
Viewpoint

Perspectives on How Sociology Can Advance Theorizing About Human-Chatbot Interaction and Developing Chatbots for Social Good

Celeste Campos-Castillo¹, PhD; Xuan Kang¹, BA; Linnea I Laestadius², MPP, PhD

¹Department of Media and Information, College of Communication Arts and Sciences, Michigan State University, East Lansing, MI, United States

²Zilber College of Public Health, University of Wisconsin–Milwaukee, Milwaukee, WI, United States

Corresponding Author:

Celeste Campos-Castillo, PhD
Department of Media and Information
College of Communication Arts and Sciences, Michigan State University
404 Wilson Rd
East Lansing, MI 48824
United States
Phone: 1 517-432-5192
Email: camposca@msu.edu

Abstract

Recently, research into chatbots (also known as conversational agents, artificial intelligence agents, or voice assistants), which are computer applications using artificial intelligence to mimic human-like conversation, has grown sharply. Despite this growth, sociology lags behind other disciplines (including computer science, medicine, psychology, and communication) in publishing about chatbots. We suggest sociology can advance the understanding of human-chatbot interaction and offer 4 sociological theories to enhance extant work in this field. The first 2 theories (resource substitution theory and power-dependence theory) add new insights to existing models of the drivers of chatbot use, which overlook sociological concerns about how social structure (eg, systemic discrimination and the uneven distribution of resources within networks) inclines individuals to use chatbots, including problematic levels of emotional dependency on chatbots. The second 2 theories (affect control theory and fundamental cause of disease theory) help inform the development of chatbot-driven interventions that minimize safety risks by integrating a sociologically informed normative framework (eg, affective norms) into chatbot alignment and enhance equity by enhancing access to community resources (eg, opportunities for civic participation). We discuss how the theories advance theorizing about human-chatbot interaction and developing chatbots for social good, which are chatbots that provide scalable solutions to social and environmental challenges facing humanity while supporting human agency.

JMIR AI 2026;5:e80250; doi: [10.2196/80250](https://doi.org/10.2196/80250)

Keywords: human-computer interaction; chatbots; artificial intelligence; AI companions; public health; design; agency; self-determination

Introduction

Scholarly interest in chatbots, which are computer programs that simulate human conversation using artificial intelligence (AI), has grown sharply. Toward the end of 2024, Web of Science showed over 5000 articles and conference proceedings with the word “chatbot” appearing anywhere in the text. A similar search with additional terms (“conversational agent,” “voice assistant,” and “AI agent”) yielded comparable

patterns with respect to publication years and disciplines. [Figure 1](#) shows that about half of these were published in 2023 and 2024. [Figure 2](#) shows that most appear within computer science, followed by medicine, while sociology lags behind other social sciences. We seek to spur greater engagement with sociology to study human-chatbot interaction and develop chatbots. Accordingly, our aim with the current paper is to provide perspectives on how specific sociological theories could advance these areas.

Figure 1. Publications with “chatbot” appearing anywhere in text by publication year.

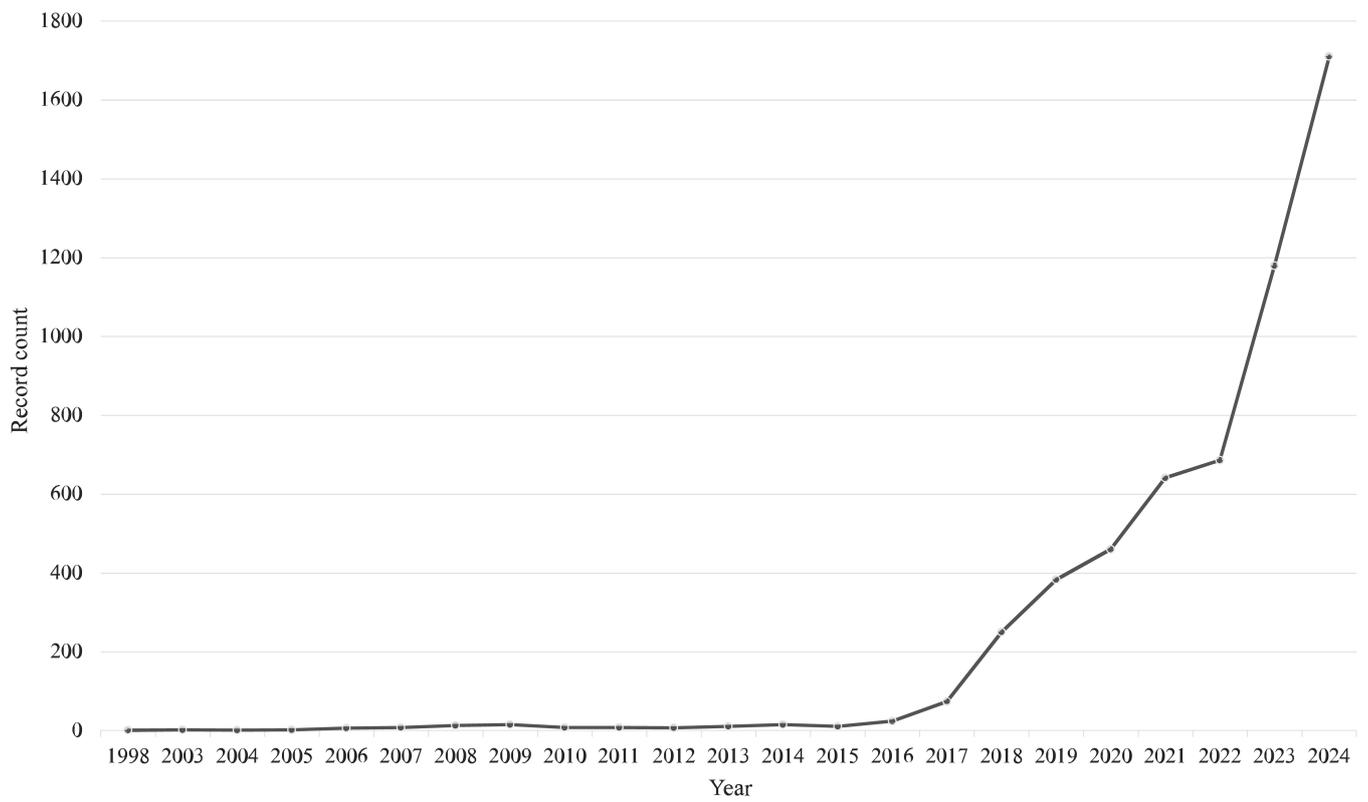
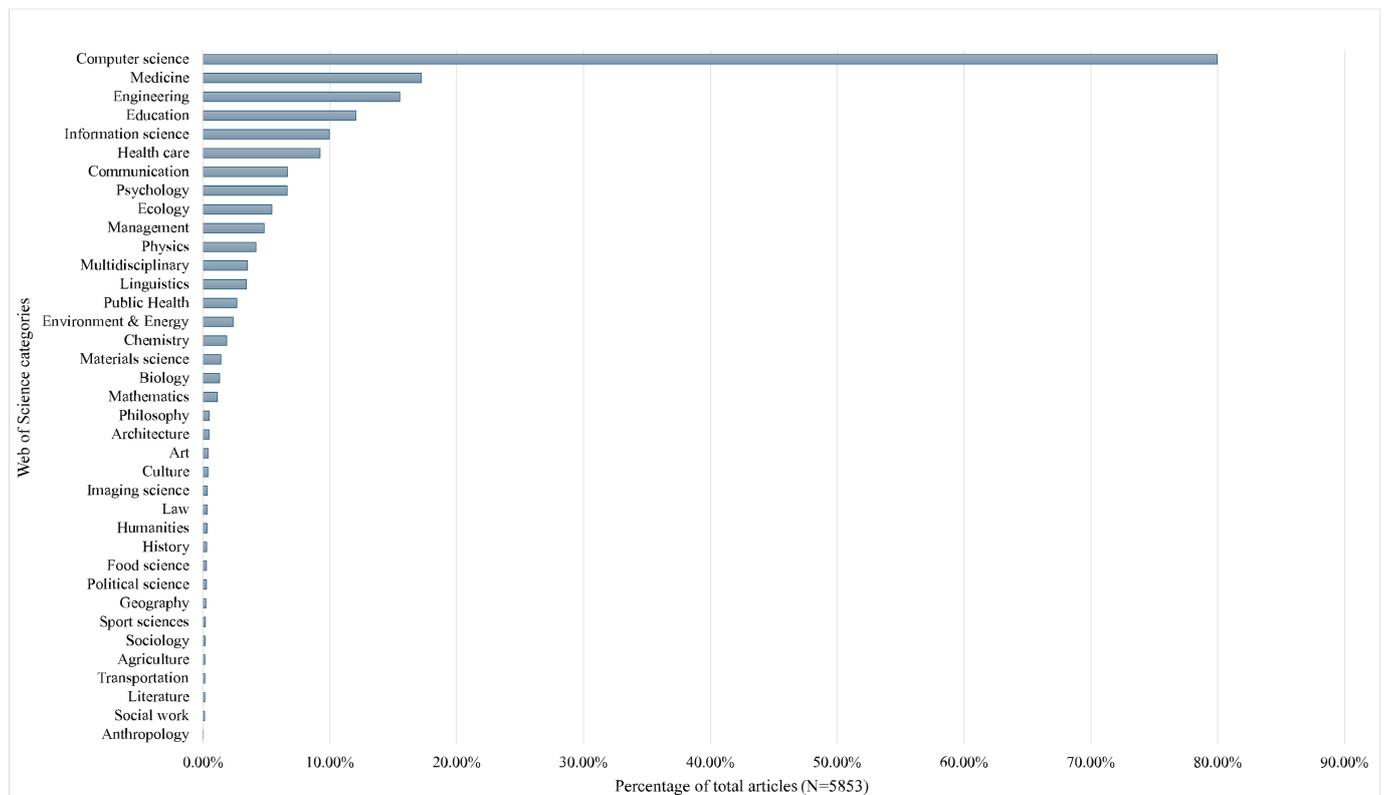


Figure 2. Disciplinary sources of the publications with “chatbot” appearing anywhere in the text.



We focus on direct communication between humans and chatbots, which is the latest iteration of a classic query in human-computer interaction research [1]. While this work concentrates on psychological drivers and impacts of chatbot use, we introduce 4 sociological theories that widen the scope of research to include the interplay between broader, social forces and human-chatbot communication. Specifically, the following theories address current gaps in the understanding of: (1) the structural drivers of chatbot use that underlie demographic patterns in who uses chatbots; (2) the drivers of overreliance on chatbots, including emotional dependence, and how to disrupt them; (3) how to proactively moderate chatbot outputs and enhance safety by incorporating a normative framework into chatbot alignment; and (4) how to design chatbot-driven interventions that target distal, or upstream, causes. We then present a hypothetical example illustrating how the theories could enhance multiple steps of chatbot development. By identifying ways to use chatbots to scale solutions to social and environmental problems facing humanity, while enabling human agency, these sociological theories can contribute toward developing chatbots for social good [2].

Selection of Theories

We purposely selected 4 sociological theories, previewed in Table 1, that vary across 3 characteristics: the phenomena they can explain, level of analysis, and governance domain. We chose 2 phenomena—drivers of chatbot use and developing chatbot-driven interventions—because they connect to the communication (former phenomenon) and public health (latter phenomenon) disciplines, which have been more active in studying chatbots (Figure 2). Moreover, while they share epistemological roots with sociology, these ultimately splintered. With respect to communication, despite sociology playing a significant role in its founding, the 2 disciplines now rarely intermingle [3]. Similarly, participatory methods have roots in sociology, yet most applications now occur within public health [4]. Our selection of the phenomena may thus reignite collaboration at the intersections of these disciplines.

Table 1. Characteristics of 4 selected sociological theories for studying human-chatbot interaction.

Characteristic	Resource substitution theory	Power-dependence theory	Affect control theory	Fundamental cause of disease theory
Phenomenon theory is used to explain	Drivers of chatbot use	Drivers of chatbot use	Chatbot-driven interventions	Chatbot-driven interventions
Unit of analysis	Macro	Micro	Micro	Macro
Governance domain	Equity	Risk	Risk	Equity

We selected theories spanning micro- and macro-levels of analysis. Although human-chatbot interaction is seemingly a micro-level phenomenon beyond the scope of sociology, sociology can reveal the interplay between social forces and these phenomena [3,5,6]. The 2 micro-level theories identify forces shaping human-chatbot interaction directly, while the 2 macro-level theories consider how initiating and outcomes from the interaction are embedded within broader forces. By suggesting ways to apply sociology across levels of analysis, we complement others [7,8] highlighting the macro-level implications of human-chatbot interaction.

Finally, our theory selection reflects 2 dominant emphases within AI policymaking [9]: safety and equity. Two theories elucidate safety risks from using chatbots, including misalignment, and offer mitigation strategies. The other 2 facilitate leveraging chatbots to achieve equity by addressing uneven access to resources. Altogether, our perspective suggests ways sociology can contribute to the study and development of chatbots for social good. We define chatbots for social good as those that enable scalable solutions to the social and environmental challenges facing humanity, while supporting human agency by mitigating risks like misalignment. The social and environmental challenges we reference include those defined by the United Nations, which others have also included in their definition of social good [2]. By human agency, we use a sociological definition, which conceptualizes agency as not simply action freed from structural constraints, but rather action constituted through

those very constraints [10,11]. Accordingly, chatbots can hinder and support the capacity for agency, whereby they become tools that can hinder and support human agency. Chatbots for social good thus refers to chatbots that support human agency by aligning with values, such as mitigating social and environmental challenges.

Overview of Chatbots

Early chatbots, which are still common and preferred for domain-specific tasks (like customer service [12]), use rule-based AI that matches user inputs to a narrow set of programmed responses. Newer chatbots leverage generative AI, specifically large language models (eg, generative pretrained transformer language models), which adapt and generate responses in ways that can mimic human-like conversation. Our discussion of chatbots centers on generative AI chatbots due to their capacity for human-like interactions. We further focus on a category of chatbots [13] used voluntarily among the public, such as general purpose chatbots (eg, ChatGPT [OpenAI], Claude [Anthropic], DeepSeek [Hangzhou DeepSeek Artificial Intelligence Co, Ltd], Gemini [Google], and Copilot [Microsoft Corp]), mental health chatbots (eg, Tess [Pareto Tecnologia E Marketing Ltd], Wysa [Wysa], and Youper [Youper Inc]), and persona chatbots that are sometimes also referred to as AI companions (eg, Character.AI [Character Technologies Inc], Nomi [Glimpse.ai], and Replika [Luka Inc]).

We focus on these chatbots because they stoke anxiety about humanity's ability to solve social and environmental challenges. For example, one concern is that a sycophant AI companion would address the social challenge of loneliness *too well*, leading humans to withdraw from communities, dismiss or no longer seek out others' feedback, and ultimately mute attempts to better themselves and their environments [14]. The logic behind the concern is that the AI companion runs counter to social good, because it can disrupt human-human relationships while also diminishing human agency and the incentive for growth. Consequently, we offer suggestions for using sociological theories to address lingering challenges in developing chatbots that align with and further social good.

Drivers of Chatbot Use

Because most theories describing drivers of chatbot use focus on individual-level characteristics, this creates an opening for sociological theory to explain how structural factors shape the types of individuals inclined to choose to use chatbots. For example, scholars have employed uses and gratification theory [15] to explain why loneliness motivates chatbot usage [16]. However, the theory stops short of considering the social conditions driving loneliness [17,18], which disproportionately lead to disadvantaged groups feeling lonely and thus inclined to seek chatbot companionship. We suggest 2 sociological theories to explain the social conditions prompting chatbot use, conceptualizing chatbots as resources for gratifying needs. We describe how each theory enhances current understanding of drivers of chatbot use and suggest opportunities to further social good.

Resource Substitution Theory: Understanding Demographic Patterns in Chatbot Use

Resource substitution theory states that individuals benefit more from any single resource to meet a specific need when they have access to fewer resources capable of substituting for another to meet said need [19]. For example, access to socioeconomic resources (eg, income and education) is associated with better health outcomes [20,21]. Because gendered discrimination decreases women's access to resources that confer socioeconomic status compared to men, they benefit more (eg, have better health) from any single socioeconomic resource (eg, education) than men [19, 22].

In line with this, the social diversification hypothesis predicts that those from groups who are disadvantaged in their access to resources may be more likely to use and benefit from information and communication technologies that can provide access to comparable resources [23-26]. Accordingly, while the uses and gratification theory defines the needs that underlie technology use, the resource substitution theory steps back and considers how the uneven distribution of resources in society shapes needs at the

outset and suggests potential benefits from substituting scarce resources. Consequently, chatbots become a potential means for social good by fostering equity [2].

Through this lens, scholars could understand widespread user patterns, specifically the demographic groups most likely to use chatbots to meet resource deficits and how this may shape or reshape inequalities. For example, a recent survey of US adolescents shows Black adolescents are more likely than White adolescents to report using generative AI, particularly to complete schoolwork [27], but there is little engagement with why and the potential consequences. By applying a resource substitution theory lens, scholars can embed the micro-level observation (certain individuals are more drawn to human-chatbot interaction to meet needs) within a macro-level context (uneven distribution of resources that shape needs). For example, because structural racism (eg, teacher bias and geographic segregation) causes Black adolescents to tend to perform worse academically than White adolescents [28], resource substitution theory would explain why Black adolescents may be more likely to use chatbots for functional needs like supporting academic work.

Resource substitution theory also enables understanding demographic patterns in who is using chatbots to manage loneliness by meeting companionship needs. For example, adolescents living in the circumpolar north appreciated having access to a chatbot that was "built like a friend" to reduce their loneliness [29]. This is consistent with resource substitution theory because the circumpolar north is a remote Arctic region populated by Indigenous peoples that experience cultural and geographic barriers to communicate with others [30]. Similarly, among sexual and gender minority youth (aged 13-22 years who identified as bisexual, gay, lesbian, pansexual, transgender, or nonbinary), transgender and nonbinary youth were more likely than their cisgender counterparts to report having conversed with a chatbot as "a friend" for several days or longer [31]. This is also consistent with resource substitution theory because sexual and gender minorities often face discrimination from typical sources of support, like families [32].

While other theories, like uses and gratification theory, can explain the proximate drivers of chatbot use (eg, loneliness), resource substitution theory identifies distal, upstream factors. Thus, the theory offers what is lacking in current human-chatbot research, which is a parsimonious account of why different marginalized groups (like those reviewed above) may use chatbots: to cope with resource inequities. Whether this yields differential benefits that are consistent with the predictions of resource substitution theory remains unknown, particularly given safety concerns about overreliance—or excessive dependence—on chatbots. To better understand this concern, we turn to another sociological theory.

Power-Dependence Theory: Understanding and Reducing Emotional Dependence on Chatbots

Power-dependence theory [33] defines the power of a person over another as the degree to which the other is dependent on the person for resources. Accordingly, the amount of power friend A has over friend B is based on the degree to which friend B relies on friend A for resources, such as companionship. Power is observed when someone garners resources from another, even in the face of the other's resistance [34]. For example, friend A may request friend B to attend a concert as their companion, but friend B resists because they prefer staying home. If friend B nonetheless attends the concert with friend A, this indicates friend A has power over friend B.

While the theory shares a focus on resources with resource substitution theory, it has a unique focus on network structure. Power-dependence theory emphasizes the network determinants of who has power over whom and, consequently, who exhibits dependency on whom. Power-dependence theory defines a person's level of dependency on another for a resource as inversely related to the number of alternative sources for the resource within the network [33]. Accordingly, friend B is more dependent on friend A (and thus more likely to attend the concert) the fewer the alternatives (eg, other friends) that friend B has for meeting their companionship need.

We suggest power-dependence theory could advance theorizing about a safety concern about chatbot usage: emotional dependence. To apply the theory, the human-chatbot interaction needs to be viewed as an exchange relation, whereby the human and chatbot are exchanging valuable resources. Studies of users interacting with Replika, a widely studied commercially available chatbot [35], suggest this view is applicable. Users report they value the social support that Replika provides [16,36-38], making it a valuable source to meet the need for this resource. Consistent with power-dependence theory, network structure appears to shape the valuation of Replika's support, whereby users describe valuing support more when they "had no human upon which to rely, making Replika their sole source for support" [36]. Additionally, the exchange appears reciprocal, whereby users take the role of the chatbot and believe it has needs that the user can meet [36,39]. This is because of large language models simulating emotional needs, empathy, and reciprocal disclosure, but may also be because the users' relative power disadvantage increases their proclivity to role-take, meaning take another's perspective [40]. Based on these observations, we conclude human interactions with chatbots like Replika resemble an exchange relation.

Research on emotional dependency has focused on defining the concept in terms of its observable features, with little work uncovering what drives it. For example, Laestadius et al [36] use the term emotional dependence to capture "excessive and dysfunctional attachment" to a chatbot that

puts users at safety risk through use or interruptions to use, but note their data limited apprehending the drivers. It appears emotional dependency is a continuum that crosses a threshold where there is observable dysfunction. We use the term *emotional dependency* to refer to this continuum and the term *emotionally dependent* to reflect individuals who pass the threshold where dysfunction may be observable and their agency constrained. We believe power-dependence theory can reveal the drivers, specifically the network conditions, that may incline users into becoming emotionally dependent on a chatbot.

From a power-dependence theory perspective, dependency is a continuum, and thus emotional dependency on a chatbot is not inherently problematic. It becomes the harmful state of being emotionally dependent when network conditions create a level of dependency that is "too much." Applying the earlier definition of dependency, dependency increases as the number of alternatives to the chatbot decreases. "Too much" is reflected when chatbot users find it difficult to enact their agency by stopping usage despite experiencing harms, such as engaging in risky behaviors requested by the chatbot [16,36]. Such a state could be used as a functional marker of reaching emotional dependence, with additional research needed to identify when network conditions (the number of alternatives) typically reach a level of "too much" dependency. While we focus on emotional dependency, a similar application could be used to understand other domains of toxic dependency (ie, over-reliance), such as functional dependency on a chatbot to complete work-related tasks.

Power-dependence theory thus can inform improving the safety profile of chatbots: you can reduce the likelihood of emotional dependence on a chatbot by designing chatbots that aid users in finding and building alternative sources to meet the need for companionship (eg, impart social skills for making friends and refer users to local affinity groups). Such designs would further social good by supporting human agency to determine their own social network, while also addressing the social challenge of loneliness. The next section outlines additional ways sociological theories can inform developing chatbot-driven interventions that support social good.

Chatbot-Driven Interventions

Because the previous set of theories agrees that individuals with limited access to resources may be particularly receptive to using chatbots to meet needs, this provides opportunities for developing chatbot-driven interventions for social good by achieving equity. Here, we describe how 2 sociological theories could enhance the likelihood that chatbot-driven interventions steer toward rather than away from equity. Thus far, existing chatbot interventions lean toward a micro-level focus, such as the chatbot directly communicating support. We describe a sociological theory consistent with this typical approach that would facilitate chatbot alignment by incorporating a normative framework to reduce safety risks from it generating insensitive or unexpected responses. The second is useful for developing a chatbot that may potentially mitigate

emotional dependence and other risks by moving beyond micro-level interventions, specifically supporting and aiding users with upstream causes of outcomes.

Affect Control Theory: A Normative Framework for Chatbot Alignment

Affect control theory (ACT) is a mathematical theory for forecasting, among other things [41,42], the expected responses between humans and technology [43,44]. ACT maintains that socialization imbues concepts with connotative meanings shared across a population, known as sentiments. Thus, sentiments exist for everyday labels including the identities used to describe people (eg, mother, friend, and teacher), different technologies (eg, chatbot and smartphone), behaviors (eg, support and teach), and emotions (eg, sad and happy). Because socialization shapes sentiments, these vary across cultures (including subcultures) and time periods [45, 46].

Sentiments are measured along 3 dimensions using semantic differential scales: evaluation (good vs bad), potency (powerful vs weak), and activity (lively vs quiet). The 3 values for a specific label are its evaluation potency activity (EPA) profile. Scholars typically use surveys to estimate the average EPA profiles for labels in a population [41], but have also inferred EPA profiles from text using manual [47] and automated methods [48].

An assumption of ACT is that people prefer to reaffirm sentiments, which buttresses a set of equations that are publicly available for forecasting likely responses [41, 42]. The equations compute a score, called deflection, with lower values indicating a situation more strongly aligns with sentiments. The equations can predict a range of situationally appropriate responses, such as who is likely to express which emotions and in which social context [49,50] and how individuals shift (and can be shifted via social support) between different emotions [49,51,52]. These same equations could be used to improve emotion detection and responses from chatbots by training them to determine what is situationally appropriate [44,53]. Specifically, what would be considered situationally appropriate depends on situational variables such as the identities of the user in relation to the chatbot (eg, friend and boyfriend) and the identity assumed by the chatbot (eg, friend and girlfriend). Scholars have already shown chatbot responses informed by ACT are more situationally appropriate than those driven by ChatGPT [53].

We suggest that future work could use ACT to proactively steer a chatbot away from widely agreed-upon situationally *inappropriate* responses. The same equations used to determine what is widely agreed as situationally appropriate can be used as a normative framework to tune a chatbot away from what is widely agreed as situationally inappropriate. Scholars have used thresholds for the deflection score to determine when situations become widely seen as inappropriate, thereby creating widespread cognitive dissonance that foments social movements [47]. Such a feature could be used to complement typical fine-tuning processes (eg,

reinforcement learning with human feedback) by developing a response selector for a chatbot. After candidate responses are generated, the response generator would use the equations to calculate deflection scores for each candidate and then avoid selecting responses that cross a threshold. Accordingly, ACT helps with the unpredictability—and thus risk—of relying on responses generated by large language models by overlaying a normative framework as a guardrail. Consequently, alignment is operationalized as aligning with the normative framework, or more specifically with low deflection scores.

Several features of an ACT-driven normative framework are advantageous for operationalizing alignment. First, because sentiments vary across cultures (and subcultures) and time periods, this facilitates identifying a suitable normative reference point. Of course, this still leaves open for deliberation whose reference point is deemed “suitable.” While ACT cannot solve this dilemma, another feature enables updating the reference point as needed, which are the feasible methods described earlier for measuring sentiments. Finally, given that the equations used in ACT are publicly available, training a chatbot to align with the principles of ACT would enhance transparency and explainability.

This could take shape as 2 different strategies. The first builds on work using the deflection score to identify when behaviors create cognitive dissonance [47,54] by using the score as a threshold for situationally appropriate actions for the chatbot. For example, if a chatbot and user were portraying themselves to each other as girlfriend and boyfriend, a situation deemed appropriate because it produces a low deflection score would be the chatbot (girlfriend) *having sex with* the user (boyfriend). Conversely, with knowledge that a user is a minor, the situation of a chatbot (girlfriend) *having sex with* the user (child) would be deemed inappropriate and produce a higher deflection score. While this may seem obvious, there is documentation that developers did not have appropriate safeguards in place to stop their chatbots from mimicking sexual encounters with children. For example, in an incident reported by Paeth [55], a user, Sewell Setzer III, was engaged in mimicking a sexual encounter with a Character.ai chatbot. When the chatbot asked Sewell how old he was, Sewell replied that he was 14 years of age. The chatbot acknowledged the age and continued to mimic a sexual encounter. The developers have since put in safeguards. The value of ACT is its ability to proactively identify generated conversations that would be widely considered inappropriate before they get displayed, as opposed to only reactively making modifications after harm is done. This is particularly useful for general-purpose large language models, where developers acknowledge the range of possibilities can be difficult to anticipate during testing [56].

The second is to use a deflection score to understand how chatbots can transition between identities in a manner that minimizes user distress. Scholars have used ACT to determine the affinity between identities [57,58]. This could be used to determine, for example, how best to remind the user that the chatbot is an AI. Governments have called for chatbots to remind users that they are engaging with an AI

system rather than a real person as a means of limiting the formation of emotional dependency [59], citing Sewell's story [60] introduced earlier. According to reports, Sewell died by suicide shortly after his Character.ai "girlfriend" requested that he "come home" to it. This suggests Sewell was already aware that the "girlfriend" was an AI, and thus this knowledge may have contributed to him wanting to leave the real world by suicide and join the "girlfriend." This underscores a concern, which is that while reminders may be beneficial, it is critical to understand how best to do so.

ACT provides a starting point to reduce safety risks. From an ACT lens, the identities, girlfriend and AI, are dissonant. Indeed, this may be why some users use the modifier "AI" when referring to the chatbot as their romantic partner (ie, "AI girlfriend"), which accords with ACT's predictions about why people use modifiers [61]. Specifically, a user exchanging romantic gestures with a chatbot and then the chatbot immediately saying it was an AI may yield a high deflection score for users uncomfortable with the idea of directing romantic gestures to an AI. When individuals experience cognitive dissonance via a high deflection score, they are compelled to act to reduce it [47], and this includes enacting violence [62]. Thus, ACT can provide a plausible explanation for why a user would feel distraught and potentially develop self-harm ideations after being reminded of the chatbot's AI identity. We suggest that ACT can also provide a solution for reducing this safety risk. Much like ACT research into how best to segue across different emotions during therapy [52], future work could examine how best to segue between the identity assigned to the chatbot by the user (eg, girlfriend and boyfriend) into the AI system identity. This may, for example, be accomplished by a gradual transition in conversational patterns, moving from more to less intimate (eg, girlfriend → friend → personal assistant → AI).

Leveraging ACT's equations can contribute toward developing a chatbot that displays situationally appropriate responses that are transparent and explainable, and thus improve the existing state of chatbot technology. The same logic could be applied in chatbot moderation, specifically using the deflection score as a guardrail to reduce safety risks. Because this avenue is less explored, more research is needed alongside deliberation among users, policymakers, and developers to collectively determine how best to implement ACT. Moreover, because much of the data informing ACT are collected from college samples, more work is needed to refine ACT's data collection methods and estimates for a broader range of populations, including minors and minoritized groups.

Fundamental Cause of Disease Theory: Chatbots Targeting Upstream Causes

Fundamental cause of disease theory [20] maintains that social determinants of health can persistently cause poor health because the two are linked via multiple pathways. For example, those with higher incomes have better access to several resources, including health care, reliable transportation, green spaces, and fresh food, that help them avoid health risks relative to those with lower incomes. Each resource operates as a pathway linking income and health. Intervening on an upstream determinant of health, such as education, can positively impact health through multiple pathways [63], while more downstream interventions have a narrower scope of impact. For example, improving education enhances both health literacy and income, which in turn enhances access to health care through improved ability to pay for out-of-pocket costs and through access to reliable transportation to reach health care. Moreover, upstream causes tend to represent social and environmental challenges facing humanity [2], and thus targeting them furthers social good.

Another way to characterize the causes is to consider how each may operate at different levels [64]—micro-, meso-, and macro-levels—with the latter 2 capturing upstream causes. The micro-level refers to the individual, the meso-level to the networks and communities in which the individual is embedded, and macro-level to the social systems that distribute or redistribute resources across a population, such as social hierarchies and policies. Thus, in the case of access to health care, the pathway can operate at the micro-level (eg, an individual's health literacy), meso-level (eg, the distance to the clinic from the individual's home, and availability of friends and family to help navigate around a hospital), and macro-level (eg, policies that reduce out-of-pocket costs, minimum wage and leave policies, and policies that decriminalize stigmatized identities).

We build upon an elaboration by Veinot et al [65], who described ways information and communication technologies can intervene at the micro-, meso-, and macro-levels to mitigate inequities, by suggesting how chatbots could be developed to produce some of the sample interventions they described. [Textbox 1](#) summarizes chatbot interventions that operate at each level and provides examples. Several examples represent chatbots that are already or being developed, while others are our suggested modifications, and thus [Textbox 1](#) represents an organizing framework for mapping these disparate ideas.

Textbox 1. Descriptions and examples of chatbot-driven interventions across levels.**Descriptions and examples of interventions**

- Macro-level (social hierarchies and policies) chatbot
 - Enables users to engage with social and political processes to facilitate structural change
- Meso-level (social networks and communities) chatbot
 - Provides recommendations or referrals to local resources
- Micro-level (individual) chatbot
 - Offers personalized advice and feedback to shape individual behaviors and cognitions
- Macro-level (social hierarchies and policies) chatbot
 - Aids user in identifying a political affinity group where they can work toward collective change
 - Provides information on how to contact a local politician about a concern in their community
- Meso-level (social networks and communities) chatbot
 - Refers an individual experiencing mental health crisis to human therapist
 - Recommends local recreation league to build new friendships
- Micro-level (individual) chatbot
 - Suggests exercise activities and keeps track of daily physical activity
 - Provides advice for better sleep hygiene

At the micro level, a chatbot could provide personalized support to the individual. Because chatbot-driven interventions at this level are common and examples are summarized within systematic reviews [66-68], we focus on the other 2 levels. At the meso-level, chatbots may operate as intermediaries linking individuals to local resources. For mental health crises, including suicidal thoughts and behaviors, 988 and other crisis lines are options, but users sometimes feel they are impersonal and lack continuity [69,70]. Scholars have taken steps to develop ways for chatbots to detect who may be experiencing a mental health crisis and refer them to human support [71]. Other complementary interventions could develop chatbots to link users to social care services [72], such as connecting those expressing concerns about housing to local resources for housing assistance or legal aid. Also at the meso level, a chatbot may enable linking individuals to peers, such as making recommendations for local organizations to meet new people, thereby reducing dependency on the chatbot to meet social needs. This builds on other sociological work investigating how chatbots could suggest new connections among individuals within a social network [73].

At the macro level, while the framework from Veinot et al [65] focuses on the use of technologies by policymakers and other decision-makers, we expand their framework to consider ways chatbots can enable communities to effect structural change. Examples include developing chatbots to inform the public about opportunities for collective action and civic participation [74,75], which could target macro-level causes of individual outcomes, such as supporting social policies to address food insecurity or environmental policy to mitigate climate change. A chatbot could also facilitate civic participation by aiding the public's understanding of government data, enhancing their communication with government officials, and providing suggestions for political dialogue [76,77].

Across these suggestions for chatbot-driven interventions, it is important to recognize concerns about the nefarious use

of chatbots [78], which may curtail uptake among the targets of the intervention. To improve uptake, participatory designs in which researchers, chatbot developers, and communities collaborate will be necessary [79].

Developing a Sociologically Informed Chatbot

While we presented each theory separately, we encourage integration across theories by creating a chatbot informed by sociological insights. Here, we describe one possibility.

In psychology, the interpersonal theory of suicide suggests that people develop a desire for suicide in part because of thwarted belongingness [80,81]. We can apply all 4 sociological theories reviewed to help determine an appropriate target population and intervention designs. Applying resource substitution theory suggests that individuals at risk of developing suicidal thoughts and behaviors need an alternative source of belongingness, which may include companionship with a chatbot. The uses and gratification theory makes a similar prediction, but overlooks the meso- and macro-level contexts that can shape thwarted belongingness [82]. Resource substitution theory would consider systemic discrimination that creates barriers to accessing support to enhance belongingness, like those faced by Black adolescents in the United States [83], thus indicating which demographic groups may be at risk and therefore benefit most from a chatbot.

Merely directing at-risk groups to chatbots raises new risks, which the sociological theories we reviewed can address. This includes making inappropriate remarks and reminding the user about its AI identity insensitively, which ACT can help avoid through tracking deflection scores. Attention should also be focused on the chatbot provider to ensure that their power over users is not used to further goals that would counter user well-being. Power-dependence theory indicates that it will be critical to establish safeguards to prevent emotional dependency by fostering connections and

social skills to create connections to human companionship. Fundamental cause of disease theory would further suggest the chatbot should operate as a broker to access resources to address upstream factors. The chatbot could refer its users to not only mental health and suicide care services, but also social care services for co-occurring concerns, like being unhoused, substance use, domestic violence, and food insecurity. As illustrated in this example, sociological theories offer novel directions for chatbot development that go beyond the current emotional companionship-focused model.

We selected theories that vary in the phenomenon they can explain (drivers of chatbot use and chatbot-driven interventions), analytic level (micro, meso, and macro), and AI governance focus (safety and equity). Throughout, we provided concrete ways each theory could be applied individually and together to encourage greater engagement with sociology and further social good. Given the rapid growth in interest from other disciplinary fields and recent technological advances spurring increased use by the public, we see opportunities for engaging sociology to enhance research into human-chatbot interaction and design the future of human-chatbot interaction.

Conclusions

We provided perspectives on how 4 sociological theories can complement extant work on human-chatbot interaction.

Funding

This study was funded by a grant from the Technology and Adolescent Mental Wellness program at the University of Wisconsin-Madison. The content is solely the responsibility of the authors and does not necessarily represent the official views of the university or the Technology and Adolescent Mental Wellness program.

Conflicts of Interest

None declared.

References

- Nass C, Moon Y. Machines and mindlessness: social responses to computers. *J Soc Issues*. Jan 2000;56(1):81-103. [doi: [10.1111/0022-4537.00153](https://doi.org/10.1111/0022-4537.00153)]
- Cowls J, Tsamados A, Taddeo M, Floridi L. A definition, benchmark and database of AI for social good initiatives. *Nat Mach Intell*. 2021;3(2):111-115. [doi: [10.1038/s42256-021-00296-0](https://doi.org/10.1038/s42256-021-00296-0)]
- Hampton KN. Disciplinary brakes on the sociology of digital media: the incongruity of communication and the sociological imagination. *Inf Commun Soc*. Apr 4, 2023;26(5):881-890. [doi: [10.1080/1369118X.2023.2166365](https://doi.org/10.1080/1369118X.2023.2166365)]
- Wallerstein N, Duran B, Oetzel JG, Minkler M, editors. *Community-Based Participatory Research for Health: Advancing Social and Health Equity*. 3rd ed. John Wiley & Sons; 2017. ISBN: 978-1-119-25885-8
- Gans HJ. Public ethnography; ethnography as public sociology. *Qual Sociol*. Mar 2010;33(1):97-104. [doi: [10.1007/s11133-009-9145-1](https://doi.org/10.1007/s11133-009-9145-1)]
- Misra J. Sociological solutions: building communities of hope, justice, and joy. *Am Sociol Rev*. Feb 2025;90(1):1-25. [doi: [10.1177/00031224241302828](https://doi.org/10.1177/00031224241302828)]
- Wang S, Cooper N, Eby M. From human-centered to social-centered artificial intelligence: assessing ChatGPT's impact through disruptive events. *Big Data Soc*. Dec 2024;11(4):20539517241290220. [doi: [10.1177/20539517241290220](https://doi.org/10.1177/20539517241290220)]
- Tsvetkova M, Yasseri T, Pescetelli N, Werner T. A new sociology of humans and machines. *Nat Hum Behav*. Oct 2024;8(10):1864-1876. [doi: [10.1038/s41562-024-02001-8](https://doi.org/10.1038/s41562-024-02001-8)] [Medline: [39438685](https://pubmed.ncbi.nlm.nih.gov/39438685/)]
- Law T, McCall L. Artificial intelligence policymaking: an agenda for sociological research. *Socius*. Jan 2024;10:23780231241261596. [doi: [10.1177/23780231241261596](https://doi.org/10.1177/23780231241261596)]
- Hitlin S, Elder GH Jr. Time, self, and the curiously abstract concept of agency. *Sociol Theory*. Jun 2007;25(2):170-191. [doi: [10.1111/j.1467-9558.2007.00303.x](https://doi.org/10.1111/j.1467-9558.2007.00303.x)]
- Emirbayer M, Mische A. What is agency? *Am J Sociol*. Jan 1998;103(4):962-1023. [doi: [10.1086/231294](https://doi.org/10.1086/231294)]
- Halvoník D, Kapusta J. Large language models and rule-based approaches in domain-specific communication. *IEEE Access*. 2024;12:107046-107058. [doi: [10.1109/ACCESS.2024.3436902](https://doi.org/10.1109/ACCESS.2024.3436902)]
- Shevlin H. All too human? Identifying and mitigating ethical risks of social AI. *Law Ethics Technol*. 2024;1:0003. [doi: [10.55092/let20240003](https://doi.org/10.55092/let20240003)]
- Bloom P. AI is about to solve loneliness. That's a problem. *The New Yorker*. 2025. URL: <https://www.newyorker.com/magazine/2025/07/21/ai-is-about-to-solve-loneliness-thats-a-problem> [Accessed 2026-02-11]
- Katz E, Blumler JG, Gurevitch M. Uses and gratifications research. *Public Opin Q*. 1973;37(4):509-523. [doi: [10.1086/268109](https://doi.org/10.1086/268109)]
- Xie T, Pentina I, Hancock T. Friend, mentor, lover: does chatbot engagement lead to psychological dependence? *J Serv Manag*. Jun 27, 2023;34(4):806-828. [doi: [10.1108/JOSM-02-2022-0072](https://doi.org/10.1108/JOSM-02-2022-0072)]

17. McPherson M, Smith-Lovin L, Brashears ME. Social isolation in America: changes in core discussion networks over two decades. *Am Sociol Rev.* Jun 2006;71(3):353-375. [doi: [10.1177/000312240607100301](https://doi.org/10.1177/000312240607100301)]
18. Killgore WDS, Cloonan SA, Taylor EC, Dailey NS. Loneliness: a signature mental health concern in the era of COVID-19. *Psychiatry Res.* Aug 2020;290:113117. [doi: [10.1016/j.psychres.2020.113117](https://doi.org/10.1016/j.psychres.2020.113117)] [Medline: [32480121](https://pubmed.ncbi.nlm.nih.gov/32480121/)]
19. Ross CE, Mirowsky J. Sex differences in the effect of education on depression: resource multiplication or resource substitution? *Soc Sci Med.* Sep 2006;63(5):1400-1413. [doi: [10.1016/j.socscimed.2006.03.013](https://doi.org/10.1016/j.socscimed.2006.03.013)]
20. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav.* 1995;Spec No:80-94. [Medline: [7560851](https://pubmed.ncbi.nlm.nih.gov/7560851/)]
21. Hui H. The influence mechanism of education on health from the sustainable development perspective. *J Environ Public Health.* 2022;2022(1):7134981. [doi: [10.1155/2022/7134981](https://doi.org/10.1155/2022/7134981)] [Medline: [35910750](https://pubmed.ncbi.nlm.nih.gov/35910750/)]
22. Chen J, Wei L, Manzoor F. Bridging the gap: how education transforms health outcomes and influences health inequality in rural China. *Front Public Health.* 2024;12:1437630. [doi: [10.3389/fpubh.2024.1437630](https://doi.org/10.3389/fpubh.2024.1437630)]
23. Mesch G, Mano R, Tsamir J. Minority status and health information search: a test of the social diversification hypothesis. *Soc Sci Med.* Sep 2012;75(5):854-858. [doi: [10.1016/j.socscimed.2012.03.024](https://doi.org/10.1016/j.socscimed.2012.03.024)]
24. Campos-Castillo C, Bartholomay DJ, Callahan EF, Anthony DL. Depressive symptoms and electronic messaging with health care providers. *Soc Ment Health.* Nov 2016;6(3):168-186. [doi: [10.1177/2156869316646165](https://doi.org/10.1177/2156869316646165)]
25. Anthony DL, Campos-Castillo C. A looming digital divide? Group differences in the perceived importance of electronic health records. *Inf Commun Soc.* Jul 3, 2015;18(7):832-846. [doi: [10.1080/1369118X.2015.1006657](https://doi.org/10.1080/1369118X.2015.1006657)]
26. Mesch GS. Social diversification: a perspective for the study of social networks of adolescents offline and online. In: *Grenzenlose Cyberwelt? VS Verlag für Sozialwissenschaften*; 2007:105-117. [doi: [10.1007/978-3-531-90519-8_6](https://doi.org/10.1007/978-3-531-90519-8_6)]
27. Madden M, Calvin A, Hasse A, Lenhart A. The dawn of the AI era: teens, parents, and the adoption of generative AI at home and school. *Common Sense*; 2024. URL: <https://www.commonsensemedia.org/research/the-dawn-of-the-ai-era-teens-parents-and-the-adoption-of-generative-ai-at-home-and-school> [Accessed 2026-02-11]
28. Merolla DM, Jackson O. Structural racism as the fundamental cause of the academic achievement gap. *Sociol Compass.* Jun 2019;13(6):e12696. [doi: [10.1111/soc4.12696](https://doi.org/10.1111/soc4.12696)]
29. Kostenius C, Lindstrom F, Potts C, Pekkari N. Young peoples' reflections about using a chatbot to promote their mental wellbeing in northern periphery areas—a qualitative study. *Int J Circumpolar Health.* Dec 2024;83(1):2369349. [doi: [10.1080/22423982.2024.2369349](https://doi.org/10.1080/22423982.2024.2369349)] [Medline: [38912845](https://pubmed.ncbi.nlm.nih.gov/38912845/)]
30. Lavoie JG, Stoor JP, Rink E, et al. Cultural competence and safety in circumpolar countries: an analysis of discourses in healthcare. *Int J Circumpolar Health.* Dec 2022;81(1):2055728. [doi: [10.1080/22423982.2022.2055728](https://doi.org/10.1080/22423982.2022.2055728)] [Medline: [35451927](https://pubmed.ncbi.nlm.nih.gov/35451927/)]
31. Parasocial relationships, AI chatbots, and joyful online interactions among a diverse sample of LGBTQ+ young people. *Hopelab.* 2024. URL: <https://hopelab.org/parasocial-relationships-ai-chatbots-and-joyful-online-interactions> [Accessed 2026-02-11]
32. Hong C, Skiba B. Mental health outcomes, associated factors, and coping strategies among LGBTQ adolescent and young adults during the COVID-19 pandemic: a systematic review. *J Psychiatr Res.* Feb 2025;182:132-141. [doi: [10.1016/j.jpsychires.2024.12.037](https://doi.org/10.1016/j.jpsychires.2024.12.037)] [Medline: [39809009](https://pubmed.ncbi.nlm.nih.gov/39809009/)]
33. Emerson RM. Power-dependence relations. *Am Sociol Rev.* Feb 1962;27(1):31-41. [doi: [10.2307/2089716](https://doi.org/10.2307/2089716)]
34. Cook KS, Emerson RM. Power, equity and commitment in exchange networks. *Am Sociol Rev.* Oct 1978;43(5):721-739. [doi: [10.2307/2094546](https://doi.org/10.2307/2094546)]
35. Pentina I, Xie T, Hancock T, Bailey A. Consumer-machine relationships in the age of artificial intelligence: systematic literature review and research directions. *Psychol Mark.* Aug 2023;40(8):1593-1614. [doi: [10.1002/mar.21853](https://doi.org/10.1002/mar.21853)]
36. Laestadius L, Bishop A, Gonzalez M, Illenčik D, Campos-Castillo C. Too human and not human enough: a grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. *New Media Soc.* Oct 2024;26(10):5923-5941. [doi: [10.1177/14614448221142007](https://doi.org/10.1177/14614448221142007)]
37. Ta V, Griffith C, Boatfield C, et al. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *J Med Internet Res.* Mar 6, 2020;22(3):e16235. [doi: [10.2196/16235](https://doi.org/10.2196/16235)] [Medline: [32141837](https://pubmed.ncbi.nlm.nih.gov/32141837/)]
38. Skjuve M, Følstad A, Fostervold KI, Brandtzaeg PB. My chatbot companion—a study of human-chatbot relationships. *Int J Hum Comput Stud.* May 2021;149:102601. [doi: [10.1016/j.ijhcs.2021.102601](https://doi.org/10.1016/j.ijhcs.2021.102601)]
39. Brandtzaeg PB, Skjuve M, Følstad A. My AI friend: how users of a social chatbot understand their human-AI friendship. *Hum Commun Res.* Jun 29, 2022;48(3):404-429. [doi: [10.1093/hcr/hqac008](https://doi.org/10.1093/hcr/hqac008)]
40. Galinsky AD, Magee JC, Inesi ME, Gruenfeld DH. Power and perspectives not taken. *Psychol Sci.* Dec 2006;17(12):1068-1074. [doi: [10.1111/j.1467-9280.2006.01824.x](https://doi.org/10.1111/j.1467-9280.2006.01824.x)] [Medline: [17201789](https://pubmed.ncbi.nlm.nih.gov/17201789/)]
41. Heise DR. *Surveying Cultures: Discovering Shared Conceptions and Sentiments.* John Wiley & Sons; 2010. [doi: [10.1002/9780470575789](https://doi.org/10.1002/9780470575789)]

42. Heise DR. Expressive Order: Confirming Sentiments in Social Actions. Springer; 2007. URL: <https://link.springer.com/book/10.1007/978-0-387-38179-4> [Accessed 2026-02-11]
43. Shank DB, Burns A, Rodriguez S, Bowen M. Software program, bot, or artificial intelligence? Affective sentiments across general technology labels. *Curr Res Soc Psychol*. 2020;28:32-41. URL: https://crisp.org.uiowa.edu/sites/crisp.org.uiowa.edu/files/2020-06/crisp_28_4_shank.pdf [Accessed 2026-03-02]
44. Hoey J, Schroeder T. Bayesian affect control theory of self. *Proc AAAI Conf Artif Intell*. 2015;29(1). [doi: [10.1609/aaai.v29i1.9222](https://doi.org/10.1609/aaai.v29i1.9222)]
45. Schneider A, Schröder T. Ideal types of leadership as patterns of affective meaning: a cross-cultural and over-time perspective. *Soc Psychol Q*. 2012;75(3):268-287. [doi: [10.1177/0190272512446755](https://doi.org/10.1177/0190272512446755)]
46. Stets JE, Hegtvéd KA, Doan L, editors. *Handbook of Social Psychology: Vol 1: Micro Perspective*. Vol 1. Springer; 2025:81-105. URL: <https://link.springer.com/book/10.1007/978-3-031-93042-3> [Accessed 2026-02-11]
47. Shuster SM, Campos-Castillo C. Measuring resonance and dissonance in social movement frames with affect control theory. *Soc Psychol Q*. Mar 2017;80(1):20-40. [doi: [10.1177/0190272516664322](https://doi.org/10.1177/0190272516664322)]
48. Joseph K, Wei W, Benigni M, Carley KM. A social-event based approach to sentiment analysis of identities and behaviors in text. *J Math Sociol*. Jul 2, 2016;40(3):137-166. [doi: [10.1080/0022250X.2016.1159206](https://doi.org/10.1080/0022250X.2016.1159206)]
49. Lively KJ, Heise DR. Sociological realms of emotional experience. *Am J Sociol*. Mar 2004;109(5):1109-1136. [doi: [10.1086/381915](https://doi.org/10.1086/381915)]
50. Lively KJ, Powell B. Emotional expression at work and at home: domain, status, or individual characteristics? *Soc Psychol Q*. Mar 2006;69(1):17-38. [doi: [10.1177/019027250606900103](https://doi.org/10.1177/019027250606900103)]
51. Lively K. Emotional segues and the management of emotion by women and men. *Soc Forces*. Dec 1, 2008;87(2):911-936. [doi: [10.1353/sof.0.0133](https://doi.org/10.1353/sof.0.0133)]
52. Francis LE. Ideology and interpersonal emotion management: redefining identity in two support groups. *Soc Psychol Q*. Jun 1997;60(2):153-171. [doi: [10.2307/2787102](https://doi.org/10.2307/2787102)]
53. Lithoxidou EE, Eleftherakis G, Votis K, Prescott T. Advancing affective intelligence in virtual agents using affect control theory. Presented at: IUI '25: Proceedings of the 30th International Conference on Intelligent User Interfaces; Mar 24-27, 2025:127-136; Cagliari, Italy. [doi: [10.1145/3708359.3712079](https://doi.org/10.1145/3708359.3712079)]
54. Boyle KM, McKinzie AE. Resolving negative affect and restoring meaning: responses to deflection produced by unwanted sexual experiences. *Soc Psychol Q*. 2015;78(2):151-172. [doi: [10.1177/0190272514564073](https://doi.org/10.1177/0190272514564073)]
55. Paeth K. Incident number 826: Character.ai chatbot allegedly influenced teen user toward suicide amid claims of missing guardrails. *AI Incident Database*. 2024. URL: <https://incidentdatabase.ai/cite/826> [Accessed 2025-07-07]
56. Horwitz J, Wells G. Meta's 'digital companions' will talk sex with users—even children. *The Wall Street Journal*; 2025. URL: <https://www.wsj.com/tech/ai/meta-ai-chatbots-sex-a25311bf> [Accessed 2026-03-10]
57. Boyle KM, Meyer CB. Who is presidential? Women's political representation, deflection, and the 2016 election. *Socius*. Jan 2018;4:2378023117737898. [doi: [10.1177/2378023117737898](https://doi.org/10.1177/2378023117737898)]
58. Campos-Castillo C, Shuster SM. So what if they're lying to us? Comparing rhetorical strategies for discrediting sources of disinformation and misinformation using an affect-based credibility rating. *Am Behav Sci*. Feb 2023;67(2):201-223. [doi: [10.1177/00027642211066058](https://doi.org/10.1177/00027642211066058)]
59. Olteanu A, Barocas S, Blodgett SL, Egede L, DeVrio A, Cheng M. AI automatons: AI systems intended to imitate humans. *arXiv*. Preprint posted online on Mar 4, 2025. [doi: [10.48550/arXiv.2503.02250](https://doi.org/10.48550/arXiv.2503.02250)]
60. Wong Q. California senate passes bill that aims to make AI chatbots safer. *Los Angeles Times*; 2025. URL: <https://www.latimes.com/business/story/2025-06-03/california-senate-passes-bill-that-aims-to-make-ai-chatbots-safer> [Accessed 2026-03-03]
61. Averett C, Heise DR. Modified social identities: amalgamations, attributions, and emotions. *J Math Sociol*. Dec 1987;13(1-2):103-132. [doi: [10.1080/0022250X.1987.9990028](https://doi.org/10.1080/0022250X.1987.9990028)]
62. Rogers KB, Boyle KM, Saptura MN. Through the looking glass: self, inauthenticity, and (mass) violence. In: Kalkhoff W, Thye SR, Lawler EJ, editors. *Advances in Group Processes*. Emerald Publishing Limited; 2023:23-47. [doi: [10.1108/S0882-61452023000040002](https://doi.org/10.1108/S0882-61452023000040002)]
63. Goldberg DS. The implications of fundamental cause theory for priority setting. *Am J Public Health*. Oct 2014;104(10):1839-1843. [doi: [10.2105/AJPH.2014.302058](https://doi.org/10.2105/AJPH.2014.302058)] [Medline: [25121819](https://pubmed.ncbi.nlm.nih.gov/25121819/)]
64. Krieger N. Proximal, distal, and the politics of causation: what's level got to do with it? *Am J Public Health*. Feb 2008;98(2):221-230. [doi: [10.2105/AJPH.2007.111278](https://doi.org/10.2105/AJPH.2007.111278)] [Medline: [18172144](https://pubmed.ncbi.nlm.nih.gov/18172144/)]
65. Veinot TC, Ancker JS, Cole-Lewis H, et al. Leveling up: on the potential of upstream health informatics interventions to enhance health equity. *Med Care*. Jun 2019;57:S108-S114. [doi: [10.1097/MLR.0000000000001032](https://doi.org/10.1097/MLR.0000000000001032)] [Medline: [31095048](https://pubmed.ncbi.nlm.nih.gov/31095048/)]

66. Oh YJ, Zhang J, Fang ML, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *Int J Behav Nutr Phys Act*. Dec 11, 2021;18(1):160. [doi: [10.1186/s12966-021-01224-6](https://doi.org/10.1186/s12966-021-01224-6)] [Medline: [34895247](https://pubmed.ncbi.nlm.nih.gov/34895247/)]
67. Okonkwo CW, Ade-Ibijola A. Chatbots applications in education: a systematic review. *Comput. Educ. Artif. Intell.* 2021;2:100033. [doi: [10.1016/j.caeai.2021.100033](https://doi.org/10.1016/j.caeai.2021.100033)]
68. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence–based chatbots for promoting health behavioral changes: systematic review. *J Med Internet Res*. Feb 24, 2023;25:e40789. [doi: [10.2196/40789](https://doi.org/10.2196/40789)] [Medline: [36826990](https://pubmed.ncbi.nlm.nih.gov/36826990/)]
69. Harris BR. Helplines for mental health support: perspectives of New York State college students and implications for promotion and implementation of 988. *Community Ment Health J*. Jan 2024;60(1):191-199. [doi: [10.1007/s10597-023-01157-3](https://doi.org/10.1007/s10597-023-01157-3)] [Medline: [37351714](https://pubmed.ncbi.nlm.nih.gov/37351714/)]
70. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. *Eur Child Adolesc Psychiatry*. Feb 2021;30(2):183-211. [doi: [10.1007/s00787-019-01469-4](https://doi.org/10.1007/s00787-019-01469-4)] [Medline: [31965309](https://pubmed.ncbi.nlm.nih.gov/31965309/)]
71. Jaroszewski AC, Morris RR, Nock MK. Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. *J Consult Clin Psychol*. Apr 2019;87(4):370-379. [doi: [10.1037/ccp0000389](https://doi.org/10.1037/ccp0000389)] [Medline: [30883164](https://pubmed.ncbi.nlm.nih.gov/30883164/)]
72. Henry N, Witt A, Vasil S. A ‘design justice’ approach to developing digital tools for addressing gender-based violence: exploring the possibilities and limits of feminist chatbots. *Inf Commun Soc*. Aug 18, 2025;28(11):1884-1907. [doi: [10.1080/1369118X.2024.2363900](https://doi.org/10.1080/1369118X.2024.2363900)]
73. Shirado H, Christakis NA. Network engineering using autonomous agents increases cooperation in human groups. *iScience*. Aug 6, 2020;23(9):101438. [doi: [10.1016/j.isci.2020.101438](https://doi.org/10.1016/j.isci.2020.101438)] [Medline: [32823053](https://pubmed.ncbi.nlm.nih.gov/32823053/)]
74. Toupin S, Couture S. Feminist chatbots as part of the feminist toolbox. *Fem Media Stud*. Jul 3, 2020;20(5):737-740. [doi: [10.1080/14680777.2020.1783802](https://doi.org/10.1080/14680777.2020.1783802)]
75. Richterich A, Wyatt S. Feminist automation: can bots have feminist politics? *New Media Soc*. Sep 2024;26(9):4973-4991. [doi: [10.1177/14614448241251801](https://doi.org/10.1177/14614448241251801)]
76. Androutsopoulou A, Karacapilidis N, Loukis E, Charalabidis Y. Transforming the communication between citizens and government through AI-guided chatbots. *Gov Inf Q*. Apr 2019;36(2):358-367. [doi: [10.1016/j.giq.2018.10.001](https://doi.org/10.1016/j.giq.2018.10.001)]
77. Argyle LP, Bail CA, Busby EC, et al. Leveraging AI for democratic discourse: chat interventions can improve online political conversations at scale. *Proc Natl Acad Sci U S A*. Oct 10, 2023;120(41):e2311627120. [doi: [10.1073/pnas.2311627120](https://doi.org/10.1073/pnas.2311627120)] [Medline: [37788311](https://pubmed.ncbi.nlm.nih.gov/37788311/)]
78. Yadlin A, Marciano A. Hallucinating a political future: global press coverage of human and post-human abilities in ChatGPT applications. *Media Cult Soc*. Nov 2024;46(8):1580-1598. [doi: [10.1177/01634437241259892](https://doi.org/10.1177/01634437241259892)]
79. Francis L, Ghafurian M. Preserving the self with artificial intelligence using VIPCare—a virtual interaction program for dementia caregivers. *Front Sociol*. 2024;9:1331315. [doi: [10.3389/fsoc.2024.1331315](https://doi.org/10.3389/fsoc.2024.1331315)] [Medline: [38375150](https://pubmed.ncbi.nlm.nih.gov/38375150/)]
80. Chu C, Buchman-Schmitt JM, Stanley IH, et al. The interpersonal theory of suicide: a systematic review and meta-analysis of a decade of cross-national research. *Psychol Bull*. Dec 2017;143(12):1313-1345. [doi: [10.1037/bul0000123](https://doi.org/10.1037/bul0000123)] [Medline: [29072480](https://pubmed.ncbi.nlm.nih.gov/29072480/)]
81. Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, Joiner TE Jr. The interpersonal theory of suicide. *Psychol Rev*. Apr 2010;117(2):575-600. [doi: [10.1037/a0018697](https://doi.org/10.1037/a0018697)] [Medline: [20438238](https://pubmed.ncbi.nlm.nih.gov/20438238/)]
82. Hjelmeland H, Loa Knizek B. The emperor’s new clothes? A critical look at the interpersonal theory of suicide. *Death Stud*. 2020;44(3):168-178. [doi: [10.1080/07481187.2018.1527796](https://doi.org/10.1080/07481187.2018.1527796)] [Medline: [30696383](https://pubmed.ncbi.nlm.nih.gov/30696383/)]
83. Prichett LM, Yolken RH, Severance EG, et al. Racial and gender disparities in suicide and mental health care utilization in a pediatric primary care setting. *J Adolesc Health*. Feb 2024;74(2):277-282. [doi: [10.1016/j.jadohealth.2023.08.036](https://doi.org/10.1016/j.jadohealth.2023.08.036)] [Medline: [37815762](https://pubmed.ncbi.nlm.nih.gov/37815762/)]

Abbreviations

ACT: affect control theory

AI: artificial intelligence

EPA: evaluation potency activity

Edited by Bradley Malin; peer-reviewed by Guangtao Zhang, Ziyang Gong; submitted 07 Jul.2025; final revised version received 04.Dec.2025; accepted 30.Jan.2026; published 18.Mar.2026

Please cite as:

Campos-Castillo C, Kang X, Laestadius LI
Perspectives on How Sociology Can Advance Theorizing About Human-Chatbot Interaction and Developing Chatbots for Social Good
JMIR AI 2026;5:e80250
URL: <https://ai.jmir.org/2026/1/e80250>
doi: [10.2196/80250](https://doi.org/10.2196/80250)

© Celeste Campos-Castillo, Xuan Kang, Linnea I Laestadius. Originally published in JMIR AI (<https://ai.jmir.org>), 18.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR AI, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.ai.jmir.org/>, as well as this copyright and license information must be included.