportamentos com eficácia semelhante à de comunicadores humanos, embora com menor impacto sobre intenções comportamentais (Huang; Wang, 2023). Em experimentos, mensagens geradas por LLMs conseguiram mudar opiniões políticas (Bai et al., 2025) e foram avaliadas como mais persuasivas e positivas do que comunicados oficiais de saúde (Karinshak et al., 2023). O avanço dos modelos de linguagem de grande escala aprofundou essa mudança, com a introdução da chamada *persuasão mediada por*

LLM, termo proposto por Rogiers et al. (2024) para descrever pesquisas que analisam efeitos persuasivos mediados por sistemas de IA generativa.

O crescimento de sistemas baseados em inteligência artificial, como *chatbots*, tem transformado a forma como as pessoas interagem com a tecnologia. Ferramentas conversacionais baseadas em LLMs, como o ChatGPT, alcançaram ampla adoção desde seu lançamento (Reuters, 2023). No Brasil, 63% da população já utilizou alguma aplicação de IA generativa, e 16% relatam uso diário (Nexus, 2025).

Em escala global, estudos apontam uma atitude ambivalente diante da inteligência artificial, marcada por otimismo quanto às suas potencialidades e preocupação com seus riscos (Kaya et al., 2024). Ao mesmo tempo, esse avanço suscita questões éticas e sociais relacionadas à privacidade, à desinformação e à autonomia (Hendrycks et al., 2023), além do surgimento da chamada ansiedade de IA, caracterizada por medo ou desconforto diante da crescente integração desses sistemas à vida cotidiana (Kaya et al., 2024). Considerando o potencial dos *chatbots* voltados à saúde mental para atenuar parte dos desafios contemporâneos em saúde mental, compreender como as pessoas percebem, adotam e interagem com essas ferramentas tornou-se fundamental para sua aceitação e efetividade.

Dados recentes da OMS (2025) mostram que mais de um bilhão de pessoas no mundo convivem com algum transtorno mental, cerca de 14% da população. Entre esses quadros, ansiedade e depressão têm os diagnósticos mais frequentes, somando cerca de dois terços dos casos globais. Na América Latina e no Caribe, as taxas chegam a 17%, cenário que, no Brasil, se torna mais crítico devido às desigualdades regionais e às dificuldades enfrentadas pelas políticas públicas de atenção psicossocial (Araújo; Torrenté, 2023).

O estresse diário refere-se a pequenas demandas e pressões rotineiras que exigem adaptação constante do indivíduo e compõem parte natural da vida cotidiana (Mauriello et al., 2021). Diferentemente do estresse crônico, esses estressores são geralmente breves e não persistem de um dia para o outro (Almeida, 2005). Quando, porém, ocorrem de forma repetida e prolongada, podem contribuir para sobrecarga fisiológica e aumento do risco de condições associadas, como alterações imunológicas, cardiovasculares e metabólicas (McEwen, 1998).

Nesse contexto, *chatbots* voltados à saúde mental, também denominados agentes conversacionais autônomos ou assistentes virtuais baseados em IA, têm

sido estudados como ferramentas digitais de apoio psicológico que utilizam inteligência artificial para interagir com o usuário em linguagem natural, oferecendo suporte emocional, psicoeducação e exercícios terapêuticos inspirados em abordagens psicológicas validadas (Casu et al., 2024). Um estudo conduzido nos serviços do sistema público de saúde inglês, mostrou que um *chatbot* de autorreferência baseado em IA aumentou em 15% o número de encaminhamentos (versus 6% em serviços-controle) e ampliou o acesso entre minorias de gênero e étnicas (Habicht et al., 2024; n= 129.400). Pesquisas indicam que *chatbots* como Woebot (Fitzpatrick et al., 2017), Wysa (Chaudhry; Debi, 2024) e Tess (Fulmer et al., 2018) reduzem sintomas de ansiedade e estresse e contribuem para a diminuição do estigma associado à busca por apoio psicológico (Durden et al., 2023). Essas ferramentas também se destacam pela disponibilidade contínua, anonimato e baixo custo, fatores que favorecem sua aceitação e o engajamento de usuários em contextos de autocuidado (Casu et al., 2024).

Apesar dos avanços na personalização mediada por inteligência artificial, ainda são escassas as evidências sobre como mensagens adaptadas influenciam atitudes e intenções de uso em contextos de saúde mental digital. Embora a Abertura à Experiência pareça favorecer a aceitação de tecnologias inteligentes, permanece uma lacuna quanto à sua interação com a personalização de mensagens geradas por IA. Investigar essa relação pode ajudar a compreender de que forma traços de personalidade modulam os efeitos da persuasão personalizada em contextos de cuidado psicológico digital.

5.
Objetivos e Hipóteses

5.1
Objetivo Geral

Avaliar a eficácia persuasiva de mensagens geradas por LLMs, personalizadas pelo traço Abertura à Experiência, na mudança de atitudes em relação ao uso de *chatbots* para gerenciamento de estresse.

5.2
Objetivos Específicos

A) Comparar a eficácia persuasiva de mensagens de texto personalizadas versus genéricas geradas por LLM.
B) Analisar o efeito da personalização na mudança de atitude em função de níveis altos e baixos de Abertura à Experiência.
C) Explorar possíveis diferenças nas atitudes frente aos *chatbots* entre estudantes de Psicologia e de outros cursos.

5.3 Hipóteses

H1a: Mensagens congruentes com o nível de Abertura à Experiência (*match*) terão maior eficácia persuasiva do que mensagens genéricas e mensagens incongruentes em relação ao traço de personalidade (*mismatch*). Espera-se observar um efeito principal de magnitude média (f = 0,25).

H1b: O efeito da personalização na mudança de atitude será moderado pelo nível de Abertura à Experiência, sendo mais forte para pessoas com alta Abertura do que para aquelas com baixa Abertura. Espera-se encontrar uma diferença no efeito também de magnitude média (f= 0,25).

H2: Estudantes de Psicologia apresentarão atitudes menos favoráveis frente aos *chatbots* para gerenciamento de estresse do que estudantes de outros cursos, independentemente do tipo de mensagem. Espera-se um efeito de magnitude média (f= 0,25).

6. Método

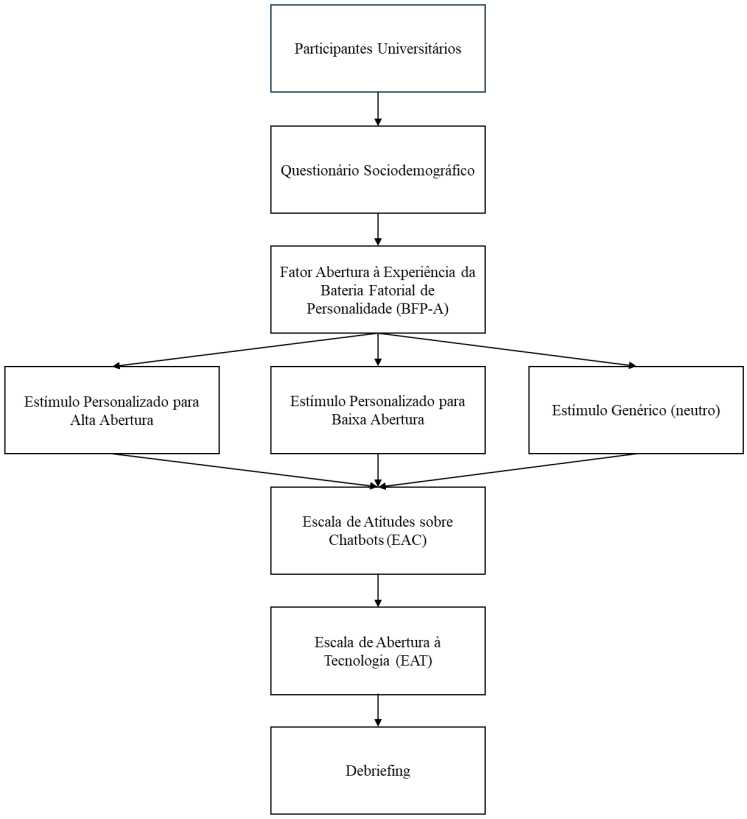
6.1 Delineamento Experimental

O estudo adotou um delineamento experimental entre sujeitos, com três condições independentes de exposição: mensagem personalizada para alta Abertura à Experiência, mensagem personalizada para baixa Abertura à Experiência e mensagem genérica não personalizada. Os participantes foram distribuídos entre as condições de forma aleatória e em proporções equivalentes, sendo expostos a apenas um estímulo textual.

O fator independente foi o tipo de mensagem, e a variável dependente foi o escore médio de atitude em relação ao uso de chatbots para gerenciamento de estresse. A Abertura à Experiência foi considerada variável moderadora, operacionalizada por meio da categorização em grupos de baixa e alta Abertura. A Abertura à Tecnologia foi incluída como covariável, com o objetivo de controlar diferenças individuais de familiaridade tecnológica na avaliação das mensagens.

Esse delineamento permitiu examinar o efeito principal do tipo de mensagem sobre as atitudes frente aos *chatbots*, bem como o efeito de interação entre o tipo de mensagem e o nível de Abertura à Experiência.

Figura 1 – Delineamento Experimental.



6.2 Participantes

Participaram da pesquisa 451 estudantes universitários de diferentes áreas do conhecimento, com idades entre 18 e 64 anos (M= 25,9; DP= 8,9). A maioria era do gênero feminino (60%) e todos os participantes eram maiores de 18 anos e falantes nativos de português. A amostra foi composta principalmente por

estudantes de graduação (81%), seguidos por participantes de mestrado (11%), doutorado (5%) e especialização (3%).

A distribuição por curso indicou leve predominância de estudantes de Psicologia (21%) e Engenharia (13%), seguidos por Design (6%), Ciência da Computação (5%), Direito (4%) e Administração (4%). Os demais participantes (47%) foram agrupados na categoria "Outros", que reúne cursos de diferentes campos das ciências humanas, exatas e biológicas, refletindo a heterogeneidade do público universitário incluído no estudo.

Tabela 5 – Dados sociodemográficos da amostra.

Participantes	**Frequência**	**Porcentagem (%)**
Gênero		
Masculino	168	37,3%
Feminino	269	59,6%
Não-binário	10	2,2%
Outro	1	0,2%
Prefiro não responder	3	0,7%
Faixa Etária		
18-24	275	61,1%
25-34	111	24,7%
35-49	52	11,6%
50+	12	2,7%
Nível de Formação (cursando)		
Graduação	366	81,2%
Mestrado	48	10,6%
Doutorado	22	4,9%
Especialização	15	3,3%
Curso		
Psicologia	95	21,1%
Engenharia	60	13,3%
Design	25	5,5%
Ciência da Computação	24	5,3%
Direito	18	4,0%
Administração	17	3,8%
Letras	14	3,1%
Ciências Econômicas	10	2,2%
Outros	188	41,7%

6.3. Instrumentos

Utilizou-se um questionário *online* elaborado na plataforma Qualtrics, composto por perguntas sociodemográficas (Anexo A) e três instrumentos

psicológicos. As perguntas iniciais abordaram dados gerais dos participantes (idade, gênero, curso e nível de formação). Em seguida, foram aplicadas as escalas que compõem as variáveis do estudo: o Fator Abertura à Experiência da Bateria Fatorial de Personalidade (BFP-A), a Escala de Atitudes sobre Chatbots (EAC) e a Escala Breve de Abertura à Tecnologia (EAT). Cada instrumento é descrito a seguir.

Fator Abertura à Experiência da Bateria Fatorial de Personalidade

A Bateria Fatorial de Personalidade (Nunes et al., 2010) é um instrumento psicológico desenvolvido para avaliar a personalidade com base no Modelo dos Cinco Grandes Fatores (CGF). Diferentemente de escalas internacionais adaptadas ao português, a BFP foi criada especificamente para o contexto brasileiro, considerando particularidades culturais, linguísticas e regionais. O instrumento é composto por 126 itens que descrevem sentimentos, opiniões e atitudes relacionadas às cinco dimensões da personalidade: Extroversão, Socialização, Realização, Neuroticismo e Abertura à Experiência.

Neste estudo, foi utilizado o fator Abertura à Experiência (BFP-A), que avalia a predisposição dos indivíduos a buscar novas experiências, aceitar ideias inovadoras e demonstrar curiosidade intelectual. Esse fator é composto por 23 itens, organizados em três facetas: Interesse por Novas Ideias, Liberalismo e Busca por Novidades. As respostas foram registradas em escala Likert de sete pontos, variando de “Descreve-me muito mal” (1) a “Descreve-me muito bem” (7).

O escore total de Abertura à Experiência foi calculado pela média aritmética dos 23 itens, após recodificação dos itens invertidos, de modo que valores mais altos indicassem maior nível de Abertura. Exemplo de item: “Acho que não existe uma verdade absoluta.” No estudo original (Nunes et al., 2010), o fator apresentou consistência interna satisfatória (α= 0,74). No presente estudo, a consistência interna foi α= 0,80 (M= 4,73; DP= 0,75), indicando boa confiabilidade.

Escala de Atitudes sobre Chatbots

Instrumento adaptado de Hackenburg et al. (2023) para avaliar atitudes em relação ao estímulo experimental, o uso de *chatbots* para gerenciamento de estresse diário. A escala foi composta por nove itens avaliados em escala Likert de 5 pontos, variando de “Discordo totalmente” (1) a “Concordo totalmente” (5), sendo que os sete primeiros mediram atitudes avaliativas e os dois últimos mediram intenção

comportamental sobre o uso de *chatbots*. Exemplo de item: “O uso de *chatbots* para gerenciamento de estresse diário traz boas consequências.”

O escore total foi calculado pela média aritmética dos nove itens, após a recodificação do item inverso. O resultado variou de 1 a 5, com valores mais altos indicando atitudes mais favoráveis ao uso de *chatbots* para gerenciamento de estresse. A escala apresentou boa consistência interna α= 0,91 (M= 2,69; DP= 0,98) e suas instruções aos participantes encontram-se no Anexo B.

Escala Breve de Abertura à Tecnologia

Para controlar diferenças individuais de familiaridade tecnológica na avaliação das atitudes frente aos *chatbots*, foi incluída uma medida breve de Abertura à Tecnologia (EAT). A escala consiste em três itens avaliados em escala Likert de cinco pontos, variando de “Discordo totalmente” (1) a “Concordo totalmente” (5). Exemplo de item: “Eu costumo adotar novas tecnologias antes da maioria das pessoas.” A escala apresentou consistência interna adequada (α= 0,73; M= 3,85; DP= 0,85), indicando boa confiabilidade dos itens. A escala encontra-se no Anexo C.

6.4 Materiais

Estímulos Textuais

Os estímulos textuais consistiram em três mensagens curtas sobre o uso de *chatbots* para gerenciamento de estresse diário, elaboradas especificamente para este estudo. As mensagens foram integralmente geradas por meio do ChatGPT-4o, então considerado entre os modelos mais avançados em geração automática de texto (Shahriar et al., 2024).

O desenho dos *prompts* baseou-se em abordagens apresentadas em estudos prévios sobre persuasão personalizada mediada por IA (Hackenburg; Margetts, 2024; Matz et al., 2024; Meguellati et al., 2024; Simchon et al., 2024). Optou-se pelo uso de *prompts* simples, visando reproduzir um cenário mais próximo da realidade de usuários ou organizações com pouca experiência em engenharia de prompts.

Foram definidas três condições experimentais: (a) mensagem personalizada para alta Abertura à Experiência, (b) mensagem personalizada para baixa Abertura à Experiência e (c) mensagem genérica não personalizada (condição controle). Cada participante foi exposto a apenas uma das versões.

As mensagens apresentaram extensão e estrutura argumentativa semelhantes, diferenciando-se apenas no enquadramento do conteúdo conforme o perfil de Abertura (Tabela 6). Os *prompts* e estímulos utilizados são apresentados a seguir.

Tabela 6 – Prompts utilizados.

Baixo nível de Abertura
Prompt Escreva uma mensagem persuasiva com cerca de 150 palavras para encorajar de forma sutil o uso de chatbots para lidar com o estresse cotidiano, usando linguagem natural, espontânea e não promocional. Direcione o texto especialmente a pessoas com baixo nível de Abertura à Experiência (Big Five).
Estímulo Às vezes, o dia parece puxado demais. Entre tarefas, compromissos e preocupações, a cabeça não pára, e o estresse vai se acumulando sem a gente perceber. Nessas horas, ajuda ter à mão algo simples e confiável. Os chatbots funcionam assim: são diretos, fáceis de usar e ajudam a colocar as ideias no lugar ou até a desabafar um pouco, sem complicação. Não é preciso aprender nada novo, nem mudar a rotina. É como fazer uma lista ou anotar um lembrete — só que com respostas na hora, quando você precisa. E o melhor: estão sempre disponíveis, sem precisar marcar hora ou esperar. Para quem prefere praticidade no dia a dia, é uma forma tranquila de aliviar a mente sem mudar o que já funciona. Às vezes, o que mais ajuda é justamente o que não exige muito da gente — e, só de dar esse respiro, o dia já fica mais leve.
Alto nível de Abertura
Prompt Escreva uma mensagem persuasiva com cerca de 150 palavras para encorajar de forma sutil o uso de chatbots para lidar com o estresse cotidiano, usando linguagem natural, espontânea e não promocional. Direcione o texto especialmente a pessoas com alto nível de Abertura à Experiência (Big Five).
Estímulo Você já reparou como, às vezes, só de escrever o que estamos sentindo, as coisas ficam mais claras? É curioso como a mente encontra caminhos quando tem espaço para experimentar, sem julgamentos. Tem gente que usa diários, outros gravam áudios pra si mesmos… e há quem converse com um chatbot. Não é sobre respostas prontas, mas sobre brincar com ideias, testar possibilidades, escutar os próprios pensamentos ganhando forma. Pode parecer inusitado no começo, mas muitos acabam descobrindo ali um jeito novo — e surpreendentemente íntimo — de organizar emoções, refletir ou até criar. É como um caderno que responde, sem pressa e sem regras. Só você, suas palavras, e a liberdade de explorar o que vier. Não é terapia, não substitui gente de verdade, mas pode ser um pequeno laboratório mental para quem gosta de pensar diferente. Afinal, por que não experimentar uma nova forma de conversar consigo mesmo?
Mensagem neutra (não personalizada)
Prompt Escreva uma mensagem persuasiva com cerca de 150 palavras para encorajar de forma sutil o uso de chatbots para lidar com o estresse cotidiano, usando linguagem natural, espontânea e não promocional.

Estímulo
Sabe aqueles dias em que tudo parece estar dando errado, e a mente fica a mil? Nessas horas, ter um espaço para soltar um pouco do peso pode fazer diferença. Não precisa ser algo grandioso ou complicado – às vezes, só conversar, nem que seja com um chatbot, já ajuda. Ele não julga, não interrompe, e está sempre disponível. Pode parecer estranho no começo, mas escrever ou até mesmo falar com ele sobre o que está passando na sua cabeça pode ser um jeito de organizar os pensamentos e aliviar um pouco do estresse. É como se a gente tirasse as palavras de dentro e colocasse pra fora, mesmo que ninguém esteja "ouvindo" de verdade. Vale a pena experimentar quando sentir aquele aperto no peito ou a cabeça cheia demais. Afinal, um pouco de leveza no dia a dia nunca faz mal, não é?

Validação dos Estímulos Textuais

Os estímulos textuais foram validados previamente em um procedimento independente com 8 juízes mestres e doutores em Psicologia (Anexo E). O objetivo foi verificar se as mensagens representavam corretamente os níveis alto e baixo de Abertura à Experiência. Cada avaliador leu três mensagens curtas geradas pelo LLM, sendo duas sobre o uso de *chatbots* para gerenciamento de estresse e uma mensagem de controle sobre *chatbots* para treinamento de idiomas. Em seguida, indicaram em uma escala de cinco pontos, que variava de "muito baixo" a "muito alto", o quanto cada traço de personalidade do modelo dos Cinco Grandes Fatores parecia presente em cada texto. As avaliações indicaram que as mensagens de alta e baixa Abertura foram reconhecidas conforme o perfil esperado, confirmando a adequação das manipulações usadas no experimento.

6.5. Procedimentos

Aspectos Éticos

O estudo foi aprovado pela Câmara de Ética em Pesquisa da PUC-Rio e registrado na Plataforma Brasil sob o número CAAE 87448525.1.0000.8137. Todos os participantes foram voluntários e concordaram em participar após ler e aceitar o Termo de Consentimento Livre e Esclarecido (TCLE, Anexo F). As respostas foram coletadas de forma anônima, sem identificação pessoal, garantindo sigilo e confidencialidade dos dados conforme as diretrizes éticas da pesquisa em Psicologia. Ao final do questionário, foi apresentado um *debriefing* informando os reais objetivos do estudo e reforçando o compromisso ético e de confidencialidade dos dados (Anexo G).

Coleta de dados

A coleta de dados foi realizada de forma virtual, entre junho e setembro de 2025, utilizando a plataforma Qualtrics. O recrutamento ocorreu por conveniência, a partir da divulgação em *e-mails* institucionais, grupos acadêmicos e redes sociais de universidades.

Após aceitar o Termo de Consentimento Livre e Esclarecido (TCLE), os participantes foram informados de que o estudo investigava percepções sobre *chatbots* para gerenciamento de estresse, descrição utilizada para manter o objetivo real encoberto e reduzir vieses de expectativa. Em seguida, responderam ao questionário sociodemográfico e à Escala de Abertura à Experiência (BFP-A), que incluiu o primeiro item de atenção para controle da qualidade das respostas.

Os participantes foram alocados a uma das três condições experimentais (mensagem personalizada para alta Abertura, personalizada para baixa Abertura ou mensagem genérica). Antes da leitura do estímulo textual, foi apresentada uma explicação breve e padronizada sobre o conceito de *chatbots* de gerenciamento de estresse diário, garantindo compreensão uniforme do tema (Anexo D).

Após a exposição à mensagem correspondente, os participantes responderam à Escala de Atitudes sobre Chatbots, que incluiu o segundo item de atenção, e, em seguida, à Escala Breve de Abertura à Tecnologia. O tempo médio total de resposta foi de aproximadamente 44 minutos (DP= 240 minutos), com mediana de 9 minutos.

Análise de dados

Inicialmente, foram realizados procedimentos de limpeza da base e exclusão de casos com respostas inválidas. Dos 534 participantes que concluíram todas as etapas do questionário, foram excluídos 50 que falharam em uma ou mais questões de atenção e 33 que não estavam cursando atividades de formação universitária no momento da coleta, resultando em uma amostra final de 451 participantes.

Em seguida, foi avaliada a randomização para verificar a equivalência entre as condições experimentais em termos de variáveis sociodemográficas e demais medidas coletadas no pré-teste. Essa verificação também foi repetida após a dicotomização da variável de Abertura à Experiência, a fim de assegurar que a exclusão do grupo intermediário (tercil médio) não alterou a distribuição aleatória entre os braços experimentais. Os testes indicaram ausência de diferenças

significativas quanto a gênero, idade, curso e nível de formação ($p > 0,05$ em todos os casos).

As variáveis psicológicas contínuas apresentaram distribuições adequadas para análises paramétricas. O escore de Abertura à Experiência ($M = 4,73$; $DP = 0,75$) apresentou leve assimetria negativa e pequena violação da normalidade ($p = 0,022$), considerada irrelevante em função do tamanho da amostra. O escore de Atitude frente aos *chatbots* ($M = 2,69$; $DP = 0,98$) apresentou distribuição simétrica, sem *outliers*, e desvio não substancial de normalidade ($p < 0,001$). O escore de Abertura à Tecnologia ($M = 3,85$; $DP = 0,85$) apresentou leve assimetria negativa ($-0,64$) e violação de normalidade ($p < 0,001$), também considerada aceitável diante do tamanho amostral e da robustez dos testes paramétricos.

Para fins analíticos, o escore médio de Abertura à Experiência foi posteriormente categorizado em percentis, conforme o critério adotado por Carrella et al., (2024), Matz et al., (2024) e Ross et al., (2009), permitindo a comparação entre grupos de baixa (≤33%) e alta (≥66%) Abertura. Devido à categorização em grupos extremos de Abertura à Experiência, as análises que incluíram essa variável foram conduzidas com 294 participantes (de um total de 451). A variável Abertura à Tecnologia foi mantida como covariável contínua.

Para testar essa hipótese, utilizou-se uma análise de covariância (ANCOVA), com tipo de mensagem como fator, Abertura à Experiência como moderadora categorizada e Abertura à Tecnologia como covariável. O pressuposto de homogeneidade das variâncias foi atendido (teste de Levene: $F(11, 282) = 0,56$; $p = 0,859$).

Para o terceiro objetivo específico, referente à comparação entre estudantes de Psicologia e de outros cursos, a variável "Curso" foi recodificada em dois grupos ("Psicologia" e "Outros cursos") e incluída como fator fixo adicional na ANCOVA. Essa inclusão permitiu examinar possíveis diferenças nas atitudes frente aos *chatbots* associadas à área de formação e testar eventuais interações com os demais fatores experimentais.

O poder estatístico foi estimado para detectar um efeito de magnitude média ($f = 0,25$), considerando $\alpha = 0,05$ e o tamanho amostral previsto para as análises com os tercis extremos de Abertura ($n = 294$), resultando em probabilidade de aproximadamente 97,6% de detecção de efeitos dessa magnitude. O cálculo foi

realizado no G*Power 3.1 (Faul et al., 2007), e as análises foram conduzidas no Jamovi 2.6 (Jamovi, 2025), ambos de acessos livres.

7. Resultados

7.1. Efeitos principais do tipo de mensagem (H1a)

A análise de covariância (ANCOVA) foi conduzida para examinar o efeito principal do tipo de mensagem sobre as atitudes frente aos *chatbots* para gerenciamento de estresse, controlando a Abertura à Tecnologia e considerando a Abertura à Experiência como variável moderadora categorizada.

Tabela 7 – Resultados da ANCOVA para os efeitos principais e de interação nas atitudes frente aos chatbots.

Efeito	gl	F	p	η^2_p	f
Tipo de mensagem (H1a)	2, 286	1,52	0,221	0,01	0,1
Interação Mensagem × Abertura (H1b)	2, 286	0,28	0,759	0,002	0,04
Alunos de Psicologia (H2)	1, 286	6,72	0,01	0,008	0,15
Abertura à Experiência (principal)	1, 286	4,31	0,039	0,015	0,12
Abertura à Tecnologia (covariável)	1,286	4,6	0,033	0,016	0,13

Os resultados (Tabela 7) indicaram que o tipo de mensagem não exerceu efeito significativo sobre as atitudes dos participantes, $F(2, 286)= 1,52$; $p= 0,221$; $\eta^2_p= 0,01$; $f= 0,1$. O tamanho de efeito observado foi pequeno, sugerindo diferenças mínimas entre as condições experimentais. Assim, a Hipótese 1a não foi suportada: o tipo de mensagem não produziu variação significativa nas atitudes frente aos *chatbots*, indicando que mensagens personalizadas e não personalizadas apresentaram níveis semelhantes de eficácia persuasiva.

7.2. Interação entre tipo de mensagem e Abertura à Experiência (H1b)

A ANCOVA examinou o efeito de interação entre o tipo de mensagem e os níveis de Abertura à Experiência sobre as atitudes frente aos *chatbots* para gerenciamento de estresse, controlando a Abertura à Tecnologia como covariável (Figura 2). O resultado não indicou interação significativa entre as variáveis, $F(2, 286)= 0,28$; $p= 0,759$; $\eta^2_p= 0,002$; $f= 0,04$. O tamanho de efeito observado foi muito pequeno, sugerindo que o impacto do tipo de mensagem sobre as atitudes não variou

em função do nível de Abertura à Experiência dos participantes. Em outras palavras, tanto indivíduos com baixa quanto com alta Abertura apresentaram padrões semelhantes de atitude em relação ao uso de *chatbots*, independentemente do tipo de mensagem recebida. Dessa forma, a Hipótese 1b não foi suportada, pois não se observou evidência de moderação do traço de Abertura à Experiência na eficácia persuasiva das mensagens.

Figura 2 – Interação entre tipo de mensagem e Abertura à Experiência nas atitudes frente aos chatbots.

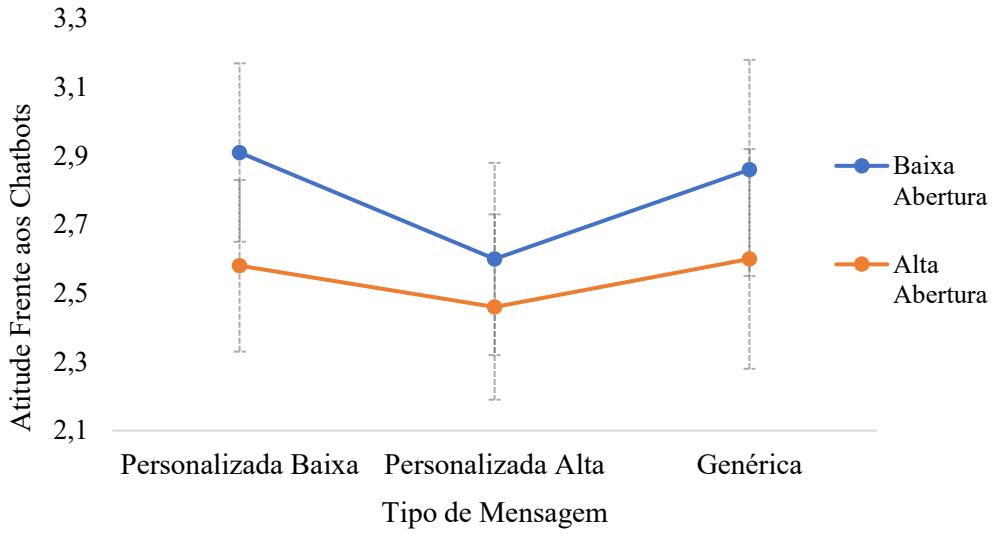


Tabela 8 – Médias marginais estimadas.

Abertura à Experiência	Tipo de Mensagem	Congruência	Média	Erro-padrão	IC a 95% Lim. Inferior	IC a 95% Superior
Baixo	Personalizada Baixa	Alinhada	2,91	0,13	2,65	3,17
	Personalizada Alta	Desalinhada	2,60	0,16	2,55	3,18
	Genérica	Neutra	2,86	0,14	2,32	2,88
Alto	Personalizada Alta	Alinhada	2,46	0,14	2,19	2,73
	Personalizada Baixa	Desalinhada	2,58	0,16	2,28	2,92
	Genérica	Neutra	2,60	0,13	2,33	2,83

7.3. Diferença entre estudantes de Psicologia e de outras áreas (H2)

A Hipótese 2 foi parcialmente suportada. O efeito de curso foi estatisticamente significativo, $F(1, 286)= 6,72$; $p= 0,010$, embora de pequena magnitude ($f= 0,15$). Esse resultado indicou que estudantes de Psicologia tenderam

a avaliar o uso de *chatbots* de forma ligeiramente mais cautelosa do que estudantes de outras áreas.

7.4. Efeitos principais das variáveis individuais

Além das hipóteses testadas, foram examinados os efeitos principais das variáveis individuais incluídas no modelo, independentemente das condições experimentais.

A Abertura à Tecnologia foi incluída como covariável na ANCOVA para controlar diferenças individuais de familiaridade tecnológica. O resultado indicou um efeito significativo dessa variável sobre as atitudes frente aos *chatbots*, $F(1, 286)= 4,60$; $p= 0,033$; $\eta^2_p= 0,016$; $f= 0,13$. O tamanho de efeito foi pequeno, mas estatisticamente relevante, sugerindo que participantes com maior familiaridade e interesse por tecnologias digitais apresentaram atitudes mais favoráveis ao uso de *chatbots* para gerenciamento de estresse, independentemente do tipo de mensagem recebida.

Além disso, observou-se um efeito principal significativo da Abertura à Experiência, $F(1, 286)= 4,31$; $p= 0,039$; $\eta^2_p= 0,015$; $f= 0,12$. Também de pequena magnitude, esse resultado indicou que, surpreendentemente, participantes com níveis mais baixos de Abertura tenderam a avaliar os *chatbots* de maneira ligeiramente mais positiva ($M= 2,78$; $DP= 1,00$) do que aqueles com alta Abertura ($M= 2,56$; $DP= 0,99$), independentemente da condição experimental.

De modo geral, os resultados da ANCOVA indicaram que o tipo de mensagem não exerceu efeito significativo sobre as atitudes frente aos *chatbots*, e não houve interação com a Abertura à Experiência. Por outro lado, observaram-se efeitos principais de Abertura à Tecnologia e de Abertura à Experiência, além da diferença entre estudantes de Psicologia e de outras áreas.

8. Discussão

O presente estudo teve como objetivo avaliar a eficácia de um modelo de linguagem de grande escala (LLM) na personalização de mensagens persuasivas de acordo com o nível de Abertura à Experiência, no contexto do uso de *chatbots* para gerenciamento de estresse diário. Os resultados permitiram examinar se o tipo de

mensagem influenciou as atitudes frente aos *chatbots* e se esse efeito variou em função da Abertura à Experiência, controlando diferenças individuais de Abertura à Tecnologia.

Os resultados não apoiaram as hipóteses H1a e H1b. O tipo de mensagem não exerceu efeito significativo sobre as atitudes dos participantes, nem houve interação com a Abertura à Experiência. Os tamanhos de efeito foram pequenos, indicando que, neste contexto, a personalização das mensagens geradas por IA não produziu impacto persuasivo relevante. Por outro lado, a H2 foi parcialmente suportada: estudantes de Psicologia apresentaram atitudes ligeiramente mais cautelosas em relação ao uso de *chatbo*ts do que estudantes de outras áreas, independentemente do tipo de mensagem recebida.

Embora os resultados deste estudo contrariem o padrão observado em investigações recentes sobre persuasão personalizada mediada por IA (Matz et al., 2024; Salvi et al., 2024; Simchon et al., 2024), com exceção de Hackenburg e Margetts (2024), é importante reconhecer que se trata de uma área emergente, na qual diferentes estratégias de personalização e contextos de aplicação ainda estão sendo testados e comparados. A literatura mais ampla sobre personalização motivacional, já discutida na introdução, mostrou que efeitos nulos não são incomuns. A meta-análise de Joyal-Desmarais et al. (2022), que havia identificado um efeito robusto (r= 0,20) e menor em estudos voltados à saúde (r= 0,125), também apontou uma heterogeneidade considerável, com cerca de 25% dos experimentos relatando efeitos nulos ou até negativos.

Diante dessa divergência, alguns fatores teóricos, metodológicos e contextuais podem ajudar a compreender a ausência de efeito observada. Sendo o primeiro estudo empírico da área a gerar mensagens em português brasileiro, uma possível explicação para o efeito nulo é que o LLM talvez não tenha conseguido personalizar de forma efetiva para esse traço de personalidade. Como referência, cerca de 93% do *corpus* usado para treinar o GPT-3 foi em língua inglesa (Brown et al., 2020) e grande parte do GPT-4 também teve o inglês como idioma principal (OpenAI, 2023). É plausível, portanto, que o modelo não possuísse o mesmo nível de sutileza e precisão ao gerar textos em português. LLMs costumam apresentar variações expressivas de desempenho conforme as referências culturais e linguísticas em que são utilizados (Almeida et al., 2025). Em outras palavras, é

possível questionar se o modelo realmente conseguiu produzir mensagens congruentes para pessoas com níveis distintos de Abertura à Experiência.

Além dessas limitações linguísticas e culturais do modelo, o efeito nulo também pode ter refletido aspectos do próprio delineamento experimental, especialmente se a personalização foi de fato percebida. Conforme demonstrado por Li (2016), a eficácia de mensagens personalizadas depende menos da personalização real e mais da personalização percebida pelo destinatário, sendo este o verdadeiro mecanismo psicológico que impulsiona a persuasão. As instruções incluídas nos *prompts*, como "encorajar de forma sutil" e "usar linguagem natural, espontânea e não promocional", foram inseridas para evitar que as mensagens parecessem artificiais ou invasivas. Embora essa escolha tenha contribuído para que o texto soasse mais natural, pode ter tornado as diferenças entre as condições pouco perceptíveis aos participantes.

Conforme discutido por Teeny e Matz (2024), a personalização precisa ser percebida o bastante para gerar relevância pessoal, mas sem que se torne excessivamente evidente. Quando é sutil demais, deixa de transmitir a sensação de relevância que sustenta o efeito persuasivo. É plausível, portanto, que as mensagens geradas neste estudo não tenham sido percebidas como suficientemente distintas entre si, limitando a percepção de personalização e, por consequência, o impacto persuasivo esperado.

O padrão das médias marginais ajudou a interpretar os resultados (Tabela 8). Entre participantes com baixa Abertura, a mensagem alinhada apresentou a maior média (2,91), seguida da neutra (2,86) e da desalinhada (2,60), reproduzindo o gradiente esperado no contínuo de personalização (Joyal-Desmarais, 2025). No grupo de alta Abertura, porém, a mensagem alinhada exibiu a menor média (2,46), ficando abaixo da neutra (2,60) e da desalinhada (2,58). Esse resultado sugeriu que a mensagem voltada à Abertura pode ter sido pouco sensível ao traço, mas não indica que tenha gerado rejeição ou conflito com o conteúdo apresentado. A condição desalinhada também não produziu respostas negativas e permaneceu próxima da neutra, indicando que o desalinhamento não foi percebido como conflitivo.

Por fim, a ausência de um teste de manipulação impediu confirmar se os participantes perceberam a personalização das mensagens. Como apontaram Rothman; Rogers; Mann (2025), definir a dosagem adequada de direcionamento é

um desafio constante, e muitos estudos acabam dependendo apenas da validade aparente dos estímulos. Efeitos nulos ou de pequena magnitude podem surgir quando o grau de correspondência entre o traço e a mensagem é insuficiente.

Ademais, um aspecto do delineamento experimental que pode ajudar a interpretar os resultados vem de estudos sobre inoculação ("*boosting*") de personalização persuasiva (Lorenz-Spreen et al., 2021; n= 828). Os autores observaram que o simples preenchimento de um pequeno questionário de personalidade, mesmo sem *feedback*, aumentou a consciência metacognitiva e a percepção de direcionamento psicológico. No presente estudo, os participantes também responderam à Escala de Abertura à Experiência antes da exposição às mensagens, o que pode ter gerado um efeito semelhante, ampliando a consciência sobre o próprio perfil e, com isso, levando a um processamento mais atento e metacognitivo do estímulo.

Esse aumento de consciência pode ter elevado o nível de escrutínio dos argumentos das mensagens (Teeny et al., 2021). Como o estímulo experimental era breve e continha argumentos relativamente simples, é plausível que essa reflexão tenha levado parte dos participantes, especialmente aqueles com maior Abertura à Experiência, traço que já foi associado à necessidade de cognição (Hu, 2022), a avaliar o conteúdo como pouco convincente. A ausência de um teste de suspeita, contudo, impediu verificar se os participantes perceberam o real objetivo da pesquisa ou associaram o questionário de personalidade às mensagens apresentadas.

Assim, a estratégia que buscava aumentar a eficácia persuasiva por meio da personalização pode ter produzido o efeito oposto, um leve "*backfire*", reduzindo a influência das mensagens personalizadas. Além disso, pesquisas no campo político indicaram que o público tende a rejeitar ações de microdirecionamento (Smith, 2018), o que reforça a possibilidade de reações defensivas à personalização se isso se torna saliente.

Esse mesmo processo pode ter ativado um efeito de reatância psicológica (Fransen et al., 2015), caracterizado por reações contrárias quando o indivíduo percebe uma tentativa de influência que ameaça sua autonomia. Assim, esse aumento de consciência metacognitiva pode ter deixado os participantes mais desconfiados, levando textos percebidos como direcionados ou artificiais a serem interpretados como uma tentativa de controle.

Uma possibilidade, ainda que pequena, é que parte dos participantes tenha percebido as mensagens como geradas por inteligência artificial. No período da coleta, houve destaque na mídia sobre o uso excessivo de travessões em textos produzidos pelo ChatGPT (CNN, 2025), padrão também presente nas três mensagens utilizadas. Embora Matz et al. (2024) indiquem que a autoria por IA não afeta significativamente a persuasão, pesquisas mostram que, quando as pessoas são informadas de que um argumento foi gerado por IA, sua eficácia persuasiva tende a diminuir (Palmer e Spirling, 2023). Além disso, estudos apontam baixa confiança pública em aplicações de IA na saúde mental, especialmente quanto à privacidade e à falta de contato humano (Varghese et al., 2024).

No caso da Hipótese 2, observou-se que participantes da área de Psicologia apresentaram médias significativamente mais baixas de atitude frente aos *chatbot*s ($p= 0,01$), mesmo exibindo níveis ligeiramente superiores de Abertura à Experiência ($M= 4,77$; $DP= 0,79$) e menores de Abertura à Tecnologia ($M= 3,60$; $DP= 0,89$) em comparação a estudantes de outras áreas ($M= 4,72$; $DP= 0,73$; $M= 3,92$; $DP= 0,82$, respectivamente), embora essas variáveis tenham sido controladas no modelo. Esse padrão sugere que o debate recente sobre o impacto da IA nas profissões de saúde mental (Conselho Federal de Psicologia, 2025) possa ter ampliado, entre esses estudantes, a percepção de risco de substituição profissional, levando a atitudes mais cautelosas diante de *chatbots* voltados ao apoio emocional. Quando a tecnologia se aproxima de práticas tipicamente humanas, como o cuidado psicológico, é possível que preocupações éticas e profissionais reduzam os efeitos positivos esperados de traços como a Abertura à Experiência ou da familiaridade com tecnologias digitais.

O padrão inverso observado para Abertura à Experiência também merece atenção. Embora revisões indiquem associações positivas entre esse traço e atitudes favoráveis frente à inteligência artificial (Riedl, 2022), estudos recentes têm reportado efeitos inconsistentes ou até invertidos (Stein et al., 2024). Além disso, variáveis mais específicas, como percepção de risco ou ansiedade em relação à IA, podem explicar atitudes de forma mais direta do que traços gerais de personalidade, sugerindo que a influência da Abertura pode variar conforme o tipo de tarefa e o contexto de interação proposto (Kaya et al., 2024).

8.
Considerações Finais

O presente estudo representou um passo inicial na investigação empírica da persuasão personalizada mediada por inteligência artificial aplicada à saúde mental. Essa é uma área ainda pouco explorada, embora apresente potencial relevante para ampliar o acesso a intervenções e compreender como diferentes perfis psicológicos respondem a comunicações automatizadas sobre bem-estar e autocuidado.

Embora as mensagens personalizadas não tenham se mostrado mais eficazes do que as demais condições experimentais, diferenças individuais, especialmente Abertura à Experiência e Abertura à Tecnologia, influenciaram diretamente as atitudes frente aos *chatbots*. Além disso, a Hipótese 2 foi parcialmente suportada. Estudantes de Psicologia apresentaram atitudes ligeiramente mais cautelosas frente aos *chatbots* do que estudantes de outras áreas, independentemente do tipo de mensagem, sugerindo que a familiaridade conceitual e o envolvimento profissional com temas de saúde mental podem influenciar a forma como esses sistemas são avaliados. Em conjunto, esses achados indicaram que, neste contexto, as variações de atitude dependeram mais das características pessoais do que das manipulações automáticas de personalização.

Teoricamente, os achados reforçaram a necessidade de compreender melhor como a personalização é percebida e até que ponto pequenas variações de linguagem são suficientes para gerar relevância psicológica. Avançar nesse campo exigirá desenvolver métodos capazes de medir e ajustar a dosagem de personalização, equilibrando sutileza e saliência para que ela seja de fato percebida e eficaz.

Metodologicamente, o estudo seguiu um delineamento considerado completo na literatura (Petty et al., 2024; Joyal-Desmarais et al., 2025) ao incluir mensagens correspondentes, não correspondentes e neutras, o que permitiu testar a congruência psicológica de forma controlada. Também demonstrou a viabilidade de empregar LLMs na geração de estímulos experimentais em português e indicou que a Escala de Atitudes sobre Chatbots, criada para este estudo, apresentou excelente consistência interna e adequação ao contexto investigado.

Do ponto de vista aplicado, os resultados sugerem que estratégias de comunicação sobre o uso de agentes conversacionais em saúde mental podem se

beneficiar de abordagens de personalização mais amplas, que considerem não apenas o perfil psicológico, mas também o estágio de familiaridade tecnológica do indivíduo. Para parte do público, a barreira principal pode ser o desconhecimento; para outros, a manutenção do engajamento. Adaptar a mensagem a esses estágios pode favorecer a clareza, a confiança e a disposição para adotar ferramentas de IA voltadas ao manejo de estresse.

Limitações e Pesquisas Futuras

Algumas limitações devem ser consideradas na interpretação dos resultados. O estudo analisou apenas um traço de personalidade, a Abertura à Experiência, selecionado pela sua relevância empírica para a pesquisa em persuasão personalizada. Ainda assim, outras variáveis psicológicas, como traços de personalidade distintos, necessidade de cognição (Williams-Piehota et al., 2003) ou foco regulatório (Fransen e Hoven, 2013), também já demonstraram potencial persuasivo e poderiam produzir padrões diferentes de resposta. Além disso, foi investigado um único objeto atitudinal, a aceitação de *chatbots* voltados ao manejo de estresse, o que restringe a generalização dos resultados a outros temas e contextos.

Outra limitação refere-se ao uso de apenas um LLM e de uma estratégia simples de *prompting*, em formato *zero-shot* (sem exemplos), baseada em breves descrições do traço psicológico. Não se pode descartar que outros *prompts*, formatos de instrução ou tópicos de mensagem produzam efeitos distintos. Pesquisas futuras poderiam comparar diferentes modelos de IA, testar abordagens de *few-shot prompting* (breves exemplos fornecidos ao modelo) e avaliar como variações no nível de detalhamento do perfil psicológico, no estilo e no tipo de argumento influenciam a eficácia das mensagens.

Do ponto de vista metodológico, a pesquisa utilizou uma amostra de conveniência composta por estudantes universitários, o que reduz a representatividade em termos de curso, gênero e região. As medidas de atitude e intenção basearam-se em autorrelato, sem verificação comportamental ou acompanhamento longitudinal. Além disso, as mensagens foram apresentadas em um ambiente experimental controlado, possivelmente gerando maior atenção do que ocorreria em situações reais de exposição. Estudos futuros podem recorrer a delineamentos de campo, incorporar medidas de comportamento efetivo e ampliar

a diversidade amostral para testar a robustez dos efeitos em contextos naturais de interação com tecnologias de IA.

Outra direção potencial envolve o uso de LLMs desenvolvidos no Brasil, como o Sabiá-3, da Maritaca AI, treinado em corpus de língua portuguesa e contexto cultural nacional (Abonizio et al., 2025). O modelo apresentou desempenho médio de 79% em 93 exames acadêmicos e profissionais, resultado próximo ao GPT-4o e superior ao Llama-3.1, com custo até quatro vezes menor por *token*. Em tarefas que exigem conhecimento contextual brasileiro, como provas do ENADE (Exame Nacional de Desempenho de Estudantes) e do CPNU (Concurso Público Nacional Unificado), o Sabiá-3 superou o GPT-4o, sugerindo o potencial desses sistemas para gerar mensagens mais contextualizadas e socialmente adequadas em pesquisas de persuasão personalizada.

Futuras investigações também poderiam explorar outros formatos de estímulo, como imagens, vídeos ou diálogos contínuos, e incluir novas formas de comparação, envolvendo interações com humanos, especialistas ou humanos em colaboração com IA (*human in the loop*), que podem superar o desempenho isolado de cada parte quando o humano mantém papel ativo nas decisões (Vaccaro el al., 2024). Alguns estudos recentes exploraram abordagens distintas: Hackenburg et al. (2023) compararam o desempenho persuasivo de IAs e de especialistas humanos em debates simulados, enquanto Costello et al. (2024) utilizaram diálogos com um *chatbot*, o que levou a uma redução de cerca de 20% em crenças conspiratórias, indicando o potencial das interações conversacionais em relação a estímulos únicos.

Considerando que o Brasil já figura entre os países que mais utilizam o ChatGPT semanalmente (Chatterji et al., 2025), com aproximadamente 140 milhões de mensagens enviadas por dia, sem contar outras plataformas conversacionais de IA, torna-se ainda mais relevante compreender como essas interações se traduzem em atitudes e comportamentos. Esse entendimento é essencial para que a persuasão personalizada mediada por IA evolua de um conceito emergente para uma abordagem empiricamente fundamentada de promoção do bem-estar psicológico

9.
Conclusão

A saúde mental segue marcada por alta prevalência de transtornos e por um acesso desigual aos serviços, agravado por estigma, baixa disponibilidade de profissionais, custo e distância (Durden et al., 2023). Essas barreiras têm estimulado o uso de ferramentas digitais que oferecem apoio psicológico de forma contínua e a baixo custo, como os *chatbots* baseados em inteligência artificial voltados ao manejo do estresse. Essas tecnologias têm sido avaliadas como capazes de ampliar o alcance do cuidado e facilitar o engajamento de grupos que costumam enfrentar maiores obstáculos para buscar apoio presencial (Habicht et al., 2024). Ainda assim, a adoção desses recursos não é automática e depende de como as pessoas percebem sua utilidade e adequação ao cotidiano. Por isso, a presente dissertação buscou investigar se a estratégia de persuasão personalizada mediada por inteligência artificial poderia contribuir para melhorar atitudes em relação a essas ferramentas.

O primeiro estudo consistiu em uma revisão semi-sistemática que identificou oito artigos empíricos publicados entre 2023 e 2025, abrangendo os domínios de política, consumo e saúde e totalizando 26.584 participantes. Entre os 56 testes extraídos desses subestudos, 66% apresentaram resultados significativos, com destaque para o domínio da saúde, que atingiu 86%. Os efeitos foram mais consistentes quando a personalização se apoiou em características psicológicas, sobretudo Abertura à Experiência e Extroversão. Apesar das diferenças entre os delineamentos, os achados indicam que a relevância psicológica do atributo usado na personalização pesa mais do que a quantidade de informações disponíveis.

No segundo estudo, buscou-se avaliar, de forma experimental, se mensagens personalizadas a partir do traço Abertura à Experiência, geradas pelo modelo de linguagem ChatGPT-4o, produziriam diferenças nas atitudes sobre o uso de *chatbots* para manejo de estresse. Participaram 451 estudantes universitários, distribuídos aleatoriamente entre três condições de mensagem (alta Abertura, baixa Abertura e versão genérica). As análises incluíram a Abertura à Tecnologia como covariável, e a moderação foi testada apenas com os participantes nos níveis extremos de Abertura (n= 294). As três condições apresentaram médias semelhantes de atitude e não houve moderação pelo traço, indicando ausência de efeito persuasivo da personalização textual. Em contraste, variáveis individuais como

abertura à tecnologia e área de formação mostraram influência mais clara. A diferença observada entre estudantes de Psicologia e de outras áreas, com avaliações ligeiramente mais cautelosas, confirma a hipótese de que esse grupo tenderia a perceber os *chatbots* de forma mais reservada. Isso sugere que predisposições prévias pesaram mais do que as diferenças entre os estímulos e que a sutileza da manipulação pode ter limitado a percepção da personalização, contribuindo para a ausência de efeitos.

Considerando os dois estudos, observa-se que a eficácia da personalização mediada por IA variou conforme o contexto e o modo como a estratégia é aplicada**.** Na revisão, a correspondência baseada em características psicológicas apresentou resultados mais uniformes, sobretudo quando relacionada a Abertura à Experiência e Extroversão. No experimento, porém, as mensagens personalizadas não alteraram atitudes, possivelmente porque as diferenças entre os textos foram sutis e pouco perceptíveis para os participantes. Além disso, a influência de fatores como abertura à tecnologia e área de formação aponta que respostas a esse tipo de intervenção variam conforme predisposições individuais e familiaridade com agentes de IA. Nos contextos de bem-estar psicológico, o efeito da personalização depende do quanto esse ajuste é de fato percebido pelo participante, o que provavelmente não ocorreu no estudo experimental.

Os resultados oferecem contribuições em dois níveis. No campo teórico, a revisão reforça a importância de basear a personalização em variáveis psicológicas com relação direta ao conteúdo persuasivo. No campo metodológico, o experimento mostra limites de personalizações textuais sutis em português e destaca a necessidade de verificar se a manipulação é reconhecida, etapa decisiva para compreender por que certos efeitos não emergem. Esses pontos ajudam a entender em quais condições de maior eficácia e situações em que os efeitos tendem a ser reduzidos ou inexistentes.

Algumas limitações devem ser consideradas. O experimento avaliou apenas o traço Abertura à Experiência, o que restringe a generalização dos resultados para outros perfis psicológicos. A personalização textual foi breve e pode ter produzido diferenças pouco perceptíveis entre as condições. O delineamento contou com medidas de autorrelato e um único momento de avaliação, o que impede verificar efeitos duradouros. Além disso, a amostra foi composta majoritariamente por estudantes universitários, o que limita a extrapolação para outros públicos.

Pesquisas futuras podem explorar diferentes traços e combinações de variáveis, testar formatos multimodais ou baseados em diálogo, comparar modelos de linguagem treinados em português e avaliar abordagens de *fine-tuning* voltadas à personalização psicológica. Também seria útil incluir medidas comportamentais ou delineamentos longitudinais para examinar a estabilidade dos efeitos ao longo do tempo.

10.
Referências

ABONIZIO, Hugo *et al.* **Sabiá-3 Technical Report**. arXiv, , 1 abr. 2025. Disponível em: <http://arxiv.org/abs/2410.12049>. Acesso em: 25 out. 2025

ALMEIDA, David M. Resilience and Vulnerability to Daily Stressors Assessed via Diary Methods. **Current Directions in Psychological Science**, v. 14, n. 2, p. 64–68, abr. 2005.

ARAÚJO, Tânia Maria De; TORRENTÉ, Mônica De Oliveira Nunes De. Mental Health in Brazil: challenges for building care policies and monitoring determinants. **Epidemiologia e Serviços de Saúde**, v. 32, n. 1, p. e2023098, 2023.

ARISTOTLE; REEVE, C. D. C. **Rhetoric**. Cambridge: Hackett Publishing Company, Incorporated, 2018.

BAI, Hui *et al.* LLM-generated messages can persuade humans on policy issues. **Nature Communications**, v. 16, n. 1, p. 6037, 1 jul. 2025.

BROWN, Tom B. *et al.* **Language Models are Few-Shot Learners**. arXiv, , 22 jul. 2020. Disponível em: <http://arxiv.org/abs/2005.14165>. Acesso em: 24 out. 2025

CARRELLA, Fabio *et al.* **Transparency is not enough: Warning people that they are being microtargeted fails to eliminate persuasive advantage**. PsyArXiv, , 5 set. 2024. Disponível em: <https://osf.io/akfsu>. Acesso em: 11 out. 2024

CASU, Mirko *et al.* AI Chatbots for Mental Health: A Scoping Review of Effectiveness, Feasibility, and Applications. **Applied Sciences**, v. 14, n. 13, p. 5889, 5 jul. 2024.

CHATTERJI, Aaron *et al.* How People Use ChatGPT. 2025.

CHAUDHRY, Beenish Moalla; DEBI, Happy Rani. User perceptions and experiences of an AI-driven conversational agent for mental health support. **mHealth**, v. 10, p. 22–22, jul. 2024.

COHEN, Jacob. QUANTITATIVE METHODS IN PSYCHOLOGY. 1992.

Conselho Federal de Psicologia. Disponível em: <https://site.cfp.org.br/wp-content/uploads/2025/09/SEI_CFP-2402857-Nota.pdf>. Acesso em: 22 nov. 2025.

DRUCKMAN, James N. A Framework for the Study of Persuasion. 2022.

DURDEN, Emily *et al.* Changes in stress, burnout, and resilience associated with an 8-week intervention with relational agent “Woebot”. **Internet Interventions**, v. 33, p. 100637, set. 2023.

FAUL, Franz *et al.* G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. **Behavior Research Methods**, v. 39, n. 2, p. 175–191, maio 2007.

FITZPATRICK, Kathleen Kara; DARCY, Alison; VIERHILE, Molly. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. **JMIR Mental Health**, v. 4, n. 2, p. e19, 6 jun. 2017.

FRANCO, Víthor Rosa; PRIMI, Ricardo. **O Futuro das Avaliações por Meio das Redes Sociais**. Disponível em: <https://pepsic.bvsalud.org/pdf/avp/v21n4/08.pdf>. Acesso em: 27 nov. 2025.

FRANSEN, Marieke L.; SMIT, Edith G.; VERLEGH, Peeter W. J. Strategies and motives for resistance to persuasion: an integrative framework. **Frontiers in Psychology**, v. 6, 14 ago. 2015.

FULMER, Russell *et al.* Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. **JMIR Mental Health**, v. 5, n. 4, p. e64, 13 dez. 2018.

HABICHT, Johanna *et al.* Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. **Nature Medicine**, v. 30, n. 2, p. 595–602, fev. 2024.

HACKENBURG, Kobi *et al.* **Comparing the persuasiveness of role-playing large language models and human experts on polarized U.S. political issues**. OSF, , 13 dez. 2023. Disponível em: <https://osf.io/ey8db>. Acesso em: 19 jun. 2024

HACKENBURG, Kobi *et al.* **The Levers of Political Persuasion with Conversational AI**. arXiv, , 18 jul. 2025. Disponível em: <http://arxiv.org/abs/2507.13919>. Acesso em: 24 jul. 2025

HACKENBURG, Kobi; MARGETTS, Helen. Evaluating the persuasive influence of political microtargeting with large language models. **Proceedings of the National Academy of Sciences**, v. 121, n. 24, p. e2403116121, 11 jun. 2024.

HU, Haoyang. The Need for Cognition as it Relates to Personality Traits of Openness to Experience and Conscientiousness. **BCP Education & Psychology**, v. 7, p. 241–244, 7 nov. 2022.

HU, Krystal. ChatGPT sets record for fastest-growing user base - analyst note. **Reuters**, 2 fev. 2023.

HUANG, Guanxiong; WANG, Sai. Is artificial intelligence more persuasive than humans? A meta-analysis. **Journal of Communication**, v. 73, n. 6, p. 552–562, 14 dez. 2023.

IBM. O que é inteligência LLM (grandes modelos de linguagem)? Disponível em: <https://www.ibm.com/br-pt/think/topics/large-language-models>. Acesso em: 13 nov. 2025.

IBM. O que é inteligência artificial (IA)? Disponível em: <https://www.ibm.com/br-pt/think/topics/artificial-intelligence>. Acesso em: 12 nov. 2025.

JOYAL-DESMARAIS, Keven. Do Message Matching Interventions Work. *[S.d.]*.

JOYAL-DESMARAIS, Keven; ROTHMAN, Alexander J.; SNYDER, Mark. Motivational Message Matching and the Functional Approach to Personalized Persuasion. *In*: **The Handbook of Personalized Persuasion**. *[S.l.]*: Routledge, 2025.

KARINSHAK, Elise *et al.* Working With AI to Persuade: Examining a Large Language Model's Ability to Generate Pro-Vaccination Messages. **Proceedings of the ACM on Human-Computer Interaction**, v. 7, n. CSCW1, p. 1–29, 14 abr. 2023.

KAYA, Feridun *et al.* The Roles of Personality Traits, AI Anxiety, and Demographic Factors in Attitudes toward Artificial Intelligence. **International Journal of Human–Computer Interaction**, v. 40, n. 2, p. 497–514, 17 jan. 2024.

KOSINSKI, Michal; STILLWELL, David; GRAEPEL, Thore. Private traits and attributes are predictable from digital records of human behavior. **Proceedings of the National Academy of Sciences**, v. 110, n. 15, p. 5802–5805, 9 abr. 2013.

LATIMER, Amy E. *et al.* Motivating Cancer Prevention and Early Detection Behaviors using Psychologically Tailored Messages. **Journal of Health Communication**, v. 10, n. sup1, p. 137–155, nov. 2005.

LATIMER, Amy E. *et al.* A field experiment testing the utility of regulatory fit messages for promoting physical activity. **Journal of Experimental Social Psychology**, v. 44, n. 3, p. 826–832, maio 2008.

LI, Cong. When does web-based personalization really work? The distinction between actual personalization and perceived personalization. **Computers in Human Behavior**, v. 54, p. 25–33, 1 jan. 2016.

LIBERATI, Alessandro *et al.* The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. **PLOS Medicine**, v. 6, n. 7, p. e1000100, 21 jul. 2009.

LORENZ-SPREEN, Philipp *et al.* Boosting people's ability to detect microtargeted advertising. **Scientific Reports**, v. 11, n. 1, p. 15541, 30 jul. 2021.

LU, Hang. Generative AI for vaccine misbelief correction: Insights from targeting extraversion and pseudoscientific beliefs. **Vaccine**, v. 54, p. 127018, abr. 2025.

LUSTRIA, Mia Liza A. *et al.* A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions. **Journal of Health Communication**, v. 18, n. 9, p. 1039–1069, 1 set. 2013.

LUTTRELL, Andrew; TEENY, Jacob D.; PETTY, Richard E. An Introduction to Personalized Persuasion. *In*: **The Handbook of Personalized Persuasion**. *[S.l.]*: Routledge, 2025.

MATZ, S. C. *et al.* Psychological targeting as an effective approach to digital mass persuasion. **Proceedings of the National Academy of Sciences**, v. 114, n. 48, p. 12714–12719, 28 nov. 2017.

MATZ, S. C. *et al.* The potential of generative AI for personalized persuasion at scale. **Scientific Reports**, v. 14, n. 1, p. 4692, 26 fev. 2024.

MATZ, Sandra C. *et al.* Personality Science in the Digital Age: The Promises and Challenges of Psychological Targeting for Personalized Behavior-Change Interventions at Scale. **Perspectives on Psychological Science**, p. 17456916231191774, 29 ago. 2023.

MATZ, Sandra C.; APPEL, Ruth E.; KOSINSKI, Michal. Privacy in the age of psychological targeting. **Current Opinion in Psychology**, v. 31, p. 116–121, fev. 2020.

MAURIELLO, Matthew Louis *et al.* A Suite of Mobile Conversational Agents for Daily Stress Management (Popbots): Mixed Methods Exploratory Study. **JMIR Formative Research**, v. 5, n. 9, p. e25294, 14 set. 2021.

MCCRAE, Robert R.; JOHN, Oliver P. An Introduction to the Five-Factor Model and Its Applications. **Journal of Personality**, v. 60, n. 2, p. 175–215, jun. 1992.

MCEWEN, Bruce S. Stress, Adaptation, and Disease: Allostasis and Allostatic Load. **Annals of the New York Academy of Sciences**, v. 840, n. 1, p. 33–44, maio 1998.

MEGUELLATI, Elyas *et al.* How Good are LLMs in Generating Personalized Advertisements? *In*: : WWW '24.New York, NY, USA: Association for Computing Machinery, Maio 2024. Disponível em: <https://doi.org/10.1145/3589335.3651520>. Acesso em: 8 jul. 2024

NEXUS. **51% avaliam: IA pode tomar decisões melhores que um ser humano**. **Nexus**, 20 out. 2025. Disponível em: <https://www.nexus.fsb.com.br/estudos-divulgados/51-dos-brasileiros-avaliam-que-ia-pode-tomar-decisoes-melhores-que-um-ser-humano-em-certas-situacoes/>. Acesso em: 6 nov. 2025

NOAR, Seth M.; BENAC, Christina N.; HARRIS, Melissa S. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. **Psychological Bulletin**, v. 133, n. 4, p. 673–693, 2007.

NUNES, Carlos Henrique; HUTZ, Claudio; NUNES, Maiana Farias Oliveira. Bateria Fatorial de Personalidade. 2010.

OAKLEY-GIRVAN, Ingrid *et al.* What Works Best to Engage Participants in Mobile App Interventions and e-Health: A Scoping Review. **Telemedicine and e-Health**, p. tmj.2021.0176, 11 out. 2021.

OPENAI. **GPT-4 Technical Report**. Disponível em: <https://cdn.openai.com/papers/gpt-4.pdf?utm_source=chatgpt.com>. Acesso em: 24 out. 2025.

OPENAI. **OpenAI**. Disponível em: <https://openai.com/pt-BR/index/introducing-gpt-5/>. Acesso em: 27 nov. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **World mental health today: Latest data**. Geneva: World Health Organization, 2025.

PALMER, Alexis; SPIRLING, Arthur. Large Language Models Can Argue in Convincing Ways About Politics, But Humans Dislike AI Authors: implications for Governance. **Political Science**, v. 75, n. 3, p. 281–291, 2 set. 2023.

PARK, Gregory *et al.* Automatic personality assessment through social media language. **Journal of Personality and Social Psychology**, v. 108, n. 6, p. 934–952, jun. 2015.

PETTY, Richard E.; CACIOPPO, John T. **Communication and persuasion: central and peripheral routes to attitude change**. New York Berlin Heidelberg: Springer, 1986.

PETTY, Richard E.; LUTTRELL, Andrew; TEENY, Jacob D. **The Handbook of Personalized Persuasion: Theory and Application**. *[S.l.]*: Routledge, 2024.

PETTY, Richard E.; WEGENER, Duane T.; FABRIGAR, Leandre R. ATTITUDES AND ATTITUDE CHANGE. **Annual Review of Psychology**, v. 48, n. Volume 48, 1997, p. 609–647, 1 fev. 1997.

PODSAKOFF, Philip M.; ORGAN, Dennis W. Self-Reports in Organizational Research: Problems and Prospects. **Journal of Management**, v. 12, n. 4, p. 531–544, dez. 1986.

RIEDL, René. Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. **Electronic Markets**, v. 32, n. 4, p. 2021–2051, dez. 2022.

ROGIERS, Alexander *et al.* PERSUASION WITH LARGE LANGUAGE MODELS: A SURVEY. 2024.

ROSS, Craig *et al.* Personality and motivations associated with Facebook use. **Computers in Human Behavior**, v. 25, n. 2, p. 578–586, mar. 2009.

ROTHMAN, Alexander J.; ROGERS, Maya G.; MANN, Traci. Using Message Matching Strategies to Promote Health: Opportunities and Challenges. *In*: **The Handbook of Personalized Persuasion**. *[S.l.]*: Routledge, 2025.

Saiba como usar corretamente o travessão, alvo de polêmica por IA. Disponível em: <https://www.cnnbrasil.com.br/educacao/saiba-como-usar-corretamente-o-travessao-alvo-de-polemica-por-ia/>. Acesso em: 27 out. 2025.

SALVI, Francesco *et al.* On the conversational persuasiveness of GPT-4. **Nature Human Behaviour**, v. 9, n. 8, p. 1645–1653, 19 maio 2025.

SHAHRIAR, Sakib *et al.* Putting GPT-4o to the Sword: A Comprehensive Evaluation of Language, Vision, Speech, and Multimodal Proficiency. 2024.

SIMCHON, Almog; EDWARDS, Matthew; LEWANDOWSKY, Stephan. The persuasive effects of political microtargeting in the age of generative artificial intelligence. **PNAS Nexus**, v. 3, n. 2, p. pgae035, 1 fev. 2024.

SMITH, Aaron. **Algorithms in action: The content people see on social media**. **Pew Research Center**, 16 nov. 2018. Disponível em: <https://www.pewresearch.org/internet/2018/11/16/algorithms-in-action-the-content-people-see-on-social-media/>. Acesso em: 20 nov. 2025

SNYDER, Hannah. Literature review as a research methodology: An overview and guidelines. **Journal of Business Research**, v. 104, p. 333–339, 1 nov. 2019.

STEIN, Jan-Philipp *et al.* Attitudes towards AI: measurement and associations with personality. **Scientific Reports**, v. 14, n. 1, p. 2909, 5 fev. 2024.

TAPPIN, Ben M. *et al.* Quantifying the potential persuasive returns to political microtargeting. **Proceedings of the National Academy of Sciences**, v. 120, n. 25, p. e2216261120, 20 jun. 2023.

TEENY, Jacob D. *et al.* A Review and Conceptual Framework for Understanding Personalized Matching Effects in Persuasion. **Journal of Consumer Psychology**, v. 31, n. 2, p. 382–414, 2021.

TEENY, Jacob D.; LUTTRELL, Andrew; PETTY, Richard E. The Present and Future Landscape of Personalized Persuasion. *In*: **The Handbook of Personalized Persuasion**. *[S.l.]*: Routledge, 2025.

TEENY, Jacob D.; MATZ, Sandra C. We need to understand "when" not "if" generative AI can enhance personalized persuasion. **Proceedings of the National Academy of Sciences**, v. 121, n. 43, p. e2418005121, 22 out. 2024.

THE JAMOVI PROJECT. **Jamovi**, 2025. Disponível em: <https://www.jamovi.org>

VACCARO, Michelle; ALMAATOUQ, Abdullah; MALONE, Thomas. When combinations of humans and AI are useful: A systematic review and meta-analysis. **Nature Human Behaviour**, v. 8, n. 12, p. 2293–2303, 28 out. 2024.

VARGHESE, Mahima Anna; SHARMA, Poonam; PATWARDHAN, Maitreyee. Public Perception on Artificial Intelligence–Driven Mental Health Interventions: Survey Research. **JMIR Formative Research**, v. 8, p. e64380, 28 nov. 2024.

VOELKEL, Jan G.; FEINBERG, Matthew. Morally Reframed Arguments Can Affect Support for Political Candidates. **Social Psychological and Personality Science**, v. 9, n. 8, p. 917–924, nov. 2018.

WANG, Yilun; KOSINSKI, Michal. Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation From Facial Images. **Journal of Personality and Social Psychology**, v. 114, p. 246–257, 1 fev. 2018.

WILLIAMS-PIEHOTA, Pamela *et al.* Matching Health Messages to Information-Processing Styles: Need for Cognition and Mammography Utilization. **Health Communication**, v. 15, n. 4, p. 375–392, out. 2003.

World Bank Group. Disponível em: <https://data.worldbank.org>. Acesso em: 21 nov. 2025.

YOUYOU, Wu; KOSINSKI, Michal; STILLWELL, David. Computer-based personality judgments are more accurate than those made by humans. **Proceedings of the National Academy of Sciences**, v. 112, n. 4, p. 1036–1040, 27 jan. 2015.

ZAROUALI, Brahim *et al.* Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media. **Communication Research**, v. 49, n. 8, p. 1066–1091, 1 dez. 2022.

Anexos

Anexo A – Questionário Sociodemográfico

Instruções para os participantes:

Esse é um questionário sociodemográfico. Assinale a opção que corresponda melhor a você para cada questão.

Idade

1. **Qual é a sua idade?**
    a. ( ) Menos de 18 anos
    b. ( ) 18 a 24 anos
    c. ( ) 25 a 34 anos
    d. ( ) 35 a 44 anos
    e. ( ) 45 a 54 anos
    f. ( ) 55 a 64 anos
    g. ( ) 65 anos ou mais

2. **Gênero**
    Com qual gênero você se identifica?
    a. ( ) Masculino
    b. ( ) Feminino
    c. ( ) Não-binário
    d. ( ) Prefiro não responder

3. **Curso de Graduação ou Pós-graduação**
    Em qual curso você está matriculado(a)?
    a. ( ) Graduação (em andamento)
    b. ( ) Mestrado (em andamento)
    c. ( ) Doutorado (em andamento)
    d. ( ) Outro (especificar): _____

4. **Curso da Faculdade**
    Qual o seu curso de graduação ou pós-graduação?
    a. ( ) Administração
    b. ( ) Ciência da Computação
    c. ( ) Engenharia
    d. ( ) Pedagogia
    e. ( ) Psicologia
    f. (lista de todos cursos de graduação ou pós-graduação)

Anexo B – Escala de Atitudes sobre Chatbots

Instruções para os participantes:

Agora que você leu a mensagem sobre chatbots para gerenciamento de estresse diário, gostaríamos de saber sua opinião sobre essa tecnologia.

A seguir, você encontrará algumas afirmações e perguntas relacionadas ao uso de chatbots para essa finalidade. Leia cada item com atenção e registre sua resposta de acordo com as instruções abaixo. Não há respostas certas ou erradas. O importante é que suas respostas representem sua opinião sincera. Responda a todos os itens sem pular nenhum.

Para os itens abaixo, avalie o quanto você concorda ou discorda de cada afirmação, usando a seguinte escala:

- **1** = Discordo totalmente
- **2** = Discordo parcialmente
- **3** = Neutro
- **4** = Concordo parcialmente
- **5** = Concordo totalmente

1. O uso de chatbots para gerenciamento de estresse diário é uma boa ideia	
2. Deveríamos promover o uso de chatbots para gerenciamento de estresse diário	
3. O uso de chatbots para gerenciamento de estresse diário traz boas consequências	
4. Não me sentiria seguro(a) em compartilhar meus problemas pessoais com um chatbot de gerenciamento de estresse	
5. O sistema público de saúde poderia usar chatbots para complementar os atendimentos presenciais no gerenciamento de estresse	
6. O uso de chatbots para gerenciamento de estresse é seguro e não traz riscos para os usuários	
7. Eu apoio o uso de chatbots para o gerenciamento de estresse diário	
8. Eu votaria a favor da promoção do uso de chatbots para gerenciamento de estresse se houvesse um referendo.	
9. Eu pretendo utilizar chatbots para lidar com meu estresse diário no futuro próximo.	

Anexo C – Escala Breve de Abertura à Tecnologia

Instruções para os participantes:

Nesta seção, você encontrará algumas afirmações sobre seu envolvimento com tecnologia. Leia cada uma delas e indique o quanto você se identifica, utilizando a escala abaixo:

De 1 a 5, onde:

1 = Discordo totalmente
2 = Discordo parcialmente
3 = Nem concordo, nem discordo
4 = Concordo parcialmente
5 = Concordo totalmente

Agora, avalie as afirmações a seguir e, por favor, marque a opção que melhor representa sua opinião para cada afirmação.

1. Eu costumo adotar novas tecnologias antes da maioria das pessoas.	
2. Gosto de estar a par dos avanços tecnológicos que poderão tornar-se úteis no meu dia a dia.	
3. Tenho facilidade e conforto em usar tecnologias digitais no meu dia a dia.	

Anexo D – Explicação Breve Sobre o Estímulo Textual e Estímulos Textuais

Na próxima página, você lerá um breve texto sobre o uso de chatbots para o gerenciamento do estresse diário. Leia com atenção, pois depois você irá responder algumas perguntas sobre suas percepções.

Chatbots são programas de computador que simulam interações humanas para oferecer suporte e informações. No contexto deste estudo, o chatbot fornece orientações para o gerenciamento de estresse diário, como sugestões de técnicas de relaxamento e dicas de bem-estar.

Não há respostas certas ou erradas, apenas queremos entender suas percepções.

Instruções para os participantes:

A seguir, você lerá um breve texto sobre o uso de chatbots para gerenciamento de estresse diário. Leia com atenção, pois depois você responderá a algumas perguntas sobre ele. Não há respostas certas ou erradas, apenas queremos entender suas percepções.

Após a leitura, prossiga para a próxima etapa do estudo.

Estímulos textuais:

Mensagem Personalizada para Alta Abertura à Experiência:

Você já reparou como, às vezes, só de escrever o que estamos sentindo, as coisas ficam mais claras? É curioso como a mente encontra caminhos quando tem espaço para experimentar, sem julgamentos. Tem gente que usa diários, outros gravam áudios pra si mesmos… e há quem converse com um chatbot. Não é sobre respostas prontas, mas sobre brincar com ideias, testar possibilidades, escutar os próprios pensamentos ganhando forma. Pode parecer inusitado no começo, mas muitos acabam descobrindo ali um jeito novo — e surpreendentemente íntimo — de organizar emoções, refletir ou até criar. É como um caderno que responde, sem pressa e sem regras. Só você, suas palavras, e a liberdade de explorar o que vier. Não é terapia, não substitui gente de verdade, mas pode ser um pequeno laboratório mental para quem gosta de pensar diferente. Afinal, por que não experimentar uma nova forma de conversar consigo mesmo?

Mensagem Personalizada para Baixa Abertura à Experiência:

Às vezes, o dia parece puxado demais. Entre tarefas, compromissos e preocupações, a cabeça não pára, e o estresse vai se acumulando sem a gente perceber. Nessas horas, ajuda ter à mão algo simples e confiável. Os chatbots funcionam assim: são diretos, fáceis de usar e ajudam a colocar as ideias no lugar ou até a desabafar um pouco, sem complicação. Não é preciso aprender nada novo, nem mudar a rotina. É como fazer uma lista ou anotar um lembrete — só que com respostas na hora, quando você precisa. E o melhor: estão sempre disponíveis, sem precisar marcar hora ou esperar. Para quem prefere praticidade no dia a dia, é uma forma tranquila de aliviar a mente sem mudar o que já funciona. Às vezes, o que mais ajuda é justamente o que não exige muito da gente — e, só de dar esse respiro, o dia já fica mais leve.

Mensagem Genérica (Controle):

Sabe aqueles dias em que tudo parece estar dando errado, e a mente fica a mil? Nessas horas, ter um espaço para soltar um pouco do peso pode fazer diferença. Não precisa ser algo grandioso ou complicado – às vezes, só conversar, nem que seja com um chatbot, já ajuda. Ele não julga, não interrompe, e está sempre disponível. Pode parecer estranho no começo, mas escrever ou até mesmo falar com ele sobre o que está passando na sua cabeça pode ser um jeito de organizar os pensamentos e aliviar um pouco do estresse. É como se a gente tirasse as palavras de dentro e colocasse pra fora, mesmo que ninguém esteja "ouvindo" de verdade. Vale a pena experimentar quando sentir aquele aperto no peito ou a cabeça cheia demais. Afinal, um pouco de leveza no dia a dia nunca faz mal, não é?"

Anexo E – Validação dos Estímulos Textuais

Olá! Estamos conduzindo uma validação prévia de mensagens para um experimento sobre o uso de chatbots em diferentes contextos. Você verá três mensagens curtas, cada uma descrevendo um tipo de chatbot: duas voltadas para gerenciamento de estresse diário e uma voltada para treinamento de idiomas. As mensagens foram elaboradas com o objetivo de encorajar o uso desses chatbots, cada uma com um estilo de comunicação diferente. Essa avaliação é essencial para verificar se as mensagens transmitem corretamente a intenção original, garantindo a validade do experimento. O tempo estimado é de 5 minutos. Muito obrigado pela sua participação.

Uso e Privacidade dos Dados: Os dados coletados destinam-se estritamente a atividades de pesquisa e nos comprometemos a assegurar a privacidade das pessoas envolvidas, bem como garantimos que as informações coletadas não serão utilizadas em prejuízo das mesmas. A participação na pesquisa é voluntária e não envolve nenhum tipo de remuneração entre as partes. O envio do formulário será considerado como declaração de aceite das condições aqui citadas.

As duas perguntas a seguir são apenas para caracterizar a amostra da pesquisa. Todas as respostas são anônimas e serão utilizadas exclusivamente para fins de análise agregada. Com qual gênero você se identifica?

1. Masculino
2. Feminino
3. Não-binário
4. Prefiro não responder
5. Outro (especificar)
6.

Qual é o seu nível mais alto de formação em Psicologia?

1. Graduação completa
2. Especialização (lato sensu) completa
3. Mestrado completo
4. Doutorado completo
5. Estudante de graduação em Psicologia
6. Estudante de especialização em Psicologia
7. Estudante de mestrado em Psicologia
8. Estudante de doutorado em Psicologia
9. Outro (especificar)

As próximas duas mensagens descrevem diferentes formas de apresentar um chatbot voltado ao gerenciamento de estresse no dia a dia. Para cada uma, pedimos que você leia com atenção o parágrafo apresentado. Em seguida, será mostrada uma escala para que você avalie o quanto cada traço de personalidade do modelo Big Five parece presente na mensagem, de "Muito Baixo" a "Muito Alto". Como o tema dos parágrafos é o uso de chatbots para lidar com o estresse, evite atribuir o nível do traço com base apenas em termos ligados à preocupação ou tensão emocional.

Por favor, leia com atenção o parágrafo a seguir:

“Você já reparou como, às vezes, só de escrever o que estamos sentindo, as coisas ficam mais claras? É curioso como a mente encontra caminhos quando tem espaço para experimentar, sem julgamentos. Tem gente que usa diários, outros gravam áudios pra si mesmos… e há quem converse com um chatbot. Não é sobre respostas prontas, mas sobre brincar com ideias, testar possibilidades, escutar os próprios pensamentos ganhando forma. Pode parecer inusitado no começo, mas muitos acabam descobrindo ali um jeito novo — e surpreendentemente íntimo — de organizar emoções, refletir ou até criar. É como um caderno que responde, sem pressa e sem regras. Só você, suas palavras, e a liberdade de explorar o que vier. Não é terapia, não substitui gente de verdade, mas pode ser um pequeno laboratório mental para quem gosta de pensar diferente. Afinal, por que não experimentar uma nova forma de conversar consigo mesmo?”

Com base na mensagem acima, indique o nível em que cada traço de personalidade parece estar presente. O que deve ser avaliado é: para que tipo de pessoa essa mensagem parece ter sido escrita? Ou seja, com base na linguagem, no tom e nos argumentos utilizados, qual é o perfil psicológico da pessoa a quem essa mensagem foi dirigida? Para cada traço listado abaixo, indique o nível que você acredita caracterizar a pessoa para quem essa mensagem parece ter sido direcionada.

	Muito Baixo	Baixo	Neutro	Alto	Muito Alto
Abertura à Experiência					
Amabilidade					
Conscienciosidade					
Extroversão					
Neuroticismo					

Por favor, leia com atenção o parágrafo a seguir:

“Às vezes, o dia parece puxado demais. Entre tarefas, compromissos e preocupações, a cabeça não pára, e o estresse vai se acumulando sem a gente perceber. Nessas horas, ajuda ter à mão algo simples e confiável. Os chatbots funcionam assim: são diretos, fáceis de usar e ajudam a colocar as ideias no lugar ou até a desabafar um pouco, sem complicação. Não é preciso aprender nada novo, nem mudar a rotina. É como fazer uma lista ou anotar um lembrete — só que com respostas na hora, quando você precisa. E o melhor: estão sempre disponíveis, sem precisar marcar hora ou esperar. Para quem prefere praticidade no dia a dia, é uma forma tranquila de aliviar a mente sem mudar o que já funciona. Às vezes, o que mais ajuda é justamente o que não exige muito da gente — e, só de dar esse respiro, o dia já fica mais leve.”

Com base na mensagem acima, indique o nível em que cada traço de personalidade parece estar presente. ATENÇÃO: O que deve ser avaliado é: para que tipo de pessoa essa mensagem parece ter sido escrita? Ou seja, com base na linguagem, no tom e nos argumentos utilizados, qual é o perfil psicológico da pessoa a quem essa mensagem foi dirigida? Para cada traço listado abaixo, indique o nível que você acredita caracterizar a pessoa para quem essa mensagem parece ter sido direcionada.

	Muito Baixo	Baixo	Neutro	Alto	Muito Alto
Abertura à Experiencia					
Amabilidade					
Conscienciosidade					
Extroversão					
Neuroticismo					

A terceira e última mensagem descreve um chatbot com uma finalidade diferente, voltada ao treinamento de idiomas. Ainda assim, pedimos que você avalie da mesma forma. Para cada traço, indique o nível percebido, de "Muito Baixo" a "Muito Alto". “Aprender um novo idioma pode parecer complicado — e muitas vezes é mesmo. Mas a verdade é que não precisa ser algo fora da sua rotina ou que exija grandes mudanças. Usar um chatbot para treinar uma língua, por exemplo, pode ser uma forma prática e discreta de melhorar, sem sair do seu ritmo. Você conversa no seu tempo, sem pressa, sem julgamentos, e vai ganhando confiança aos poucos. É como praticar com alguém que está sempre disponível, mas que não cobra desempenho. E o melhor: você não precisa fazer nada de muito diferente — só incluir pequenas conversas no dia a dia, do jeito que já está acostumado. Mesmo que pareça estranho no começo, é uma forma segura e

constante de progredir. Às vezes, manter o que funciona e acrescentar uma ferramenta simples já é o bastante pra fazer diferença.”

Com base na mensagem acima, indique o nível em que cada traço de personalidade parece estar presente.

	Muito Baixo	Baixo	Neutro	Alto	Muito Alto
Abertura à Experiencia					
Amabilidade					
Conscienciosidade					
Extroversão					
Neuroticismo					

Muito obrigado pela sua colaboração! Deixe aqui quaisquer observações que desejar sobre as mensagens apresentadas.

Anexo F – Termo de Consentimento Livre e Esclarecido

Título do Projeto: "Percepções sobre Chatbots para Gerenciamento de Estresse"

Você está sendo convidado(a) a participar da pesquisa intitulada "Percepções sobre Chatbots para Gerenciamento de Estresse", sob a responsabilidade do pesquisador Bruno Grossman, aluno de mestrado do curso de Psicologia da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), sob a orientação do Professor Leonardo Fernandes Martins.

Com as interações cada vez mais mediadas por tecnologias digitais, ferramentas como chatbots para ajudar no gerenciamento do estresse estão se tornando mais comuns. Mas como essas tecnologias são percebidas pelas pessoas? Esta pesquisa busca entender as primeiras impressões sobre essa tecnologia e como diferentes aspectos da comunicação influenciam essa percepção.

A pesquisa consiste num questionário inteiramente online, com duração aproximada de 15 minutos. O questionário incluirá perguntas gerais sobre seu perfil, seguidas da apresentação de uma mensagem relacionada ao uso de chatbots para gerenciamento de estresse diário. Em seguida, você será solicitado(a) a compartilhar suas percepções e opiniões sobre a mensagem apresentada. Os dados coletados serão utilizados exclusivamente para fins acadêmicos.

A pesquisa será realizada com estudantes universitários maiores de 18 anos, falantes nativos de português. Serão excluídos participantes que não completarem o questionário ou que não atenderem aos critérios mínimos de atenção estabelecidos no estudo.

Todas as medidas de segurança são adotadas para proteger as informações coletadas, incluindo armazenamento em dispositivos protegidos por senha e autenticação biométrica. No entanto, por se tratar de uma pesquisa conduzida em ambiente virtual, existem riscos inerentes ao uso de tecnologias digitais, como eventuais falhas de segurança ou acessos não autorizados. Embora a possibilidade de vazamento de dados seja remota, todas as precauções técnicas são empregadas para prevenir esse risco. Caso ocorra, você será notificado, e medidas corretivas serão tomadas para mitigar eventuais impactos pelos mesmos canais utilizados para recrutamento dos participantes, incluindo grupos acadêmicos e redes sociais.

A pesquisa envolve a leitura de textos e o preenchimento de questionários. Embora seja um procedimento simples, você pode sentir desconforto ao refletir sobre os conteúdos apresentados. Para minimizar esse efeito, foram utilizados instrumentos validados em estudos anteriores, e ao final do experimento será apresentado uma explicação detalhada dos objetivos da pesquisa.

Não há benefícios diretos para você, como compensação financeira ou retorno imediato. No entanto, a sua participação contribuirá indiretamente para o avanço do conhecimento sobre como as pessoas percebem mensagens sobre o uso de chatbots para gerenciamento de estresse. Os resultados poderão orientar práticas de comunicação mais eficazes e éticas, especialmente para o bem-estar digital.

O seu sigilo e a privacidade serão garantidos em todas as etapas da pesquisa. Nenhuma informação pessoal que possa identificar os participantes, como nome, voz ou imagem, será coletada. Os dados serão armazenados de forma segura e utilizados exclusivamente para fins acadêmicos. Apenas a equipe responsável pelo estudo terá acesso às informações, respeitando todas as normas éticas de confidencialidade. Em nenhuma circunstância os dados serão compartilhados com terceiros. Todos os dados coletados nesta pesquisa ficarão armazenados em arquivo digital, sob guarda e responsabilidade do pesquisador, por um período mínimo de 5 (cinco) anos após o término da pesquisa. Para garantir a segurança das informações, os dados serão inicialmente armazenados em ambiente protegido e, na maior brevidade possível após a coleta, serão transferidos para um dispositivo eletrônico próprio/local, sendo então removidos de qualquer serviço de armazenamento em nuvem.

Caso você opte por interromper sua participação, suas respostas serão excluídas do estudo. Todas as medidas necessárias serão adotadas para evitar qualquer quebra de sigilo ou uso indevido das informações coletadas.

Você terá acesso aos resultados da pesquisa em um formato claro e compreensível. Ao final do estudo, um resumo dos principais achados será disponibilizado em linguagem acessível, garantindo que as conclusões da pesquisa sejam compreensíveis. Caso tenha interesse, você poderá solicitar esse resumo diretamente ao pesquisador responsável.

A participação na pesquisa não envolve qualquer tipo de custo financeiro para você, assim como não há previsão de ressarcimento ou compensação financeira.

A pesquisa não prevê acompanhamento posterior, pois se trata de um questionário pontual e anônimo. Caso o participante tenha dúvidas ou queira mais informações sobre o estudo, poderá entrar em contato com o pesquisador responsável através dos dados na seção “Contatos” nesse documento.

No evento que a pesquisa seja interrompida por qualquer motivo, você será informado sobre a justificativa da decisão.

Participação Voluntária e Direito de Desistência

A participação nesta pesquisa é totalmente voluntária. Você é livre para decidir se deseja ou não participar, podendo recusar-se a responder ao questionário ou desistir a qualquer momento e em qualquer fase da pesquisa, sem necessidade de justificativa e sem qualquer prejuízo. Em caso de desistência, seus dados serão imediatamente excluídos do estudo.

Você poderá, a qualquer momento, cancelar sua participação e/ou solicitar a retirada dos seus dados da pesquisa. Para isso, deverá entrar em contato com o pesquisador responsável utilizando os dados de correspondência fornecidos na seção "Contatos" deste documento. Após a solicitação, seus dados fornecidos serão removidos do banco de dados da pesquisa no prazo de 10 dias úteis, exceto nos casos em que os dados já tenham sido anonimizados ou incorporados a análises estatísticas impossibilitando sua exclusão. Caso o pedido de retirada não possa ser atendido por razões metodológicas, essa limitação será informada a você.

Este estudo foi revisado e aprovado pela Sistema CEP/Conep conforme as resoluções brasileiras que regulamentam pesquisa com seres humanos no Brasil: RESOLUÇÃO No 466, de 12 de DEZEMBRO de 2012, CONSELHO NACIONAL DE SAÚDE/MINISTÉRIO DA SAÚDE. RESOLUÇÃO No 510, de 7 de ABRIL de 2016, CONSELHO NACIONAL DE SAÚDE/MINISTÉRIO DA SAÚDE.

Contatos: Se desejar maiores esclarecimentos você pode também nos contactar a qualquer momento no e-mail bgrossman@puc-rio.br, ou pelo endereço Marquês de São Vicente, 225, Edifício Cardeal Leme, 2º Andar - Sala 201, Gávea, Rio de Janeiro, RJ CEP: 22451-900 e número de telefone 21 3736-1185., bem como contactar a respectiva Câmara de Ética da PUC-Rio: Câmara de Ética em Pesquisa da PUC-Rio: Rua Marquês de São Vicente, 225 – Edifício Kennedy, 2º andar. Gávea, Rio de Janeiro, RJ. CEP: 22453-900. Fone: (21) 3527-1618.

Recomendação: recomendamos que você faça o download e guarde uma cópia deste Termo de Consentimento Livre e Esclarecido em seus arquivos pessoais, para que possa consultá-lo sempre que necessário.

Declaração de Consentimento

Declaro que li e compreendi as informações contidas neste Termo de Consentimento Livre e Esclarecido. Estou ciente de que minha participação é voluntária, sem compensação financeira, e que posso desistir a qualquer momento, sem prejuízo. Autorizo, de forma livre e esclarecida, minha participação na pesquisa descrita neste documento.

[ ] Aceito participar

(Caso não queira participar, basta fechar esta página.)

Bruno Grossman

Anexo G – Debriefing

Este experimento investiga como a personalização de mensagens pode influenciar a aceitação de chatbots para gerenciamento de estresse. Os participantes são divididos em três grupos experimentais. Dois deles recebem mensagens personalizadas com base no traço de personalidade Abertura à Experiência: um grupo recebe mensagens ajustadas para indivíduos com alta Abertura, enquanto outro recebe mensagens voltadas para baixa Abertura. O terceiro grupo recebe uma mensagem genérica, sem qualquer personalização.

Todas as mensagens são elaboradas pelo modelo de inteligência artificial ChatGPT-4o, seguindo instruções específicas para adaptar o conteúdo ao perfil psicológico dos participantes. Nosso objetivo é compreender se a adaptação da mensagem ao perfil do receptor pode tornar a persuasão mais eficaz, especialmente no contexto da adoção de chatbots como ferramentas de suporte ao gerenciamento de estresse diário.

Os dados são coletados de forma anônima, sem qualquer associação com sua identidade, e tratados com total confidencialidade, conforme as normas éticas da pesquisa. Ressaltamos que este estudo não implica evidências de que chatbots para gerenciamento de estresse sejam eficazes ou recomendados como substitutos para apoio profissional especializado. O objetivo é apenas investigar como diferentes formas de comunicação podem influenciar a percepção sobre essas tecnologias.

É natural que, ao longo da vida, passemos por eventos que geram estresse. No entanto, quando isso se torna um problema, causando sofrimento ou prejuízos significativos, buscar ajuda profissional é fundamental. Se esse for o seu caso, recomendamos que procure um profissional de saúde de sua confiança para conversar sobre isso. Caso ainda não tenha um profissional de referência, procure a Unidade Básica de Saúde mais próxima da sua residência, onde poderá encontrar profissionais preparados para lhe oferecer apoio. Se você estiver passando por um momento difícil, seja por esse ou outros motivos, e tiver pensamentos relacionados à morte ou ao suicídio, ligue agora mesmo para o Centro de Valorização da Vida (CVV), pelo telefone número 188. A ligação é gratuita, e você encontrará alguém para te ouvir e orientar.

Agradecemos novamente sua colaboração, que é essencial para o avanço do conhecimento sobre personalização e novas formas de suporte à saúde mental. Sua participação contribui para pesquisas que podem tornar as estratégias de comunicação mais eficazes e acessíveis para todos. Se desejar retirar sua participação sem qualquer prejuízo ou tiver quaisquer dúvidas, entre em contato pelo e-mail bgrossman@puc-rio.br ou pelo endereço Marquês de São Vicente, 225, Edifício Cardeal Leme, 2º Andar - Sala 201, Gávea, Rio de Janeiro, RJ CEP: 22451-900 e número de telefone 21 3736-1185.