
Da empatia ao algoritmo: a cognição social frente à inteligência artificial

*From empathy to algorithm:
social cognition in the face of artificial intelligence*

*De la empatía al algoritmo:
la cognición social frente a la inteligencia artificial*

1 Tiago Fernando Figueiredo Santos  [ORCID](#) - [Lattes](#)

2 Leonardo Caixeta - [ORCID](#) - [Lattes](#)

Filiação dos autores: 1 [Pesquisador Chefe do Instituto Cognus de Ensino e Pesquisa, ICEP, Brasília, DF, Brasil]; **2** [Diretor do CERNE, Centro de Referência em Neuropsiquiatria, Hospital das Clínicas, UFG, Goiânia, GO, Brasil]

Editor Chefe responsável pelo artigo: Marsal Sanches

Contribuição dos autores segundo a [Taxonomia CRediT](#): Santos TFF [1,2,3,4,5,6,7,8]; Caixeta L [1,2,3,4,5,6]

Conflito de interesses: declaram não haver

Fonte de financiamento: não se aplica

Parecer CEP: não se aplica

Recebido em: 17/09/2025

Aprovado em: 26/11/2025

Publicado em: 13/12/2025

Como citar: Santos TFF, Caixeta L. Da empatia ao algoritmo: a cognição social frente à inteligência artificial. Debates Psiquiatr. 2025;15:1-23.

<https://doi.org/10.25118/2763-9037.2025.v15.1504>

RESUMO:

Introdução: O avanço da inteligência artificial (IA) tem remodelado profundamente as interações humanas, suscitando questões sobre sua relação com a cognição social, a capacidade de perceber, interpretar e responder a estados mentais alheios. **Objetivo:** Analisar a relação bidirecional entre cognição social e IA, abordando os desafios na modelagem computacional da inteligência social humana e os potenciais impactos da interação com máquinas inteligentes sobre o enfrentamento ou agravamento da solidão. **Método:** Artigo de opinião fundamentado em revisão narrativa da literatura, com análise de estudos empíricos e teóricos sobre neurociência social, interação humano-robô e manejo da solidão por meio de tecnologias de IA. **Resultados:** A complexidade da cognição social humana permanece como principal desafio para naturalização das IAs. Evidências neurocientíficas revelam padrões distintos de ativação cerebral na interação com agentes corporificados versus não-corporificados, sugerindo hierarquia nas funções cognitivas sociais. Tecnologias de IA demonstram potencial terapêutico no manejo da solidão, mas também riscos de amplificação desse fenômeno. **Conclusão:** A análise da interação humano-IA fornece insights valiosos sobre os mecanismos neurais da cognição social, evidenciando que o processo imaginativo interpessoal permanece como característica distintivamente humana. O uso clínico de IAs requer consideração cuidadosa de suas limitações e potenciais efeitos paradoxais.

Palavras-chave: cognição social, inteligência artificial, interação humano-máquina, solidão, psiquiatria social

ABSTRACT:

Introduction: The advancement of artificial intelligence (AI) has profoundly reshaped human interactions, raising questions about its relationship with social cognition, the ability to perceive, interpret, and respond to others' mental states. **Objective:** To analyze the bidirectional relationship between social cognition and AI, addressing challenges in computational modeling of human social intelligence and potential impacts of interaction with intelligent machines on coping with or exacerbating loneliness. **Method:** Opinion article based on narrative literature review, analyzing empirical and theoretical studies on social neuroscience, human-robot interaction, and loneliness management through AI technologies. **Results:** The complexity of human social cognition remains the main challenge for AI naturalization. Neuroscientific evidence reveals distinct patterns of brain activation in interactions with embodied versus non-

embodied agents, suggesting hierarchy in social cognitive functions. AI technologies demonstrate therapeutic potential in managing loneliness but also risks of amplifying this phenomenon. **Conclusion:** Analysis of human-AI interaction provides valuable insights into neural mechanisms of social cognition, showing that interpersonal imaginative processes remain distinctively human. Clinical use of AI requires careful consideration of its limitations and potential paradoxical effects.

Keywords: social cognition; artificial intelligence; human-machine interaction, loneliness, social psychiatry

RESUMEN:

Introducción: El avance de la inteligencia artificial (IA) ha remodelado profundamente las interacciones humanas, planteando cuestiones sobre su relación con la cognición social, la capacidad de percibir, interpretar y responder a estados mentales ajenos. **Objetivo:** Analizar la relación bidireccional entre cognición social e IA, abordando los desafíos en el modelado computacional de la inteligencia social humana y los potenciales impactos de la interacción con máquinas inteligentes sobre el afrontamiento o agravamiento de la soledad. **Método:** Artículo de opinión fundamentado en revisión narrativa de la literatura, con análisis de estudios empíricos y teóricos sobre neurociencia social, interacción humano-robot y manejo de la soledad mediante tecnologías de IA. **Resultados:** La complejidad de la cognición social humana permanece como principal desafío para la naturalización de las IAs. Evidencias neurocientíficas revelan patrones distintos de activación cerebral en la interacción con agentes corporeizados versus no corporeizados, sugiriendo jerarquía en las funciones cognitivas sociales. Las tecnologías de IA demuestran potencial terapéutico en el manejo de la soledad, pero también riesgos de amplificación de este fenómeno. **Conclusión:** El análisis de la interacción humano-IA proporciona conocimientos valiosos sobre los mecanismos neurales de la cognición social, evidenciando que el proceso imaginativo interpersonal permanece como característica distintivamente humana. El uso clínico de IAs requiere consideración cuidadosa de sus limitaciones y potenciales efectos paradójicos.

Palabras clave: cognición social, inteligencia artificial, interacción humano-máquina, soledad, psiquiatría social

"Não me ajeito com os padres, os críticos e os canudinhos de refresco: não há nada que substitua o sabor da comunicação direta" (Mario Quintana)

Introdução

O fascínio e a busca humana pela criação de um eu mecânico remonta à antiguidade, com escritores da Grécia e da China conjurando autômatos semelhantes a humanos para servir como seus trabalhadores e servos [1]. No século passado, o tipo de autômato que mais cativou a imaginação humana (e o investimento em pesquisa e desenvolvimento) foram os robôs, com alguns modelos contemporâneos se aproximando dos ideais ficcionais que surgiram há séculos. Concomitantemente aos avanços na tecnologia robótica, ocorreu o advento e o rápido desenvolvimento da tecnologia aplicada aos métodos de imagens cerebrais em estudos neurocientíficos. Esses métodos de aferição da dinâmica cerebral têm sido essenciais para o aprofundamento na compreensão dos mecanismos cognitivos que sustentam o comportamento social humano.

O avanço das tecnologias digitais e, em especial, com a implementação das ferramentas de inteligência artificial (IA), vem remodelando de maneira profunda a forma como os seres humanos interagem entre si e com as máquinas. A emergência de sistemas capazes de reconhecer padrões, responder em linguagem natural e até simular aspectos de tomada de decisão coloca em pauta uma questão central: como a cognição social, ou seja, a capacidade humana de perceber, interpretar e responder a estados mentais e comportamentos alheios, se relaciona com a presença crescente da IA em contextos cotidianos. Nesse cenário, não apenas se abrem desafios técnicos para os programadores que buscam simular a complexidade da inteligência social humana, como também emergem questões sobre os efeitos que o contato contínuo com sistemas inteligentes pode ter sobre o próprio funcionamento cognitivo dos indivíduos.

O objetivo deste artigo de opinião é analisar, à luz de evidências empíricas atuais, a relação bidirecional entre cognição social e inteligência artificial: de um lado, os obstáculos enfrentados para modelar, em linguagem computacional, os processos que sustentam a inteligência social humana; de outro, os potenciais impactos que a interação com máquinas inteligentes pode exercer sobre o manejo ou, paradoxalmente, o agravamento da solidão. Os obstáculos percebidos na modelação e naturalização da IA têm revelado insights importantes sobre os mecanismos envolvidos no comportamento social. Do outro lado, o estudo das experiências dos usuários que buscam a IA para o enfrentamento do

sofrimento mental tem acelerado o crescimento da base empírica sobre os fenômenos mentais envolvidos na percepção do suporte social.

Método

Este trabalho configura-se como um artigo de opinião fundamentado em revisão narrativa da literatura. Foram revisados estudos empíricos e teóricos publicados nas bases de dados [PubMed](#), [PsycINFO](#) e [Web of Science](#), utilizando os descritores "cognição social", "inteligência artificial", "neurociência social", "interação humano-robô", "solidão" e suas combinações em inglês. Priorizaram-se publicações dos últimos 15 anos, com ênfase em estudos de neuroimagem funcional, meta-análises e revisões sistemáticas quando disponíveis. Foram incluídos também trabalhos seminais anteriores que estabeleceram fundamentos teóricos relevantes para a discussão.

A análise integra evidências neurocientíficas sobre os substratos neurais da cognição social com dados emergentes sobre interações humano-IA, organizando-se em torno de três eixos temáticos: **1.** os desafios técnicos e conceituais para a modelagem computacional da inteligência social; **2.** os correlatos neurais diferenciais nas interações com agentes humanos versus não-humanos; e **3.** as implicações clínicas do uso de tecnologias de IA no contexto do manejo da solidão e do sofrimento psíquico. O presente artigo, de natureza ensaística, busca sintetizar criticamente esse corpo de evidências, propondo reflexões teóricas sobre as fronteiras e possibilidades da cognição social na era digital.

Resultados e Discussão

A cognição social como o principal desafio para a inteligência artificial
Tecnologias envolvendo a Inteligência Artificial (IA) tem avançado a passos largos na última década, no entanto, quando a interação natural e, principalmente de ordem social complexa é necessária, os algoritmos existentes demonstram de forma clara as suas imperfeições e limitações [2]. De maneira geral, as nuances da interação social e a subjetividade das experiências pessoais ainda são uma limitação para as ferramentas de IA [3]. Em instância principal, os desafios enfrentados no processo de naturalização da interação das máquinas inteligentes com os seres humanos se dão pela enorme complexidade da cognição social humana.

As perspectivas neurocientíficas sobre a interação entre os seres humanos e as máquinas inteligentes é uma área especialmente nova, em fase de construção e com muitos caminhos a serem descobertos. Desde o desenvolvimento do KISMET no Laboratório de Mídia do MIT, no final da

década de 1990, um dos primeiros robôs sociais, houve um progresso significativo na engenharia das máquinas inteligentes, sendo essas tecnologias atualmente capazes de interagir com humanos em um nível social relativamente complexo. Por exemplo, robôs que respondem e desencadeiam emoções humanas não apenas permitem uma colaboração mais próxima entre humanos e máquinas, mas também podem estimular usuários humanos a desenvolver laços sociais de longo prazo com esses agentes.

Ao mesmo tempo, as máquinas inteligentes têm falhado em atender às expectativas, principalmente no contexto da interação social com os humanos e esse tópico tem sido destacado como um dos dez grandes desafios que a área da robótica enfrenta atualmente [4]. Para facilitar o progresso nesse sentido, a rica literatura da neurociência cognitiva é essencial para oferecer insights sobre as nuances do comportamento social humano, não apenas em um nível superficial, mas também em relação aos mecanismos funcionais e biológicos subjacentes à interação social.

Embora haja alguma pesquisa em robótica sobre a cognição social, ela se concentra principalmente na interação humano-robô (por exemplo, como um robô interpreta as intenções de um humano) e não nas interações entre robôs [5]. É importante notar que o que parece cognição social não é necessariamente cognição social, visto que o conceito que rege a cognição social está atrelado à dinâmica social entre seres da mesma espécie [6].

A maioria das interações sociais exige que as pessoas percebam que estão na presença de outro ser humano para que elas ocorram de maneira efetiva. A presença humana instancia um contexto social, e uma interação social bem-sucedida requer um comportamento apropriado a esse contexto, onde as regras e normas sociais determinam qual seria o comportamento apropriado.

A presença real e implícita de um membro da mesma espécie (para humanos, estes são outros seres humanos) é o ambiente em que os cérebros da maioria dos animais, incluindo humanos, evoluíram. A presença física ou não de um membro da mesma espécie sinaliza oportunidades de cooperação ou competição, bem como ameaça e segurança em potencial, portanto, ser capaz de detectar a presença de outros é necessário para a sobrevivência e reprodução. Além disso, é um dos precursores críticos para a ativação da rede de mentalização, função cognitiva vital para o sucesso em contextos sociais complexos.

Embora esse campo seja marcado pela predominância de perguntas em relação às respostas já encontradas, o que é fato até aqui é que os avanços tecnológicos têm impulsionado o aprofundamento do conhecimento sobre os mecanismos envolvidos na interação social humana, mesmo que de forma indireta. Dessa maneira, aqui argumento que a análise das evidências provenientes de estudos focados em avaliar a relação entre a cognição social humana e a interação com as tecnologias inteligentes fornecerá insights para uma compreensão mais clara e profunda do lado humano da interação social.

A interação entre humanos e os tipos de IAs: as contribuições para a compreensão da hierarquia entre as funções cognitivas sociais

Está bem estabelecido na literatura atual que a cognição social se refere a um termo “guarda-chuva” que abrange uma variedade de funções cognitivas implicadas no processamento das informações sociais [6, 7, 8, 9]. A relação entre os componentes cognitivos do processamento social ainda gera debate frente a algumas evidências incipientes. Por exemplo, dados de neuroimagem demonstram atividade cerebral sobreposta ou distinta durante diferentes tarefas/processos: a ativação diferencial causada por duas tarefas sociais diferentes envolve redes neurológicas sobrepostas [10].

Além disso, há associações longitudinais nas diferenças individuais: a covariação dentro dos indivíduos ao longo do desenvolvimento tem sido usada para defender a cascata do desenvolvimento, onde a aquisição de uma habilidade social leva à aquisição de outra. Essa disposição hierárquica tem sido comprovada também no nível neural [10]. A rede de neurônios-espelho é atualmente considerada substrato no desenvolvimento dos domínios primários da cognição social, junto com as redes envolvidas na percepção social e afiliação [6, 7, 8, 9, 10, 11]. Já a rede neural da mentalização é considerada uma via associativa superior, diretamente relacionada com funções top-down e de natureza controlada [6].

O cruzamento entre os campos da neurociência e da robótica tem fornecido perspectivas empíricas valiosas para a elaboração de concepções sobre como se comporta a cognição social durante as interações entre humanos e robôs. Partindo da validação sobre a importância da corporificação das ferramentas de IA, os estudos neste campo têm fornecido dados sobre a relação intradomínio da cognição social [6, 9]. Por exemplo, a IA não corporificada (sem forma física; por exemplo, os algoritmos ChatGPT e

DeepSeek) não é capaz de acionar as vias neurais envolvidas na percepção de pistas sociais implícitas (por exemplo, o movimento corporal e as expressões faciais), ativando apenas parte dos processos cerebrais superiores envolvidos na capacidade de realizar inferências sobre outros estados mentais. Como consequência, as diferenças na ativação cerebral ao interagir com humanos e com IAs não corporificadas refletem uma ativação exclusiva das redes associativas superiores, não sendo ativadas as vias primárias cognitivas sociais, essenciais precursores de uma ativação ampla do cérebro social.

Por outro lado, o antropomorfismo de IAs corporificadas (por exemplo, robôs, avatares, andróides) é capaz de acionar as vias primárias (subconscientes) da cognição social porque robôs e avatares representam uma forma física que pode ser processada a partir da ativação dos neurocircuitos envolvidos na percepção do movimento biológico humano.

Essas considerações têm sido valiosas para a compreensão das relações intradomínio entre as funções que compõem a cognição social. Em suma, ainda é motivo de debate na literatura se a ativação das vias que sustentam as funções cognitivas sociais primárias (por exemplo, a percepção social) funcionaria como um precursor obrigatório da atividade de redes neurais superiores envolvidas na mentalização e no raciocínio social [8].

No entanto, embora já existam provas científicas robustas da disposição hierárquica intradomínio da cognição social [8, 10, 12], o que tem sido observado, incluindo a partir das análises envolvendo a interação entre os seres humanos e as diferentes formas de IA, é que a amplitude de ativação das redes neurais sociais e, conseqüentemente, o tipo de função envolvida no ato social irá variar de acordo com a natureza do estímulo social. A análise das diferenças que marcam a interação dos humanos com ferramentas de IA de acordo com o grau de corporificação já sugere que as funções cognitivas sociais podem ser recrutadas isoladamente e independentemente do grau afetivo envolvido no estímulo capturado. Esse insight pode ajudar a responder questionamentos levantados em estudos anteriores sobre o grau de dependência de participação das funções cognitivas sociais primárias para o processo de mentalização [8, 10, 12].

O processo imaginativo interpessoal: a marca do que nos torna humanos? Os seres humanos se distinguem dos não humanos porque possuem uma mente [13]. Ter uma mente implica numa experiência consciente de pensamento e emoção, que é privada e reconhecida apenas pela pessoa

que a vivência [14]. Portanto, os seres humanos nunca percebem diretamente outras mentes; eles as inferem. A inferência social é a base da habilidade que envolve a elaboração de atribuições acerca dos fatores causais envolvidos na expressão comportamental de terceiros e caracteriza a função mais complexa e elementar da cognição social: a mentalização [6, 7, 15].

Assim, reconhecer que outro ser da mesma espécie possui uma mente é um processo essencialmente imaginativo e essa é considerada uma habilidade exclusiva humana. A imaginação como um conceito abstrato separado de sua experiência fenomenológica, respectivamente, o "eu" e o "mim" [16], permite considerar que o *self* pode ser percebido como um alvo social e até mesmo desumanizado [17]. Dada essa flexibilidade especial da cognição social humana, é importante entender onde estão os limites do que as pessoas consideram humano. Controlar o componente imaginativo da cognição social revelaria os mecanismos envolvidos na presença real e implícita de agentes biológicos humanos, uma espécie de detector de humanidade. Isso levanta a questão: quais processos psicológicos e cerebrais são necessários e suficientes para que algo seja considerado humano?

A imagiologia cerebral pode lançar alguma luz sobre essa questão. O cérebro processa as formas físicas dos seres humanos usando o mapeamento espacial, função atrelada à atividade do lobo temporal ventral [18], o mesmo tipo de processamento que objetos não humanos desencadeiam, permitindo ao observador identificar o agente dissociado de si mesmo.

Numa instância mais complexa, o cérebro calcula uma inferência de estado mental, uma integração bayesiana de informações estatísticas, para inferir o conteúdo da mente de um agente social [18, 19, 20]. Dessa maneira, as mentes não podem ser percebidas diretamente, mas são apenas inferidas a partir de um conjunto complexo de sinais comportamentais (explícitos e implícitos), estando essa habilidade normalmente reservada aos seres humanos.

O compartilhamento de experiências se refere ao subdomínio da cognição social que permite ao indivíduo vivenciar a ressonância motora durante a observação das ações de terceiros e o contágio emocional, sendo esse último comumente referenciado como empatia afetiva [8 - 9]. Em nível neural, o compartilhamento de experiências depende da ativação da rede de neurônios-espelho [9].

A rede de neurônios-espelho compreende as regiões sensório-motoras, temporal e parietal [21]. De maneira geral, esse sistema permite que o sujeito seja capaz de interpretar sinais básicos relacionados com as ações e emoções de terceiros através da representação mental do que é percebido. Esse processo de representação mental é capaz de promover sensações físicas, viscerais e emocionais semelhantes ao que é percebido no outro, sendo esse fenômeno nomeado de compartilhamento de experiências motoras ou afetivas [6, 9]. Ao mesmo tempo, é válido destacar que a representação mental promovida pela ativação da rede de neurônios-espelho é considerada básica ou rudimentar e não se estende à capacidade de realizar inferências complexas.

Gazzola et al. [22] apresentaram dados que indicam que a ativação do circuito de neurônios-espelho nas regiões dos córtices parietal, pré-motor e temporal médio foi semelhante nos participantes enquanto observavam humanos agarrando e manipulando objetos ou um braço robótico industrial realizando essas mesmas ações. Essas descobertas foram replicadas no estudo de Oberman et al. [23], que usou registros eletroencefalográficos (EEG) nas regiões sensório-motoras para comparar os padrões de atividade elétrica neural dos participantes enquanto eles observavam ações expressas por agentes robóticos e humanos.

Esses achados sobre a ressonância motora durante a interação com robôs foram replicados e ampliados posteriormente [24, 25]. Dado o fato de que as ressonâncias motora e afetiva compartilham dos mesmos substratos neurais, após a publicação de evidências acerca da ocorrência de ressonância motora durante a interação com robôs, surgiram questionamentos sobre até que ponto observadores humanos também compartilham de emoções expressas mecanicamente.

Com esse propósito, estudos como os de Hortensius et al. [26] e Wang & Quadflieg [27] apresentaram registros de imagens funcionais cerebrais dos participantes durante a interação com robôs que manifestavam sinais afetivos e durante a observação de interações afetivas entre robôs e outros seres humanos, respectivamente. Em ambos os estudos, foram apresentados dados empíricos que sustentam a hipótese de que ocorre também a ressonância afetiva em contextos de expressão simulada de sinais de dor.

Numa perspectiva complementar, o estudo conduzido por Özdem et al. [28] registrou os sinais de ativação neural durante a realização de paradigmas que elicitaram habilidades de mentalização dos participantes.

Nesse estudo, a resposta cerebral (por exemplo, a ativação da junção temporoparietal anterior bilateral) ocorreu apenas quando as pessoas acreditavam que outra pessoa controlava o robô.

A análise desses achados abre debate para uma outra consideração: o quanto o estímulo social é passível de moldagem e flexibilização durante a etapa processual que envolve a representação mental da situação social. O debate sobre as diferenças percebidas nas interações com outros agentes humanos e agentes não humanos tem iluminado a importância do contexto nas relações sociais humanas. Existe uma diferença entre "presença social" e "presença humana", uma vez que a primeira inclui o antropomorfismo (presença corporificada) enquanto a presença humana remete a um fenômeno abstrato e imaginativo sobre outro agente da mesma espécie.

Em parte, por essa variabilidade complexa, a replicação de contextos sociais ainda é um desafio nos experimentos científicos e em contextos controlados de avaliação clínica e em pesquisas. Há algum tempo, as evidências indicam haver diferenças notórias na maneira como o indivíduo pode interpretar o contexto social, a depender da cultura e até mesmo do clima do dia [29, 30].

O papel do contexto social na dinâmica da cognição social é, portanto, difícil de compreender completamente, e a neurociência tem começado a desprender esforços para a maior compreensão sobre como diferentes contextos sociais impactam nos tipos específicos de comportamentos sociais das pessoas. Neurocientistas têm utilizado ferramentas de IA como linhas de base para comparar respostas ao comportamento humano e às atividades cerebrais em estudos de percepção social e tomada de decisão. IAs não impedem o processamento social, uma vez que podem ser consideradas agentes com intenções e outras capacidades mentais. No entanto, ser considerado humano envolve o uso de um conceito abstrato ou estereótipo do que significa ser humano, que tem sido considerado para dissociar agentes humanos de não humanos.

Estudos focados em detectar como o cérebro diferencia agentes biológicos humanos de agentes não humanos pode ser útil para comprovar algumas premissas teóricas atuais. Por exemplo, as regras morais e o regimento legal são considerados relevantes para regular o comportamento social das pessoas, visto que esse fator leva à consideração das pessoas como perpetradores e vítimas. Dado o fato de que o raciocínio moral é reservado às pessoas e geralmente não é considerado durante as trocas com agentes

não humanos, esse seria um dos pontos a se considerar nessa diferenciação.

Um outro ponto é que as pessoas tendem a se preocupar com a própria reputação ao interagir com outras, tornando a gestão de impressões relevante para a autorregulação social [31]. Ainda, a ressonância biológica e comportamental entre agentes humanos leva a uma sincronia comunicativa e facilita a interação social [32].

Vencidos alguns desafios na replicação contextual em experimentos científicos, dados empíricos futuros poderão fornecer insights mais aprofundados sobre essas influências. No entanto, o antropomorfismo sugere que as inferências de estado mental ocorrem de maneira hierárquica, sendo os componentes dessa hierarquia variáveis, a depender da natureza e saliência do estímulo capturado.

O uso dos agentes não humanos no manejo da solidão

Enfrentar a solidão representa um dos maiores desafios da atualidade. A solidão tem cada vez mais se firmado como um marcador de bem-estar e saúde, dado os riscos comparáveis aos do tabagismo, e é cada vez mais reconhecida como uma prioridade para as políticas de saúde pública [33, 34, 35]. A prevalência da solidão aumentou após a pandemia pelo COVID-19, impulsionada pelas medidas de distanciamento social [36]. Estudos conduzidos com a população americana revelam que cerca de 17% (o que corresponde a 44 milhões de pessoas) experimentam solidão significativa diariamente [37].

Como um potente estressor psicossocial, a vivência da solidão ativa o eixo hipotálamo-hipófise-adrenal e, conseqüentemente, está associada a uma ampla gama de desfechos adversos à saúde, incluindo a depressão, comportamento suicida, má qualidade do sono e doenças cardiovasculares [38, 39, 40, 41, 42, 43]. Conceituada em psiquiatria como um fator transdiagnóstico, a solidão permeia diversas categorias diagnósticas e contribui para o surgimento e manutenção de múltiplas condições psiquiátricas e físicas [44]. Ao contrário do isolamento social objetivo, a solidão é uma experiência subjetiva caracterizada por uma discrepância entre relacionamentos sociais significativos desejados e reais, levando a um distúrbio no senso da conexão e pertencimento social [16].

Partindo da premissa de que a conexão social é percebida como um continuum, a solidão pode ser classificada espectralmente entre leve e severa [45-46]. Apesar disso complexidade, a maioria dos ensaios clínicos

dicotomiza os indivíduos como solitários ou não, frequentemente negligenciando os graus de severidade da solidão [47, 48, 49, 50, 51, 52, 53, 54, 55]. A eficácia da intervenção é provavelmente influenciada por uma infinidade de fatores. Isso inclui o foco da intervenção, seja visando capacidades individuais (por exemplo, melhorando o enfrentamento ou a regulação emocional) ou o contexto mais amplo (por exemplo, abordando condições ambientais ou estruturais) [56, 57], bem como o modo de entrega (por exemplo, *on-line versus* presencial), que pode impactar a acessibilidade, o engajamento e a qualidade da interação social [58, 59].

Nesse cenário, as tecnologias de comunicação (por exemplo, o uso de *chatbots*) têm sido consideradas ferramentas potencialmente úteis para remediar experiências de solidão e sintomas mentais relacionados. Considerados os possíveis vieses de alguns estudos e as limitações impostas pela cognição social humana nos recursos oferecidos por ferramentas de IA, seriam as máquinas inteligentes úteis para se combater os sintomas psiquiátricos relacionados ao estado de solidão?

Nos últimos anos, os avanços tecnológicos tornaram possível à IA interagir com os humanos através de trocas comunicativas sociais, fornecendo companheirismo e orientação em momentos difíceis [60, 61]. Essas tecnologias são, inclusive, frequentemente chamadas de IA social ou robôs sociais. Ao contrário dos sistemas tradicionais de IA que se concentram em tarefas específicas ou na resolução de problemas, a IA social pode se envolver em interações naturais e significativas com as pessoas, muitas vezes imitando comportamentos sociais e comunicação humanos.

Pesquisadores já reconheceram o papel positivo que a IA social pode desempenhar no combate à solidão e em sintomas como a ansiedade, visto que pode fornecer mensagens de apoio emocional aos humanos e contribuir para um senso de conexão e pertencimento. Uma meta-análise constatou que mensagens com maior foco na pessoa foram percebidas como mais positivas do que mensagens genéricas ou gerais [62, 63]. Ou seja, mensagens personalizadas focadas nas necessidades, perspectivas e preferências individuais são percebidas como mais favoráveis para a percepção de suporte social.

Embora pesquisas anteriores tenham examinado extensivamente o papel das mensagens de apoio centradas na pessoa em um contexto de comunicação entre humanos, não está claro se essas descobertas se replicam em um contexto de comunicação interpessoal. Apesar disso, esse já pode ser considerado um insight valioso para o manejo clínico.

Outro ponto destacado em estudos é o fato de que as ferramentas de IA oferecem suporte social e redução dos níveis de solidão ao proporcionar um "espaço seguro", no qual os usuários podem discutir qualquer tópico sem medo de julgamentos morais ou retaliação. Os usuários informam haver redução nos níveis de ansiedade com o recebimento de afeto positivo por meio de mensagens inspiradoras e estimulantes fornecidas por *chatbots* e ao receber conselhos úteis ao compartilharem problemas que não se sentem seguros em compartilhar com outras pessoas [64].

A inteligência artificial como antítese da cognição social

Assim como o mundo digital pode colaborar para reduzir distâncias e minimizar a solidão, pode também operar em modo inverso. Nas redes, interagimos por meio de um filtro, o que constantemente causa a confusão entre interação e socialização; parece estarmos nos socializando, mas a capacidade de resposta e de construção de vínculos cai muito. Redes sociais estimulam o comportamento individualista e relações superficiais, tornando-nos preguiçosos para os relacionamentos [64].

A vida digital, bem como a dependência e o uso abusivo de I.A. podem amplificar o sentimento de solidão através de diversos mecanismos: **a)** substituição de interações presenciais, **b)** isolamento passivo (uso passivo das redes sociais, apenas como observador); **c)** criação de uma versão idealizada da vida (e conseqüente distanciamento e insuflação de vivências negativas como menos-valia e inadequação); **d)** vazio emocional, gerando procura por validação e gratificação instantâneos que nunca satisfazem, dada sua efemeridade. Por fim, as gerações que nasceram já imersas no mundo digital podem exibir dificuldades no desenvolvimento de habilidades sociais básicas que irão repercutir em índices mais elevados de solidão e egocentrismo [64].

Conclusões e perspectivas futuras

A análise crítica da relação entre cognição social e inteligência artificial revela um cenário de complexidade paradoxal: enquanto os avanços tecnológicos impulsionam o desenvolvimento de sistemas cada vez mais sofisticados, a natureza intrinsecamente humana da cognição social permanece como barreira fundamental à naturalização completa das interações humano-máquina. Este artigo demonstrou que a compreensão dos mecanismos neurais subjacentes à cognição social não apenas ilumina os desafios técnicos enfrentados pela IA, mas também oferece insights valiosos sobre o que nos define como humanos.

A primeira contribuição central deste trabalho reside na identificação da hierarquia funcional da cognição social como obstáculo primário para o desenvolvimento de IAs verdadeiramente sociais. As evidências neurocientíficas aqui revisadas demonstram que a ativação diferencial de redes neurais primárias (como os neurônios-espelho) versus redes associativas superiores (envolvidas na mentalização) depende criticamente da corporificação dos agentes sociais [5, 6, 22, 23]. Esta distinção sugere que o antropomorfismo, longe de ser mero artifício estético, constitui pré-requisito funcional para o engajamento pleno do aparato neurocognitivo social humano. Sistemas de IA não-corporificados, embora capazes de processar linguagem natural e simular conversação, falham em ativar os substratos neurais fundamentais do processamento social, limitando assim a profundidade e a naturalidade da interação.

A segunda contribuição refere-se ao processo imaginativo interpessoal como marca distintiva da cognição social humana. A capacidade de inferir estados mentais, de reconhecer que outro ser possui uma mente com experiências subjetivas próprias, transcende o mero processamento de sinais comportamentais e envolve um ato fundamentalmente criativo de atribuição de humanidade [13, 14, 15, 16]. Esta dimensão imaginativa, ancorada em estruturas evolutivas que privilegiam a detecção de coespecíficos, resiste à simulação algorítmica precisamente porque opera em níveis que vão além da computação de probabilidades bayesianas. O processamento social humano não se reduz a inferências estatísticas; ele incorpora dimensões afetivas, morais e contextuais que emergem da experiência vivida de ter uma mente e reconhecer mentes em outros.

No contexto clínico, as implicações são igualmente complexas. Por um lado, tecnologias de IA demonstram potencial promissor no manejo da solidão, oferecendo "espaços seguros" para expressão emocional sem julgamento e mensagens de suporte personalizadas [60, 61, 62, 63]. A eficácia dessas intervenções sugere que certos aspectos do suporte social podem ser parcialmente replicados por sistemas artificiais, particularmente aqueles relacionados à validação emocional e ao fornecimento de informações. Por outro lado, a dependência excessiva de interações mediadas por IA apresenta riscos substanciais de amplificação da solidão através da substituição de conexões humanas genuínas por simulacros que, embora confortáveis, carecem da reciprocidade e da imprevisibilidade que caracterizam relacionamentos autênticos [63, 64].

REFERÊNCIAS

1. Broadbent E. Interactions with robots: the truths we reveal about ourselves. *Annu Rev Psychol.* 2017;68:627-52.
<https://doi.org/10.1146/annurev-psych-010416-043958>
PMid:27648986
2. Schrittwieser J, Hubert T, Mandhane A, Barekatin M, Antonoglou, I, Silver D. Online and offline reinforcement learning by planning with a learned model. In: 35th Conference on Neural Information Processing [Internet]; 2021. p. 1-12. Disponível em:
https://proceedings.neurips.cc/paper_files/paper/2021/file/e8258e5140317ff36c7f8225a3bf9590-Paper.pdf
3. Bolotta S, Dumas G. Social neuro AI: social interaction as the "dark matter" of AI. *Front Comput Sci.* 2022;4:846440.
<https://doi.org/10.3389/fcomp.2022.846440>
4. Yang G-Z, Bellingham J, Dupont PE, Fischer P, Floridi L, Full R, Jacobstein N, Kumar V, McNutt M, Merrifield R, Nelson BJ, Scassellati B, Taddeo M, Taylor R, Veloso M, Wang ZL, Wood R. The grand challenges of science robotics. *Sci Robot.* 2018;3(14):eaar7650. <https://doi.org/10.1126/scirobotics.aar7650>
5. Henschel A, Hortensius R, Cross ES. Social cognition in the age of human-robot interaction. *Trends Neurosci.* 2020;43(6):373-84.
<https://doi.org/10.1016/j.tins.2020.03.013> PMid:32362399
6. Figueiredo T, Mecca TP. Um guia para a compreensão das bases do comportamento social. Belo Horizonte: Ampla; 2024.
7. Figueiredo, T. Tratado de cognição social: uma abordagem multidimensional. Belo Horizonte: Ampla; 2024
8. Happé F, Cook JL, Bird G. The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annu Rev Psychol.* 2017;68:243-67. <https://doi.org/10.1146/annurev-psych-010416-044046> PMid:27687121
9. Green MF, Horan WP, Lee J. Social cognition in schizophrenia. *Nat Rev Neurosci.* 2015;16(10):620-31.
<https://doi.org/10.1038/nrn4005> PMid:26373471

10. Schurz M, Radua J, Tholen MG, Maliske L, Margulies DS, Mars RB, Sallet J, Kanske P. Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychol Bull.* 2021 Mar;147(3):293-327. <http://doi.org/10.1037/bul0000303>. PMID: 33151703
11. Bickart KC, Dickerson BC, Barrett LF. The amygdala as a hub in brain networks that support social life. *Neuropsychologia.* 2014;63:235-48. <https://doi.org/10.1016/j.neuropsychologia.2014.08.013> PMID:25152530 PMCID:PMC4981504
12. Happé F, Frith U. Annual research review: towards a developmental neuroscience of atypical social cognition. *J Child Psychol Psychiatry.* 2014;55(6):553-77. <https://doi.org/10.1111/jcpp.12162> PMID:24963529
13. Fiske ST, Taylor SE. *Social cognition: from brains to culture.* Sage Publications; 2013. <https://doi.org/10.4135/9781529681451>
14. Wittgenstein L. *Philosophical investigations.* 3rd ed. New York: Macmillan; 1993.
15. Quesque F, Apperly I, Baillargeon R, Baron-Cohen S, Becchio C, Bekkering H, Bernstein D, Bertoux M, Bird G, Bukowski H, Burgmer P, Carruthers P, Catmur C, Dziobek I, Epley N, Erle TM, Frith C, Frith U, Galang CM, Gallese V, Grynberg D, Happé F, Hirai M, Hodges SD, Kanske P, Kret M, Lamm C, Nandrino JL, Obhi S, Olderbak S, Perner J, Rossetti Y, Schneider D, Schurz M, Schuwerk T, Sebanz N, Shamay-Tsoory S, Silani G, Spaulding S, Todd AR, Westra E, Zahavi D, Brass M. Defining key concepts for mental state attribution. *Commun Psychol,* 2024;2(29). <https://doi.org/10.1038/s44271-024-00077-6>
16. James W. *The principles of psychology.* Vol. 1. New York: Henry Holt and Company; 1890. <https://doi.org/10.1037/10538-000>
17. Tang N, Harris LT. Self-directed dehumanization: when people view themselves as less than human. *Self Identity.* 2015;14(6):654-74. <https://doi.org/10.1080/15298868.2015.1053518>
18. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci.* 2001;4(6):223-

33. [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0)
PMid:10827445

- ↑ 19. Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci*. 2006;7(4):268-77. <https://doi.org/10.1038/nrn1884> PMid:16552413
- ↑ 20. Frith CD, Frith U. The biological basis of social interaction. *Curr Dir Psychol Sci*. 2001;10(5):151-55. <https://doi.org/10.1111/1467-8721.00137>
- ↑ 21. Gallagher HL, Frith CD. Functional imaging of theory of mind. *Trends Cogn Sci*. 2003;7(2):77-83. [https://doi.org/10.1016/S1364-6613\(02\)00025-6](https://doi.org/10.1016/S1364-6613(02)00025-6) PMid:12584026
- ↑ 22. Gazzola V, Rizzolatti G, Wicker B, Keysers C. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*. 2007;35(4):1674-84. <https://doi.org/10.1016/j.neuroimage.2007.02.003> PMid:17395490
- ↑ 23. Oberman LM, McCleery JP, Ramachandran VS, Pineda JA. EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. *Neurocomputing*. 2007;70(13-15):2194-203. <https://doi.org/10.1016/j.neucom.2006.02.024>
- ↑ 24. Cross ES, Ramsey R, Liepelt R, Prinz W, Hamilton AFDC. The shaping of social perception by stimulus and knowledge cues to human animacy. *Philos Trans R Soc Lond B Biol Sci*. 2016;371(1686): 20150075. <https://doi.org/10.1098/rstb.2015.0075> PMID: 26644594 PMCID: PMC4685521
- ↑ 25. Press C. Action observation and robotic agents: learning and anthropomorphism. *Neurosci Biobehav Rev*. 2011;35(6):1410-18. <https://doi.org/10.1016/j.neubiorev.2011.03.004> PMid:21396398
- ↑ 26. Hortensius R, Hekele F, Cross ES. The perception of emotion in artificial agents. *IEEE Trans Cogn Dev Syst*. 2018;10(4):852-64. <https://doi.org/10.1109/TCDS.2018.2826921>
- ↑ 27. Wang Y, Quadflieg S. In our own image? Emotional and neural processing differences when observing human-human vs human-

robot interactions. *Soc Cogn Affect Neurosci*. 2015;10(11):1515-24.
<https://doi.org/10.1093/scan/nsv043> PMID:25911418
PMCID:PMC4631149

- 28. Özdem C, Wiese E, Wykowska A, Müller HJ, Brass M, Van Overwalle F. Believing androids - fMRI activation in the right temporo-parietal junction is modulated by ascribing intentions to non-human agents. *Soc Neurosci*. 2017;12(5):582-93.
<https://doi.org/10.1080/17470919.2016.1207702> PMID:27391213
- 29. Markus HR, Kitayama S. Culture and the self: implications for cognition, emotion, and motivation. *Psychol Rev*. 1991;98(2):224-53. <https://doi.org/10.1037/0033-295X.98.2.224>
- 30. Schwarz N, Clore GL. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J Pers Soc Psychol*. 1983;45(3):513-23.
<https://doi.org/10.1037/0022-3514.45.3.513>
- 31. Goffman E. *The presentation of self in everyday life*. Garden City: Doubleday; 1959.
- 32. McNeill WH. *Keeping together in time: dance and drill in human history*. Cambridge: Harvard University Press; 1997.
<https://doi.org/10.4159/9780674040878>
- 33. Arsenijevic J, Groot W. Does household help prevent loneliness among the elderly? An evaluation of a policy reform in the Netherlands. *BMC Public Health*. 2018;18(1):1-12.
<https://doi.org/10.1186/s12889-018-6004-6> PMID:30200920
PMCID:PMC6131846
- 34. Holt-Lunstad J. The potential public health relevance of social isolation and loneliness: prevalence, epidemiology, and risk factors. *Public Policy Aging Rep*. 2017;27(4):127-30.
<https://doi.org/10.1093/ppar/prx030>
- 35. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A*. 2013;110(15):5797-801.
<https://doi.org/10.1073/pnas.1219686110> PMID:23530191
PMCID:PMC3625264



- 36. Organização Mundial da Saúde. Mental health and COVID-19. Geneva: WHO; 2020.
https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief-Mental_health-2022.1
- 37. Witters D. Loneliness in the U.S. at a new high. Gallup Wellbeing; 2023.
- 38. Valtorta NK, Kanaan M, Gilbody S, Hanratty B. Loneliness, social isolation and risk of cardiovascular disease in the English Longitudinal Study of Ageing. *Eur J Prev Cardiol.* 2018;25(13):1387-96. <https://doi.org/10.1177/2047487318792696> PMID:30068233
- 39. Cacioppo JT, Hawkley LC. Perceived social isolation and cognition. *Trends Cogn Sci.* 2009;13(10):447-54.
<https://doi.org/10.1016/j.tics.2009.06.005> PMID:19726219
PMCID:PMC2752489
- 40. Cacioppo S, Capitanio JP, Cacioppo JT. Toward a neurology of loneliness. *Psychol Bull.* 2014;140(6):1464-1504.
<https://doi.org/10.1037/a0037618> PMID:25222636
PMCID:PMC5130107
- 41. Vanhalst J, Klimstra TA, Luyckx K, Scholte RH, Engels RC, Goossens L. The interplay of loneliness and depressive symptoms across adolescence: exploring the role of personality traits. *J Youth Adolesc.* 2012;41(6):776-87. <https://doi.org/10.1007/s10964-011-9726-7> PMID:22045508
- 42. Mezuk B, Rock A, Lohman MC, Choi M. Suicide risk in long-term care facilities: a systematic review. *Int J Geriatr Psychiatry.* 2014;29(12):1198-1211. <https://doi.org/10.1002/gps.4142> PMID:24854089 PMCID:PMC4232590
- 43. Hom MA, Stanley IH, Podlogar MC, Joiner TE Jr. "Are you having thoughts of suicide?" Examining experiences with disclosing and denying suicidal ideation. *J Clin Psychol.* 2017;73(10):1382-92.
<https://doi.org/10.1002/jclp.22440> PMID: 28085200
- 44. Nowland R, Necka EA, Cacioppo JT. Loneliness and social internet use: pathways to reconnection in a digital world? *Perspect Psychol Sci.* 2018;13(1):70-87.
<https://doi.org/10.1177/1745691617713052> PMID:28937910

45. Cacioppo S, Cacioppo JT. Decoding the invisible forces of social connections. *Front Integr Neurosci*. 2012;6:51. <https://doi.org/10.3389/fnint.2012.00051> PMID:22848192
PMCID:PMC3404426
46. Tomova L, Wang KL, Thompson T, Matthews GA, Takahashi A, Tye KM, Saxe R. Acute social isolation evokes midbrain craving responses similar to hunger. *Nat Neurosci*. 2020;23(12):1597-1605. <https://doi.org/10.1038/s41593-020-00742-z> PMID:33230328
PMCID:PMC8580014
47. Astell-Burt T, Feng X, Kolt GS. Does access to neighbourhood green space promote a healthy duration of sleep? Novel findings from a cross-sectional study of 259 319 Australians. *BMJ open*. 2013;3(8):e003094. <https://doi.org/10.1136/bmjopen-2013-003094>
48. Barreto M, Victor C, Hammond C, Eccles A, Richins MT, Qualter P. Loneliness around the world: age, gender, and cultural differences in loneliness. *Pers Individ Dif*. 2021;169:110066. <https://doi.org/10.1016/j.paid.2020.110066> PMID:33536694
PMCID:PMC7768187
49. Cohen-Mansfield J, Perach R. Interventions for alleviating loneliness among older persons: a critical review. *Am J Health Promot*. 2015;29(3):e109-e125. <https://doi.org/10.4278/ajhp.130418-LIT-182> PMID:24575725
50. Eccles AM, Qualter P. Review: Alleviating loneliness in young people-a meta-analysis of interventions. *Child Adolesc Ment Health*. 2021;26(1):17-33. <https://doi.org/10.1111/camh.12389> PMID:32406165
51. Grillich L, Titscher V, Klingenstein P, Kostial E, Emprechtinger R, Klerings I, Sommer I, Nikitin J, Laireiter AR. The effectiveness of interventions to prevent loneliness and social isolation in the community-dwelling and old population: an overview of systematic reviews and meta-analysis. *Eur J Public Health*. 2023 Apr 1;33(2):235-241. <https://doi.org/10.1093/eurpub/ckad006> PMID: 36893335; PMCID: PMC10263264.
52. Gardiner C, Geldenhuys G, Gott M. Interventions to reduce social isolation and loneliness among older people: an integrative

review. Health Soc Care Community. 2018;26(2):147-57.
<https://doi.org/10.1111/hsc.12367> PMID:27413007

- 53. Jarvis MA, Padmanabhanunni A, Chipps J, Marimuthu M. An evaluation of a low-intensity cognitive behavioral therapy intervention for loneliness among older adults. Aging Ment Health. 2020;24(8):1262-69.
- 54. Maes M, Qualter P, Vanhalst J, Van den Noortgate W, Goossens L. Gender differences in loneliness across the lifespan: A meta-analysis. Eur J Pers. 2019;33(6):642-54.
<https://doi.org/10.1002/per.2220>
- 55. Masi CM, Chen HY, Hawkley LC, Cacioppo JT. A meta-analysis of interventions to reduce loneliness. Pers Soc Psychol Rev. 2011;15(3):219-66. <https://doi.org/10.1177/1088868310377394>
PMid:20716644 PMCID:PMC3865701
- 56. Cacioppo JT, Patrick W. Loneliness: human nature and the need for social connection. New York: WW Norton & Company; 2008.
- 57. Cacioppo JT, Hawkley LC, Crawford LE, Ernst JM, Burleson MH, Kowalewski RB, Malarkey WB, Van Cauter E, Berntson GG. Loneliness and health: potential mechanisms. Psychosom Med. 2002;64(3):407-17. <https://doi.org/10.1097/00006842-200205000-00005> PMID:12021415
- 58. Hickin N, Käll A, Shafran R, Sutcliffe S, Manzotti G, Langan D. The effectiveness of psychological interventions for loneliness: a systematic review and meta-analysis. Clin Psychol Rev. 2021 Aug;88:102066. <https://doi.org/10.1016/j.cpr.2021.102066> PMID: 34339939.
- 59. Kim J, Merrill K Jr, Collins C. AI as a friend or assistant: the mediating role of perceived usefulness in social AI vs. functional AI. Telemat Inform. 2021;64:101694.
<https://doi.org/10.1016/j.tele.2021.101694>
- 60. Meng J, Dai Y. Emotional support from AI chatbots: should a supportive partner self-disclose or not? J Comput Mediat Commun. 2021;26(4):207-22. <https://doi.org/10.1093/jcmc/zmab005>

61. High AC, Dillard JP. A review and meta-analysis of person-centered messages and social support outcomes. *Commun Stud.* 2012;63(1):99-118.
<https://doi.org/10.1080/10510974.2011.598208>
62. Ta V, Griffith C, Boatfield C, Wang X, Civitello M, Bader H, DeCero E, Loggarakis A. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *J Med Internet Res.* 2020;22(3):e16235. <https://doi.org/10.2196/16235> PMID:32141837 PMCID:PMC7084290
63. Abdel Kader Mohamed SA, Shaban M. Digital dependence in aging: nomophobia as the new mental health threat for older adults. *J Emerg Nurs.* 2025;51(3):371-78.
<https://doi.org/10.1016/j.jen.2025.02.001> PMID:40340052
64. Notara V, Vagka E, Gnardellis C, Lagiou A. The emerging phenomenon of nomophobia in young adults: a systematic review study. *Addict Health.* 2021;13(2):120-36.
<https://doi.org/10.22122/ahj.v13i2.309> PMID: 34703533 PMCID: PMC8519611