RUNNING HEAD: HUMANIZING ERRORS

"Hello! How May I Helo You?":

How (Corrected) Errors Humanize a Communicator

Shirley Bluvstein Netter[a,1], Xuan Zhao[b,1,*], Alixandra Barasch[a], Juliana Schroeder[c]

Author Affiliations:

[a] New York University Stern School of Business

[b] Stanford University, Department of Psychology

[c] University of California Berkeley Haas School of Business

Corresponding author:

Xuan Zhao

450 Jane Stanford Way, Building 420, Stanford University,

Stanford, CA 94305

xuanzhao@stanford.edu

Author notes:

1. The first two authors contributed equally.

"Hello! How May I Helo You?":

How (Corrected) Errors Humanize a Communicator

Abstract

Today more than ever before, online text-based interactions (e.g., text messages, emails, social media) have become a primary means of communication. Because written communication lacks human nonverbal cues such as appearance, voice, and identity, consumers may struggle to distinguish whether they are interacting online with a human or a chatbot. The current research investigates how typographical errors ("typos"), a common yet overlooked feature in text communication, can humanize a communicator. Across five experiments (N = 2,515) that used ambiguous conversational counterparts (i.e., customer service agents that might be bots), agents (either chatbots or real humans) who made and subsequently corrected a typo were perceived to be more humanlike than ones who made no typo or did not correct the typo. Participants consequently perceived those agents as warmer and more capable of understanding and helping their issues, were more likely to endorse a reward for the agent, and even perceived the company they represented more favorably. These findings provide novel insights into how conversational features may influence customers' perception of online agents and the brands that use them. The authors discuss theoretical implications for anthropomorphism and social perception and practical implications for companies wishing to humanize their customer service agents.

Words: 199

Keywords: humanization; chatbot; typo; error; company perception

Statement of Intended Contribution

With text-based online chats prevailing the current customer service industry, our manuscript presents a novel investigation into an ever-present yet overlooked aspect of these communications: typographical errors ("typos"). Over a series of experiments featuring novel methodological tools modeled after online chats with company representatives, we examine how making and correcting typos influences customers' perceptions of service agents and companies. This research provides at least three important insights for academics, practitioners, and policy makers.

First, this work demonstrates how typos embedded in dynamic online conversations can impact interpersonal impressions in a surprisingly positive way. In particular, our research stresses that it is not just making typos *per se*, but acknowledging and addressing the typo that is key to conveying humanness over text-based online chats. Second, by focusing on humanness perception as the central construct and investigating its antecedents and consequences, this research sheds light on psychological processes that are often salient in customers' minds when engaging with online service agents, yet rarely examined in the marketing literature—i.e., "am I talking to a real person?" We provide evidence that perceiving a chat agent as human matters: it increases the agent's perceived warmth and job-related competency and may even boost a company's image. Third, with the increasing use of text-based communication, and increasing questions about how chatbot service agents affect customer experience and satisfaction, this work provides real-world implications for how companies and organizations can better communicate and connect with customers via text. Taken together, these insights will both enrich the literature and inform a wide range of stakeholders.

“Hello! How May I Helo You?”: How (Corrected) Errors Humanize a Communicator

Over the past few decades, two technological inventions have fundamentally reshaped how people connect with companies and with each other: The Internet has enabled people to communicate via text on screens instead of by looking directly into each other’s eyes, and chatbots and conversational Artificial Intelligence (AI) technologies have enabled people to talk with machines instead of only with other humans. Due to these two inventions, communicating with brands (e.g., via service agents) or other consumers (e.g., via social media) has never been easier or faster. But with such convenience also comes a challenge: When individuals type on their devices, supposedly conversing with another Internet user, they can never be completely sure about the other user’s actual identity. Is that customer service agent, Twitter user, or YouTube commenter a real person or a convincing bot? This uncertainty in the humanness of one’s chat partner is engaging billions of people in an everyday version of the Turing Test, whereby people must guess whether they are conversing with a real human or a machine (Adam 2018; Turing 1950).

The blurred line between human and bot is rooted in the fact that text communication lacks the “richness” of other communication media—one cannot see or hear one’s interaction partner in text communication. This reduces people’s ability to convey information simultaneously via nonverbal cues, to provide personal focus and individuating cues, to ground conversation in the contexts of specific people, times, or places, and to experience synchrony with their interaction partners (Chafe 1982; Wheatley et al. 2012). Moreover, while voice communication contains semantic and paralinguistic cues (e.g., intonation), it is harder to rely solely on text to discern a communicator’s mental capacities and intentions (Kruger et al. 2005; Pinker and Bloom 1990; Schroeder and Epley 2015, 2016; Schroeder, Kardas, and Epley 2017). As a result, text-based communication can often seem impersonal, cold, and dehumanizing (Kiesler, Siegel, and McGuire 1984; Schroeder and Epley 2016).

This uncertainty is further aggravated by the increasing sophistication of humanlike chatbots. Well-funded by tech giants (e.g., Google, Microsoft, and Baidu) and numerous startups around the world, chatbots, automated messaging systems, and conversational agents are constantly evolving to become more "believable" interaction partners (Gottsegen 2016; Metz and Collins 2018). Such systems provide an alternative, cost-effective means of communicating with customers (Dal Porto 2017; Reddy 2017), creating a global chatbot market estimated at $430.9 million in 2020 (Grand View Research 2021). The social distancing and stay-at-home policies during the COVID-19 pandemic lent further popularity to online communication (Kemp 2020), as much of the formerly in-person communication and consultation adopted an online format and, in many cases, became partly automated (Kniffin et al. 2021). Nowadays, it is not unusual to receive physician advice from an algorithm like Watson (Lee 2014) or therapy from a chatbot like Tess (Fulmer et al. 2018).

In the face of such inevitable ambiguity, what cues do consumers rely on to determine whether they are interacting with a human or a bot? In this paper, we investigate a common yet overlooked factor in text-based communication: typographical errors, or "typos." Across five experiments using static message displays or real-time interactions across different contexts, we demonstrate how typos—and, importantly, the correction of typos—can lead consumers to perceive a service agent as more human. Further, we examine downstream consequences of this humanization process, including how it impacts consumers' perceptions of the agent's social attributes, expectations for its job performance, and even perceptions of the company that chose to use the service agent.

Perceiving Humanness

While actually *being* human is a biological fact, *perceiving* humanness is a psychological process. Prior work has identified a number of nonverbal and verbal cues that can influence humanness perception in nonhuman agents and actual humans alike, most notably appearance, voice, and identity. For instance, people tend to attribute more humanlike characteristics or mental capacities

to robots and avatars that project more humanlike facial expressions, body configurations, or movements, and those that speak with a humanlike voice (e.g., Heider and Simmel 1944; Looser and Wheatley 2010; Morewedge, Preston, and Wegner 2007; Nass and Brave 2005; Schroeder and Epley 2016; Zhao et al. 2020). They also attribute more humanness to a communicator with demographic or identity information, such as a gender, name, or nationality (Araujo 2018; Kuchenbrandt et al. 2013; Nass & Moon 2000; Waytz, Heafner, and Epley 2014). However, most of these cues are either entirely absent from text-based communications or are not unique to written language.

To date, empirical investigation on features specific to written language has been sparse. Some studies focus on how people infer a communicator's mental capacities from the semantic content of textual communication—that is, through *what* is expressed. This includes both the extent to which a communicator expresses coherent, sophisticated, and relevant thought (Lortie and Guitton 2011; Christian 2012), as well as the social acceptability of what they express (McCoy and Ullman 2018). Other studies examine how people infer a human mind by taking into account the style in which content is communicated. For instance, textual paralanguage—written manifestations of audible, tactile, and visual elements that mimic nonverbal cues in face-to-face interaction (e.g., exclamation points, emojis, handwritten-like typefaces, and vocalizations)—have been shown to humanize communicators or products (Candello, Pinhanez, and Figueiredo 2017; Luangrath, Peck, and Barger 2017; Schroll, Schnurr, and Grewal 2018).

To Err is Human? To Correct, Divine

In contrast with prior work, our research examines how typographical errors—a usually unintended but common feature in written communication—may inadvertently humanize a communicator. Conventional wisdom about human nature suggests that making mistakes is an inevitable part of being human—that is, "to err is human." And in the specific case of typing, most people can relate to the very human experience of hitting a wrong button on a keyboard and catching it

only after pressing "send." Indeed, capitalizing on this intuition, some programmers have inserted typos in an attempt to make their chatbots appear more humanlike (Christian 2011), and social psychologists have even embraced typos as a tool to help them create written materials allegedly generated by another person (Pierce et al. 2013; Wilson et al. 2005). Yet, surprisingly little is known about the psychological impact of observing typos—or errors more broadly.

Some of the most relevant findings regarding the impact of observing errors come from research on the "Pratfall Effect," which suggests that making mistakes can increase likeability (e.g., Aronson, Willerman, and Floyd 1966; Helmreich, Aronson, and LeFan 1970). In one experiment (Aronson et al. 1966), a contestant who performed exceptionally well on a difficult quiz but then later made a blunder (i.e., spilling a cup of coffee) was rated as more attractive and likable by observers than the same contestant who made no blunder. Similarly, research on the "Blemish Effect" (Ein-Gar, Shiv, and Tormala 2011) suggests that when a single piece of subtle negative information is added to an otherwise positive product description, people may evaluate the product more favorably.

By contrast, other findings within the written communication realm indicate that errors can lead to primarily negative social perceptions. For example, participants perceived a communicator who made a series of typos to be less socially attractive and less intelligent than those who wrote without typos (Westerman, Cross, and Lindmark 2018), and they rated those who produced a large volume of spelling and grammatical errors in emails or essays as less competent, less trustworthy, and having lower writing ability (Figueredo and Varnhagen 2005; Kreiner et al. 2002; Lea and Spears 1992; Vignovic and Thompson 2010). However, experiments documenting the negative impact of errors share an important commonality: the errors were presented as part of a final product and were never corrected. By contrast, studies showing positive effects of errors almost always presented errors within a stream of actions, where the actors either corrected or at least acknowledged their blunders.

In light of this critical difference, we propose that typos may humanize a communicator when they are part of an ongoing conversation and are subsequently corrected by the communicator. We believe that correcting an error is essential for making an agent appear humanlike for at least two reasons. First, correcting one's mistake signals that the communicator is actively monitoring, assessing, and reacting to each moment of the ongoing conversation. Even though the initial occurrence of an error may result from an absent or inattentive mind, correcting one's error shows that the mind is engaged and intentional. Second, correcting one's mistake also signals that one cares about how one is being perceived. People are often aware that others form negative impressions of those who make typos, and they prefer to correct their own errors rather than be corrected by others (Schegloff, Jefferson, and Sacks 1977). Therefore, communicators who correct their own errors signal that they have social awareness and a desire to fix their errors so as to avoid being perceived as mentally "lesser." By contrast, people tend to believe that machines lack the capacity for such self-awareness or conscious intent (Gray, Gray, and Wegner 2007) and do not (consciously) care about how they will be perceived by others (Boden 2018).

Drawing on the above reasoning, we hypothesize the following:

H1a-b: People will perceive an online chat agent who makes and subsequently corrects their typographical errors as more human compared to an agent who does not make errors (H1a) and compared to an agent who makes but does not correct their errors (H1b).

Consequences of Perceiving Humanness

Attributing a human(like) mind to an ambiguous communicator can lead to a multitude of downstream consequences (for reviews, see Epley and Waytz 2010; Złotowski et al. 2015). In this paper, we focus on assessments of a customer service agent's warmth (one of the primary dimensions of social evaluation; Fiske, Cuddy, and Glick, 2007), expectations of the agent's helpfulness, willingness to endorse rewards for the agent, and perceptions of the company that utilizes the agent.

Perceptions of agent warmth. Given that online communication is often experienced as cold and impersonal, and that warmth has been shown to influence consumer satisfaction (Güntürkün, Haumann, and Mikolon, 2020; Smith, Martinez, and Sabat, 2016), it is particularly important to investigate how a communicator's text messages can influence how warm they appear.[1] According to people's lay beliefs, machines lack the capacity to experience emotions, whereas humans have the full range of emotional capabilities, such as being able to feel joy, pleasure, hunger, and love (Gray et al. 2007; Waytz and Norton 2014; Weisman, Dweck, and Markman 2017). As a result, whereas machine(-like) interaction partners are often seen as cold, detached, and impersonal, human(-like) partners are perceived as conveying more qualities linked to interpersonal warmth, such as being responsive, sincere, sociable, good-natured, and likeable (Broadbent et al. 2013; Haslam 2006; Kim, Schmitt, and Thalmann 2019; Reeves, Hancock, and Liu, 2020). Hence, actions that increase a communicator' perceived humanness should increase their perceived warmth as well. Therefore, we hypothesize the following:

H2a-b: People will perceive a chat agent who makes and corrects typographical errors to be warmer than an agent who does not make errors (or makes but does not correct errors) (H2a). This effect will be mediated by perceived humanness (H2b).

Expectations of agent helpfulness. Nowadays, many companies provide online chat options to customers who want to obtain information or assistance (Adam, Wessel and Benlian 2020). However, communicating with customers over text can be challenging, because it takes a high level of social and communicative skills—skills that typically only another human would have—to understand each

[1] To be thorough in assessing consumers' interpersonal impressions, we also examine how typos influence perception of the agent's overall competence, which is the other primary dimension of social evaluation along with warmth (Fiske et al., 2007). We had no specific hypotheses regarding how corrected typos might influence perception of competence, but note that some prior research has suggested that a large number of uncorrected typos are generally associated with lack of education and mental capacity (e.g., Lea and Spears 1992). See results and further discussion on this measure in the studies.

customer's unique situations and identify effective solutions to help them. Indeed, consumers often indicate a preference for engaging with humans over chatbots, partly because they believed a human can better understand them and solve their problems (Araujo 2018; LivePerson 2019). This suggests that people perceive humans (vs. chatbots) to be more helpful in customer service contexts. Therefore, consumers' beliefs about whether they are engaging with a human are likely to affect their expectations of the communicator's ability to help them. As such, we hypothesize the following:

H3a-b: People will expect a chat agent who makes and corrects typographical errors to be more capable of understanding and solving their problem than an agent who does not make errors (H3a). This effect will be mediated by perceived humanness (H3b).

Endorsement of reward for agent. It is important for businesses and service agents to develop and maintain positive relationships with their customers. One indicator of a positive relationship in these contexts is consumers' willingness to endorse the agent's service (e.g., providing positive ratings or reviews after an interaction). Even though we are not aware of prior studies directly linking perceived humanness to endorsing the agent for reward, accumulating evidence shows that more humanlike agents can often elicit higher levels of cooperation from consumers, such as accepting an agent's offer in an online transaction (Schanke, Burtch, and Ray 2021) and behaving more prosocially toward an agent that needs help (Oliveira et al. 2021). Furthermore, given that humanness is associated with the capacity to feel emotions (Gray et al. 2007; Waytz and Norton 2014), customers may recognize that a human(-like) agent is more likely to care about and enjoy a reward for their hard work. Taken together, we hypothesized the following:

H4a-b: People will be more likely to endorse a reward to a chat agent who makes and corrects typographical errors for their service than an agent who does not make errors (H4a). This effect will be mediated by perceived humanness (H4b).

Impressions of the company. It is crucial for companies to convey care toward their customers, and service agents act as ambassadors when people form impressions about a "faceless" company. Prior research shows that consumers believe that non-human agents have limited capacities and lead to worse service experiences, and that the primary reason for companies to deploy many new technologies is to reduce costs (Larivière et al. 2017; LivePerson 2017). Given people's negative perceptions about the motives behind companies' usage of non-human agents, and their assumptions about the negative service outcomes resulting from such usage, we expect that consumers will form more favorable impressions of businesses when they perceive their service agents to be more humanlike. Specifically, we predict that:

H5a-b: People will form more favorable impressions of a company when an online service agent of the company makes and corrects typographical errors than when it does not make errors (H5a). This effect will be mediated by perceived humanness (H5b).

Overview of Studies

We test our hypotheses across five experiments. In Study 1, we presented participants with static screenshots of a greeting message from an online customer service representative that contained either no typo or a corrected typo, and tested H1-H5. This study also examined the humanizing effect of three alternative cues: appearance, name, and gender. In Study 2, we developed a chat interface and a novel interactive paradigm to create real-time, prolonged conversation experiences, to test H1-H5. In Study 3, we employed a similar chat platform but instead of programmed chatbot we used human confederates as chat agents, thus testing H1 and H2 when the conversational counterpart is—unbeknownst to the participant—an actual human. Study 4 employed the same interactive paradigm as Study 2 to test H1 and H2 but included an additional, typo-only condition in which the chat agent's typos were left unaddressed. Finally, Study 5 investigated the moderating role of agent identity by either keeping the agent's identity ambiguous or declaring it as a chatbot or a human, examining H1-

H5; here, we predicted that corrected typos would have a stronger humanizing effect when the agent's identity was ambiguous, because people would be more likely to rely on such cues under uncertainty than when the agent's identity is already known.

For all studies, we report how we determined our sample size, all data exclusions, all manipulations, and all measures. All studies were pre-registered on AsPredicted. In some studies, we introduced various secondary measures to probe other potential effects of corrected typos, and we reported all analyses for these measures in the Web Appendix. Stimuli, measures, data, and preregistration forms from all studies are available on the Open Science Framework: [https://osf.io/6s8wm/?view_only=c2e1d7bb0efb455fa97deb11a844f6d9].

Study 1: Can Typos Humanize a Communicator?

To investigate how typographical errors in online chats influence the perceived humanness of the chat agent, Study 1 presented participants with screenshots showing a series of messages from a customer service agent and manipulated whether there was a typo and correction in the agent's message. We also examined two other common features in online chats: whether an agent's profile photo portrayed a real human, and whether the agent introduced themselves by name. Introducing these additional factors allowed us to compare the effect of typo to the effects of other potentially humanizing cues. Finally, we also manipulated the agent's gender (male or female) to test whether the effect of a typo on perceived humanness might depend on this information.

Method

Participants and design. We aimed to collect 50 participants in each of the sixteen experimental conditions, thus 800 participants across conditions. A total of 929 adults (M_{age} = 32.62, SD_{age} = 12.02; 53.1% female) from Prolific started our survey in exchange for $0.64. Among them, 114 participants never completed the survey and were thus excluded, resulting in 815 participants in

the final analysis (M_{age} = 32.60, SD_{age} = 11.95; 51.8% female). This study employed a 2 (error: typo, no-typo) × 2 (profile photo: human, avatar) × 2 (name: present vs. absent) × 2 (agent gender: male, female) between-subjects design (16 separate conditions).

Procedure. Participants learned that their task was to read and evaluate online chat messages written by a service agent of a cellular company and then saw a screenshot displaying a customer service agent's greeting messages. The typo condition included a typo in one of the messages (misspelling "else" as "esle") that was corrected in a subsequent message by the same agent ("*Else**"; see Figure 1 and OSF for stimuli presentation). Messages in the no-typo condition included no typo or correction, but were otherwise identical. Along with the messages, participants saw either an avatar or a human photo (i.e., manipulating the photo) depicting either a male or a female (i.e., manipulating the gender). The agent also either introduced him/herself ("I am Angela [Michael]") or did not (i.e., manipulating the presence of a name). After viewing the agent's greeting messages, participants completed a survey evaluating the agent.

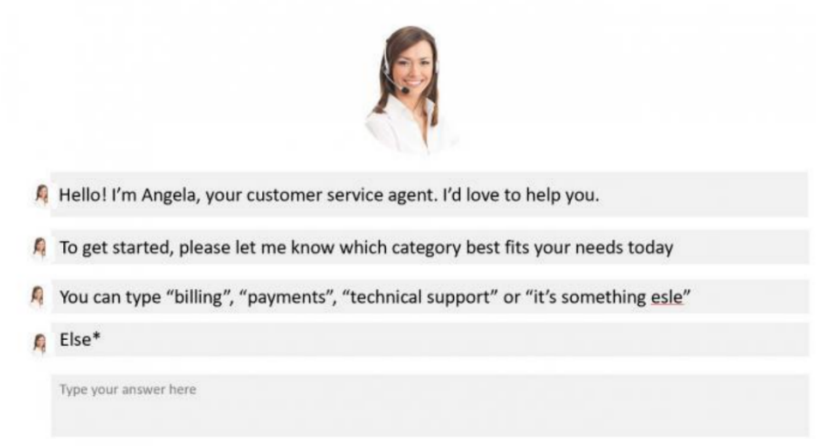

Figure 1*:* Stimuli of typo example from Study 1.

First, participants reported their *perception of the agent's humanness* on a four-item composite measure that was developed based on prior literature and administered across all studies. The first two

items focused on the uncertainty commonly experienced by online users (Adam 2018), thus providing face validity to the humanness perception construct: "To what extent did the customer service agent seem human on the previous screen?" (1 – *not at all human*; 7 – *extremely human*); "In your opinion, how likely is it that the customer service agent is a bot versus a human?" (0 – *very likely a bot; 100 – very likely a human).*[2] The next two items were developed to capture people's lay belief that robots are unable to experience emotions or bodily sensations: "If you told a joke, do you think the agent would get it?" and "Do you think this customer service agent ever gets tired when working?" (1 - *not at all likely*; 7 - *extremely likely*). These items were highly correlated (α = .89) and were combined to form our primary measure of *perceived humanness.*[3]

Participants then reported their *perceptions of the agent's warmth* (friendly/good-natured/warm/sincere; α = .92) and *competence* (confident/competent/independent/intelligent; α = .88; scale adapted from Fiske et al. 2007). Next, participants reported their e*xpectations of the agent's helpfulness* ("How likely do you think it is that this agent could help you solve your problem?" and "How likely do you think it is that this agent would be able to understand you?"; r = .87) and their *endorsement of a reward for the agent* ("To what extent do you think that this agent should be considered to receive a reward (full week of Starbucks coffee supply) from the company"). All items were rated on scales from 1 ("*not at all*") to 7 ("*very much*"). Next, participants reported their *impressions of the company* ("I hold a positive view of this cellular company," "This cellular company cares about its customers," and "This cellular company has invested a lot of effort toward helping its customers"; α = .93) on scales from 1 ("*strongly disagree"*) to 7 ("*strongly agree"*). Participants also indicated how likely they would be to share personal information with the chat agent on a scale from 1

[2] We linearly transformed ratings on this item to a range of [1, 7] in our subsequent analyses.
[3] For all statistical analyses in this and subsequent studies, when replacing the humanness composite score with the single humanness item (i.e., "To what extent did the customer service agent seem human in the previous screen?"), results remained the same. Therefore, we report analyses using the composite score throughout this manuscript.

("*not at all likely*") to 7 ("*very likely*").[4] Finally, participants completed the study by responding to *manipulation checks* regarding whether the agent had a name, what the profile picture looked like, and what typo the agent made, if any, and reported their demographic information (gender and age).

Results

Manipulation checks. Among all participants, 89.1% passed the typo manipulation check (406 out of 411 in the no-typo condition, 328 out of 412 in the typo condition), 98.9% passed the photo attention check (407 out of 411 in the no-typo condition, 407 out of 412 in the typo condition), and 77% passed the name attention check (328 out of 411 in the no-typo condition, 306 out of 412 in the typo condition). All findings remained the same when our analyses included or excluded those who failed the manipulation checks. Therefore, below we report results including all participants.

Perceived humanness. A four-way ANOVA on the composite score of perceived humanness revealed a main effect of typo, supporting H1a: Participants perceived the agent to be more human when the agent made and corrected a typo in the messages (M = 4.08, SD = 1.81) than when the agent did not make a typo (M = 2.44, SD = 1.21; $F(1, 798) = 227.33$, $p < .001$, $\eta^2 = .22$). We also found a significant, although notably smaller, main effect of profile photo ($F(1, 798) = 4.19$, $p = .041$, $\eta^2 = .005$), such that participants perceived the agent with a human photo to be slightly more human (M = 3.36, SD = 1.75) than one with an avatar photo (M = 3.18, SD = 1.74). By contrast, mentioning the agent's name only led to a marginally significant increase in perceived humanness (name: M = 3.34, SD = 1.74; no name: M = 3.20, SD = 1.74; $F(1, 798) = 2.84$, $p = .092$, $\eta^2 = .004$), and the agent's gender showed no impact on perceived humanness ($F(1, 798) = .274$, $p = .601$, $\eta^2 < .001$).

Of all possible interaction effects, only two emerged as marginally statistically significant: The interaction between typo and profile photo ($F(1, 798) = 3.04$, $p = .081$, $\eta^2 = .004$) reflected a slightly

[4] We pre-registered all measures on AsPredicted.org and describe all of them in the method sections in this manuscript. However, results related to information-disclosure (and other secondary measures in subsequent studies) are reported in the Web Appendix due to a lack of consistent main effects.

larger humanizing effect of typo when the profile photo depicted a human (typo: M = 4.29, SD = 1.72; no typo: M = 2.27, SD = 1.25) than when it depicted an avatar (typo: M = 3.88, SD = 1.88; no typo: M = 2.43, SD = 1.19), and the interaction between gender and name ($F(1, 798) = 2.93$, $p = .087$, $\eta^2 = .004$) reflected a slightly larger humanizing effect of name when the agent was a female (typo: M = 4.36, SD = 1.73; no typo: M = 2.45, SD = 1.04) than when it was male (typo: M = 4.09, SD = 1.82; no typo: M = 2.48, SD = 1.28). No other two-way, three-way, or four-way interactions reached statistical significance ($Fs < 2.13$, $ps > .144$, $\eta^2s < .001$). See Figure 2 for graphical representation of the results.

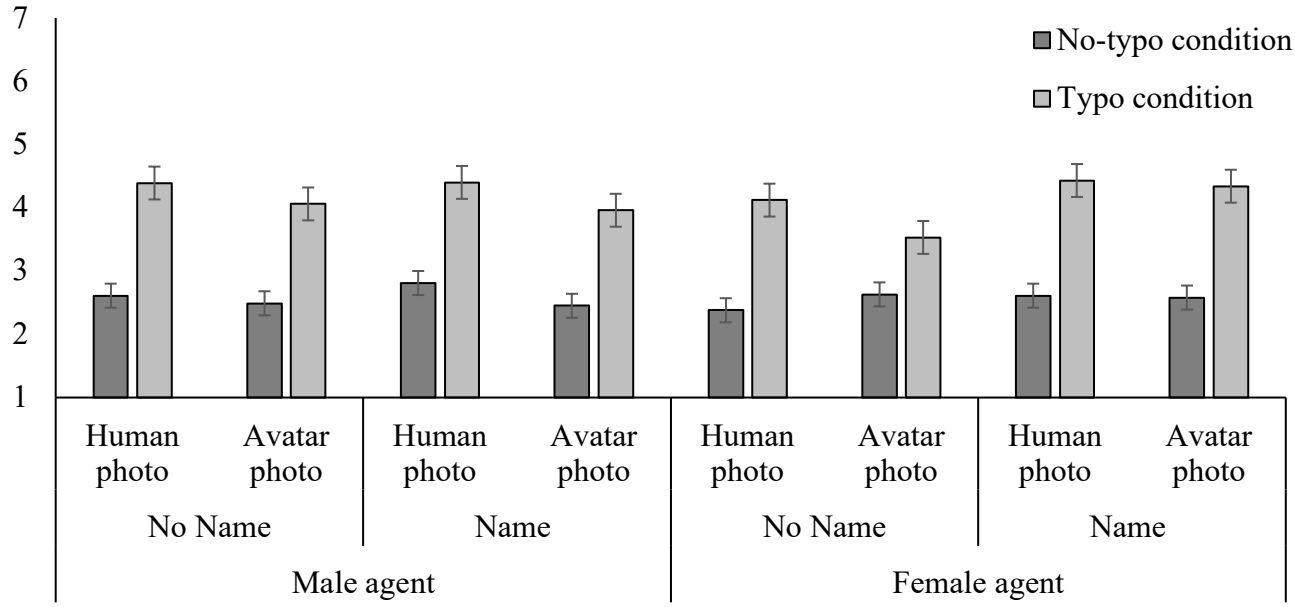

Figure 2. Perceived humanness as a function of typo, profile photo, name, and gender in Study 1. Error bars represent ±1 standard error around the means.

Perceptions of agent warmth. As to the agent's perceived warmth, supporting H2a, participants perceived the agent who made and corrected a typo to be warmer (M = 4.42, SD = 1.36) than one who did not (M = 3.85, SD = 1.54; $F(1, 798) = 33.24$, $p < .001$, $\eta^2 = .040$). Participants also perceived an agent that introduced itself by name to be slightly warmer (M = 4.24, SD = 1.48) than one that did not (M = 4.04, SD = 1.48; $F(1, 798) = 4.01$, $p = .045$, $\eta^2 = .005$). We found no main effect on warmth of

the agent's profile photo ($F(1, 798) = .481$, $p = .488$, $\eta^2 = .001$) or gender ($F(1, 798) = .419$, $p = .518$, $\eta^2 = .001$). Finally, we observed a small and unexpected three-way interaction among typo, name, and gender ($F(1, 798) = 4.51$, $p = .034$, $\eta^2 = .006$). No other interactions were significant ($Fs < 2.14$, $ps > .14$). A mediation analysis with 10,000 bootstrap samples (SPSS Macro PROCESS, Model 4) showed that, supporting H2b, perceived humanness mediated the effect of typos on warmth perceptions (b = 0.50, SE = 0.04, 95% CI = [.41, .59])

Regarding the agent's perceived competence, we found that a typo did not dampen perceptions of competence; in fact, participants perceived the agent with corrected typos in the message to be *more* competent (M = 4.30, SD = 1.27) than one with no typos in the message (M = 4.07, SD = 1.38; $F(1, 798) = 6.96$, $p = .008$, $\eta^2 = .009$). We found no main effects for other factors ($Fs < 2.16$, $ps > .14$), but we did find a small and unexpected three-way interaction among typo, name, and gender ($F(1, 798) = 6.07$, $p = .014$, $\eta^2 = .008$). No other interactions were significant ($Fs < 2.34$, $ps > .13$). Thus, making and correcting a typo led to an increase in perceptions of both agent warmth—an effect we had hypothesized and was mediated by increased perceptions of humanness—as well as competence of the agent, which we had not hypothesized.

Expectations of agent helpfulness. Supporting H3a, participants in the typo condition believed that the agent would be able to better understand them and solve their problems (M = 4.39, SD = 1.59) than did those in the no-typo condition (M = 3.43, SD = 1.65; $F(1, 798) = 69.99$, $p < .001$, $\eta^2 = .081$). We found no main effects for other factors ($Fs < 2.15$, $ps > .14$), and only a marginally significant two-way interaction between typo and name ($F(1, 798) = 3.73$, $p = .054$, $\eta^2 = .005$). Specifically, when an agent introduced themselves by name, their typo led to an even greater increase in perceived understanding and problem-solving ability (typo: M = 4.51, SD = 1.57; no typo: M = 3.33, SD = 1.71) than when they did not mention their name (typo: M = 4.28, SD = 1.60; no typo: M = 3.53, SD = 1.62). No other two-, three- or four- way interaction effects between factors were significant (Fs <

2.33, *ps* > .127, η^2s < .003). Supporting H3b, mediation analysis showed that perceived humanness mediated the relationship between the typo manipulation and participants' expectations of the agent's helpfulness (b = .59, SE = 0.04, 95% CI = [.50, .68])

Endorsement of reward for agent. Supporting H4a, participants in the typo condition were more inclined to endorse a reward for the agent (M = 3.76, SD = 1.74) than those in the no-typo condition (M = 3.43, SD = 1.79; F(1, 798) = 29.25, $p < .001$, $\eta^2 = .035$. In addition, when the agent mentioned its name, participants were also marginally more likely to endorse a reward for the agent (name: M = 3.54, SD = 1.80; no name: M = 3.31, SD = 1.77; F(1, 798) = 3.57, $p = .059$, $\eta^2 = .004$). No other main effects or interaction effects reached statistical significance (Fs < 2.68, *ps* > .10). Supporting H4b, the relationship between the typo manipulation and endorsing a reward for the agent (b = .50, SE = 0.04, 95% CI = [.42, .59]).

Impressions of the company. Supporting H5a, participants in the typo condition reported a more favorable opinion toward the company (M = 4.32, SD = 1.31) than participants in the no-typo condition (M = 3.87, SD = 1.34; F(1, 798) = 24.61, $p < .001$, $\eta^2 = .030$). We found no main effect for other factors (Fs < 1.08, *ps* > .29), and only a marginally significant two-way interaction between typo and name (F(1, 798) = 3.00, $p = .083$, $\eta^2 = .004$). Specifically, when the agent introduced themselves by name, their typo led to a greater increase in people's favorable impressions of the company (typo: M = 4.44, SD = 1.31; no typo: M = 3.84, SD = 1.34) than when the agent did not mention its name (typo: M = 4.20, SD = 1.30; no typo: M = 3.90, SD = 1.33). In addition, we found a marginally significant three-way interaction between typo, name, and gender (F(1, 798) = 3.67, $p = .056$, $\eta^2 = .005$). No other interaction effects were significant (Fs < 2.11, *ps* > .14).

Finally, supporting H5b, mediation analysis showed that perceived humanness mediated the relationship between the typo manipulation and participants' impressions of the company (b = .42, SE

= .04, 95% CI = [.34, .51])—making and correcting a typo made the communicator seem more human, which in turn led participants to view the company more favorably.

Discussion

Study 1 provides initial evidence that observing a corrected typographical error in an online chat, compared to observing no errors, can lead people to perceive the agent as more human. Further, the increased humanness perception was associated with a host of positive outcomes, such as perceiving the agent as warmer (and even, surprisingly, more competent), more helpful, and more worthy of a reward. It also led people to form a more favorable impression of the company. By contrast, the agent's profile photo appearance had a notably smaller effect on humanness perception and inconsistent effects on the downstream consequences, and neither the presence of the agent's name nor the agent's gender influenced people's perception of the agent's humanness at all.

Study 2: Real-Time Interaction with a Chatbot

In Study 2, we investigated participants' evaluations when they actually interacted with a chat agent during a live conversation. To this end, we programmed a chatbot platform and manipulated the presence of corrected typos in the agent's chat script. Modeled after actual customer service platforms, this dynamic paradigm enabled a prolonged and real-time engagement between participants and the chat agent (as opposed to a first impression of the chat agent as in Study 1), therefore ensuring the psychological realism of our method and the ecological validity of our findings.

Method

Participants and design. To ensure sufficient statistical power due to the innovative live chat platform, we aimed to recruit 200 participants in each of two experimental conditions, thus 400 participants in total. As planned, 400 participants on Prolific completed our study in exchange for $2.00 each. According to our pre-registered exclusion criteria, we excluded one participant due to suspicious spam activity, and another 14 participants who either provided nonsensible and irrelevant

responses or inadvertently asked questions that revealed the program's technological limitations (see next section for detail), resulting in 385 participants in the final analysis (M_{age} = 33.90, SD_{age} = 12.93, 53.2% female).

Chat platform. Instead of deploying marketplace AI-based chatbots, we developed our own rule-based chatbot so that we could precisely manipulate the presence or absence of typos. Learning from experts who wrote scripts for commercially successful personal assistant products (e.g., Siri and Alexa), we created a script for the chatbot where the chatbot played the role of a research assistant named "Angela" interviewing the participant with a list of personal questions. We developed the chat program using JavaScript, CSS, and HTML codes and embedded the interactive chat interface in a Qualtrics survey.[5] The program detected keywords in participants' responses and provided minimal, pre-determined responses to convey listening and understanding (see a chat example between the chat program and a participant in Figure 3). For example, when a participant's response contained a question mark, the program would interpret it as a question and respond, "*Please try your best to answer this question*." This response worked in most circumstances, yet occasionally participants used question marks for rhetorical or clarification purposes (e.g., "What is your next question?"), which triggered the automatic response and therefore exposed the non-human nature of the agent, warranting data exclusion based on our pre-registered exclusion criteria.

Besides scripting the conversation to make the chatbot appear at least somewhat "intelligent," we also implemented a few design features in the chat interface—modeled after a private chat room—that simulated a typical chat experience with a customer service agent. First, before each utterance, the chatbot would pause for a brief moment proportional to the number of words in a participant's previous message, which was meant to convey the impression that the agent spent time reading the

[5] Our OSF folder includes a demonstration video of this chatbot and a Qualtrics survey file (.qsf), to allow interested readers to download the code and create basic chatbots for their own research.

message before responding. Second, when the agent was in typing mode, the screen would display "*Angela is typing! Please wait*." above the chat box to indicate ongoing activity from the agent and thereby mimic the experience of a chat conversation with service agents. Third, the chatbot was set to be "typing" at a reasonably fast human speed (approximately 7 characters per second, spaces included), creating a realistic waiting time before each message appeared on the screen. As we developed these design features, we also conducted multiple pilot studies to ensure that participants perceived the agent to be moderately human on average.

Procedure. We informed participants that this study was part of "a nationwide project on participant well-being across behavioral labs during COVID-19," and that they would be connected to a "research agent" via a chat platform who would collect their information. Next, participants saw a loading animation and the sentence, "You will be matched with a research agent in a moment." Participants were randomly assigned to either the typo or the no-typo condition. In both conditions, they were first greeted by the chat agent, who introduced herself as Angela. In the typo condition, Angela made and corrected a typo ("helo /"*help"; see Figure 3), whereas in the no-typo condition, Angela did not make a typo. Next, Angela provided a brief task instruction and confirmed that the participant was ready to start. To strengthen the typo manipulation, we then included a second typo and correction in Angela's first question in that condition ("talking" /"*taking"; see Figure 3). No further typos were presented for the rest of the chat. After receiving a response, Angela continued with four more questions designed to elicit self-disclosure (e.g., "What's an important life lesson you learned in the past month?"). Finally, Angela concluded the chat by indicating that the time was up, thanked the participant, and instructed them to proceed to the next page.

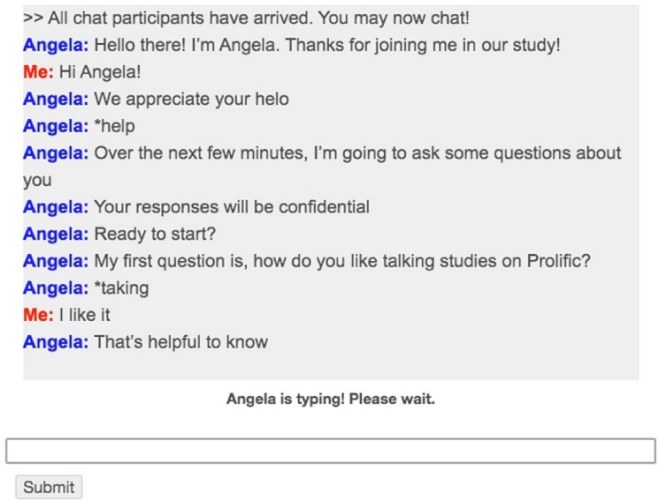

Figure 3. Screenshot of an unfolding conversation on the interactive platform in Study 2

Measures. Following the chat, participants filled out a survey measuring their evaluation of the agent and the conversation. First, participants indicated their *perceived humanness* of the agent (α = .80). Then participants responded to the same agent and company perception questions as in Study 1: *perceptions of agent warmth* (α = .90) and *competence* (α = .90), *expectations of agent helpfulness* (i.e., ability to understand/solve their problem, *r* = .81), *endorsement of reward for agent*, and *impressions of the company* (α = .89).

As secondary measures (reported in the Web Appendix), participants also completed an *information-disclosure index* indicating whether they would be willing to share certain personal information with the same chat agent in a future chat, reported their *overall satisfaction* with the conversation ("How satisfied are you with your interaction with the agent?" and "How would you rate your interaction with the agent?"; r = .73), and rated their *perceived closeness* to the agent ("To what extent did you feel connected to your agent?", "To what extent did you feel close to your agent?", and "How much do you like your agent?"; α = .91). All scales ranged from 1 (*not at all*) to 7 (*very much*)

except for rating the interaction with the agent which featured an array of seven face drawings depicting emotions from very negative (coded as "1") to very positive (coded as "7").

Finally, participants elaborated on their humanness judgment in an open-ended question, indicated what typo the agent made (if any), and answered demographic questions on their age, gender, highest education level, and race/ethnicity.

Results

Manipulation check. Among all participants, 94.3% (169 out of 191 in the typo condition, 194 out of 194 in the no-typo condition) identified their condition correctly. Results remained the same when we analyzed the data including or excluding participants who failed the manipulation checks. Hence, below we report results including all participants.

Perceived humanness. Supporting H1 and replicating the humanizing effect of typos in Study 1, participants who engaged in a live chat with a service agent perceived the agent who made and corrected their typos to be significantly more human (M = 3.99, SD = 1.72) than one who made no typos (M = 3.27, SD = 1.34; $F(1, 383) = 20.96$, $p < .001$, $\eta^2 = .052$).

Perceptions of agent warmth. As to the agent's perceived warmth, participants perceived the agent to be relatively warm on average (typo: M = 5.39, SD = 1.27; no typo: M = 5.28, SD = 1.23); the direction was consistent with our hypothesis (H2a) but did not reach statistical significance ($F(1, 383) = 0.73$, $p = .39$, $\eta^2 < .001$). Although H2a was not supported, the data revealed a significant indirect effect supporting the role of humanness perception (H2b): replicating results in Study 1, perceived humanness mediated the relationship between the typo manipulation and perceptions of agent warmth (b = 0.34, SE = .08, 95% CI = [.19, .51]). In addition, participants' perceptions of agent competence (e.g., intelligent, confident) did not differ between conditions (typo: M = 5.24, SD = 1.35; no typo: M = 5.12, SD = 1.20; $F(1, 383) = 0.92$, $p = .34$, $\eta^2 < .001$).

Expectations of agent helpfulness. Participants in both conditions believed that the agent would be somewhat able to understand and help them solve their issues (typo: M = 4.91, SD = 1.62; no typo: M = 4.70, SD = 1.46); the direction of the effect of the typo manipulation on expectations of helpfulness was consistent with our hypothesis (H3a) but did not reach statistical significance ($F(1, 383) = 1.86$, $p = .17$, $\eta^2 = .005$). Supporting H3c, perceived humanness mediated the effect of typo on expectations of the agent's helpfulness (b = .51, SE = .11, 95% CI = [.28, .74]),

Endorsement of reward for agent. Supporting H4a, participants who saw the agent make and correct typos were significantly more inclined to endorse a reward for the agent (M = 5.34, SD = 1.81) than those who saw an agent that made no typos (M = 4.76, SD = 1.84; $F(1, 383) = 9.63$, $p = .002$, $\eta^2 = .035$). Replicating results in Study 1 and supporting H4b, perceived humanness mediated the effect of typo on reward endorsement (b = .51, SE = .12, 95% CI = [.28, .75]).

Impressions of the company. Supporting H5a, participants formed a marginally more favorable opinion of the company when the chat agent made a typo (M = 5.53, SD = 1.20) than when the agent did not (M = 5.32, SD = 1.16; $F(1, 383) = 3.07$, $p = .080$, $\eta^2 = .008$). Supporting H5b, mediation analysis further showed that humanness perception mediated this relationship (b = .25, SE = .06, 95% CI = [.13, .39]).

Discussion

Study 2 employed an interactive paradigm where participants engaged in a live interaction with a pre-programmed chatbot. This chatbot was programmed to inquire about participants' recent life experiences, similar to conversations between users and personal assistants, lending ecological validity to our paradigm. Replicating Study 1, we observed that making and correcting typos led the chat agent to be perceived as more human (H1a). Furthermore, participants were more inclined to endorse a reward for the typo-correcting chat agent (H4b) and saw the company in a somewhat more positive light (H5) than a chat agent that did not make a typo. The effect of the typo manipulation on

participants' perceptions of the agent's warmth (H2a) and expectations about the agent's helpfulness (H3a) did not reach statistical significance. We suspect that the attenuated effect sizes might have resulted from the complexity of the conversation topics (e.g., discussing life lessons amid pandemic times), which often revealed the rigidity of the chat agent and introduced more variance in people's conversation experiences. Importantly, however, assessments of the agent's humanness still mediated all the above measures (supporting H1b, H2b, H3b, and H4b and H5b).

Study 3: Real-Time Interaction with a Human

Some customer service agents are powered by chatbots, yet others are actual human workers. In Study 3, we sought to replicate the humanizing effect of typos observed in Studies 1 and 2 when customers actually interacted with another human. To this end, we hired a team of human research assistants who interacted with participants via a private online chatroom. Moreover, in order to ensure more predictability in conversation flow and thus more experimental control over the chat experiences, the chat agent in this study did not ask open-ended questions about participants' lives, but instead asked questions that focused on eliciting factual personal information about the participants—such as annual income—and that participants could skip if they did not wish to respond.

Method

Participants and design. We employed a between-subjects design with two experimental conditions—typo and no-typo—and aimed to recruit 100 participants in each condition.[6] In total, 200 participants were recruited from either a university campus or from a metropolitan area with a more diverse population.[7] Participants completed this experiment in exchange for $4.00.

[6] According to a power analysis, this sample size was sufficient to detect a small-to-medium effect size of $d = .40$ given power of .80 and an alpha level of .05. We chose this arbitrary, relatively large sample size instead of an even larger one due to the sheer amount of resources required to coordinate a live interaction paradigm that involved recruiting participants to a laboratory to interact with live human confederates.

[7] Additional analyses confirmed that recruitment location did not interact with any of our key dependent variables.

Following our predetermined exclusion criteria, we excluded a total of 10 participants because they experienced prolonged disconnection from the chat due to Internet disruption or experimenter error (N = 4), lacked sufficient English proficiency to have an online chat (N = 2), provided inattentive responses (N = 1), or because the human chat agent made unintended errors other than the prescribed typos (N = 3). This resulted in a final sample of 190 participants (campus-based sample: N = 74, M_{age} = 24.5, SD_{age} = 7.7, 52.7% female; downtown city-based sample: N = 126, M_{age} = 42.7, SD_{age} = 14.5, 37.3% female).

Human confederates. We hired six experienced research assistants as human chat agents and provided rigorous standardized training based on a predetermined script to ensure identical conversations across all participants (see the OSF file for the confederate protocol). The script was largely similar to that in Study 2 and was also pretested to eliminate potential ambiguity in the agent's language and to keep participants on the same course of conversation. We trained the human confederates to time their responses to convey the impression of reading and typing, similar to the way we programmed the chatbot responses in Study 2.

Conversation interface. To enable real-time conversations with human chat agents, we employed the ChatPlat platform, a web application that resembles an actual online chatroom and allows multiple users to send and receive messages in a private chatroom in real time. This application has been used in previous research to study human social interactions (e.g., Huang et al 2017; Logg et al 2019). Similar to our chatbot interface in Study 2 (see Figure 3), ChatPlat also displays a prompt indicating "someone else is typing" when it detects activity in the other user's entry box.

Procedure. We informed participants that this study was part of "a nationwide project to compare research practices and participant populations across behavioral labs," and that they would be interacting with a "research agent" who could be either a human or a chatbot. Next, we randomly assigned participants to either the typo or the no-typo condition before matching them with an agent to

begin the chat. Participants in both conditions received greeting messages identical to those in Study 2. Before asking any questions, the agent also informed participants that they were allowed to skip questions by pressing the "S" key on the keyboard or any statement that indicated a desire to skip (e.g., "I prefer not to answer" or "next question"). If this happened, the agent was instructed to ask their next question. The number of questions that participants chose to skip was used to calculate a measure of *information-sharing behavior* (which was examined as another potential consequence of perceived humanness; see overview in the General Discussion and analyses in the Web Appendix).

During the chat, the chat agent made two typos identical to those in Study 2 (i.e., "helo"/ "*help", "talking"/ "*taking"; see Figure 3). The agent asked a total of eight personal questions that were intended to elicit factual personal information from the participants (e.g., "What is the most sensitive issue that you ever encountered at the [lab name]?", "What is your primary source of income?", and "What religion do you believe in, if any?"), and provided brief responses to indicate listening as in Study 2. When a participant asked a clarifying question, the agent provided either a brief answer (if the question was a simple yes/no question) or a canned response like "*Thanks for asking, but I can't comment on that.*"

Measures. Participants responded to the *perceived humanness* composite measure ($\alpha = .85$), followed by *perceptions of agent warmth* ($\alpha = .95$) and *competence* ($\alpha = .90$) as in previous studies. Note that this study did not measure expectations of agent helpfulness, endorsement of reward for the agent, and impressions of the company because it was conducted prior to Studies 1 and 2 (before we started testing the hypothesized consequences featured in H3 to H5).

As secondary measures (analyzed in the Web Appendix), participants indicated their overall *enjoyment* of the conversation and their *interpersonal closeness* to the agent, and responded to an extended version of the information-sharing intention index used in Study 2 (18 items; $\alpha = .86$). We also included part of the "Internet Users' Information Privacy Concerns" survey (Malhotra et al. 2004)

to explore whether individual differences in self-reported privacy concern would moderate any effects on information disclosure. Finally, participants reported their gender, age, and any comments they had about the chat experience. Then, we thanked, debriefed, and compensated participants.

Results

Manipulation check. Among all participants, 86.32% passed the manipulation check (92 out of 93 in the no-typo condition, 72 out of 97 in the corrected-typo condition). Results remained the same when we analyzed the data including or excluding participants who failed the manipulation checks. Therefore, below we report results including all participants.

Perceived humanness. Supporting H1a, a chat agent who made and corrected typos was perceived to be more human (M = 4.31, SD = 1.51) than an agent who did not make any typos (M = 3.58, SD = 1.58; F(1, 188) = 10.48, p = .001, η^2 = .053).

Perceptions of agent warmth. Supporting H2a, participants in the typo condition perceived the chat agent to be warmer (M = 5.07, SD = 1.59) than those in the no-typo condition (M = 4.52, SD = 1.73; F(1, 188) = 5.10, p = .025, η^2 = .027). Mediation analysis showed that, similar to previous studies, perceived humanness mediated the effect of the typo manipulation on perceived warmth (supporting H2b; b = .41, SE = .14; 95% CI = [.16, .70]). Finally, participants in the typo and the no-typo conditions perceived the agent to be similarly competent (Ms = 5.10 and 4.87, SDs = 1.50 and 1.49, respectively; F(1, 188) = 1.10, p = .30, η^2 = .006).

Discussion

Using a live interaction paradigm with an actual human agent, we again found that correcting a typographical error, compared to making no errors, humanized a chat agent (H1a). Moreover, we again observed that people perceived the agent to be warmer when they made and corrected their typos, and these effects were mediated by perceived humanness of the agent (H2a and H2b). We observed a non-significant effect of typo manipulation on perceived competence. Overall, by

implementing two realistic real-time conversation paradigms, similar to what consumers regularly experience when chatting with online service agents, Studies 2 and 3 provide consistent evidence that correcting typos can lead people to perceive an ambiguous communicator—who could be either a chatbot or a human employee—as more human.

Study 4: Corrected vs. Uncorrected Typos

To investigate whether it is correcting a typo, rather than just making a typo, that humanizes the chat agent (H1b), Study 4 included three experimental conditions: one in which the chat agent made and corrected typos, another in which the agent just made the typos without correcting them, and the third in which the agent made no typo at all. We again used the chat platform we programmed in Study 2 to create a real-time conversation experience, except that in this case the chat agent asked personal information questions instead of open-ended life experience questions to make the conversation flow more predictable and reduce variance in participants' experiences.

Method

Participants and design. This experiment employed a between-subjects design with three conditions: corrected-typo, uncorrected-typo and no-typo. We aimed to recruit 200 participants in each of the three experimental conditions, thus 600 participants in total. A total of 603 participants were recruited and completed our study in exchange for $2.00 each. According to our pre-registered exclusion criteria, we excluded participants whose responses to the open-ended questions were unintelligible or lacked minimal effort (N = 24), whose responses to the exploratory 18-item information-disclosure index seemed improbable (N = 4), whose chat responses exposed the mechanics of the chatbot (N = 51; 44 participants used a question mark that triggered an out-of-context automatic response from the agent and 7 participants attempted to ask the agent multiple questions during the greeting phrase that the agent could not answer), and who experienced technical

errors during the chat (N = 6), resulting in a final sample of 518 participants for final analysis (M_{age} = 36.94, SD_{age} = 11.94, 42.5% female).

Procedure. Prior to the study, participants learned that we were "constructing a new participant database to study consumer purchases" and were randomly assigned to the no-typo condition, the corrected-typo condition, or the uncorrected-typo condition prior to the conversation. After entering a virtual chatroom identical to that in Study 2, participants were greeted by a chat agent named Angela. In both the corrected-typo and the uncorrected-typo conditions, the agent made two typos during the conversation identical to those in Studies 2 and 3 (i.e., "*helo*" / "**help*", "*talking*"/ "**taking*"; see Figure 3). However, in the uncorrected-typo condition, the agent continued the conversation without addressing the typos. During the course of the conversation, the agent asked a total of ten personal questions (e.g., "What is your date of birth?", "What is your zip code?", "How much money do you make each month?", and "What are the last four digits of your phone number?"), and were again allowed to skip questions by simply pressing a skip button on their interface.

Measures. Participants reported the agent's *perceived humanness* (α = .92) and their *perceptions of the agent's warmth* (α = .93) and *competence* (α = .88) using the same items as in previous studies. As secondary measures (analyzed in the Web Appendix), participants also completed the 18-item *information-sharing intention index* and rated their overall experience in the chat (r = .84) and their *interpersonal closeness* to the agent (α = .91) on the same items as in Study 3 (reported in the Web Appendix). And once again, participants elaborated on their humanness perceptions in an open-ended question and indicated what typo the agent made, if any, in a manipulation check question. Like Study 3, this study did not test H3–H5 because it predated the generation of those hypotheses.

Results

Manipulation check. Among all participants, 17.5% (89 participants) failed the manipulation check (66 out of 173 in the uncorrected-typo condition, 23 out of 156 in the corrected-typo condition,

and 1 out of 180 in the no-typo condition). Results remained the same when we analyzed the data including or excluding participants who failed the manipulation checks. Therefore, below we report results including all participants.

Perceived humanness. A one-way ANOVA revealed a significant effect of the typo manipulation on perceived humanness of the chat agent ($F(2, 515) = 6.20$, $p < .001$, $\eta^2 = .024$). Specifically, when the chat agent corrected the typos, participants perceived the agent to be significantly more human ($M = 4.07$, $SD = 1.76$) compared to when the agent did not make a typo ($M = 3.45$, $SD = 1.67$; $t(340) = 3.32$, $p = .001$, $d = 0.36$) and when the typos were made but not corrected ($M = 3.50$, $SD = 1.85$; $t(331) = 2.83$, $p = .005$, $d = 0.31$). In fact, when the agent did not make any typos or did not correct the typos it did make, people perceived similar levels of humanness in the agent ($t(359) = .300$, $p = .764$, $d = 0.02$). Supporting H1a and H1b, it was the correction of a typo, rather than the mere presence of a typo, that led participants to perceive a chat agent as more human.

Perceptions of agent warmth. A one-way ANOVA on the agent's perceived warmth revealed a significant difference across three conditions ($F(2, 515) = .3.96$, $p = .020$, $\eta^2 = .015$). Supporting H2a, participants in the corrected-typo condition perceived the agent to be warmer ($M = 5.12$, $SD = 1.29$) than those in the no-typo condition ($M = 4.68$, $SD = 1.50$; $t(340) = 2.85$, $p = .005$, $d = 0.31$), whereas the difference between the uncorrected-typo ($M = 4.90$, $SD = 1.48$) and the no-typo conditions was nonsignificant ($t(359) = -1.39$, $p = .16$, $d = 0.14$). Supporting H2b, mediation analysis showed that perceived humanness of the agent mediated the difference in warmth perception between the corrected-typo condition and the no-typo condition ($b = .22$, $SE = .08$; 95% CI = [.06, .38]).

A one-way ANOVA on the agent's perceived competence showed no significant effect of the typo manipulation across conditions ($F(2, 515) = .709$, $p = .493$, $\eta^2 = .003$), suggesting that people's perception of the chat agent's overall competence was neither dampened nor enhanced by the presence of typos.

Discussion

Using a live chat paradigm, Study 4 supports H1a and H1b and suggests that it is correcting a typo, rather than merely making a typo, that leads to greater humanness perceptions. In other words, the presence of errors *per se* does not necessarily lead to the perception of a more human mind, but the act of correcting errors conveys a mindful agent. Bots, like humans, frequently make errors in live chats, Tweets, or spam emails, but our results show that the subsequent *correction* of those errors is the key to conveying humanness and human qualities such as interpersonal warmth.

Study 5: Ambiguous vs. Known Chat Agent Identity

Across earlier studies, the chat agent's identity has been kept ambiguous and was left up to participants' inferences. What if, however, the agent's identity is already known prior to the chat? In Study 5, we directly manipulated the identity of the agent (ambiguous, chatbot, or human) to examine whether the humanizing effect of corrected typos would persist even when the identity was unambiguous, and whether the magnitude of the humanizing effect would be moderated by knowledge of the agent's identity. We predicted that corrected typos would have a stronger humanizing effect when the agent's identity was ambiguous because people would be more sensitive to cues that can potentially reveal an agent's identity when that identity isn't already known. However, in light of people's tendency to anthropomorphize objects that merely resemble human appearances or actions (e.g., Aggarwal and McGill 2007; Zhao et al. 2020), we also predicted that correcting typos would still humanize a chat agent that is known to be a chatbot to some extent. Finally, given the dehumanizing nature of written communication for human conversations, we were also curious to examine whether correcting typos could even further humanize a human agent. Moreover, comparing responses when participants knew the agent was a human versus a chatbot can provide direct, causal evidence for the impact of perceived humanness on downstream consequences.

Method

Pilot study. Because this study was the first to compare conditions under which the customer service agent's identity was known or ambiguous, we decided to collect pilot data to ensure sufficient statistical power for the hypothesized identity × typo interaction effect. A pilot study with a 2 (agent identity: ambiguous, chatbot) × 2 (typo manipulation: corrected-typo, no typo) between-subjects design revealed an interaction effect size of f = .20 (see Web Appendix for detailed method and results). This suggested that to detect the same interaction effect at a power of .80 and an alpha level of .05, and accounting for an exclusion rate of 20%, we would need approximately 65 participants per condition in the main study.

Participants and design. Based on this pilot study, we pre-registered the above 2 × 2 factorial design as our primary conditions in the main study as well as a third agent identity condition (i.e., the human condition) as exploratory, and aimed to recruit 390 participants across all six conditions (260 for the key 2 × 2 design, as described above). A total of 386 participants were recruited via Prolific and completed our main study in exchange for $0.50 each ($M_{age}$ = 34.47, SD_{age} = 12.22, 54.9% female). Another 11 participants started the study but never completed it and were thus excluded.

Chat interface. Our primary goal was to examine whether prior beliefs about the chat agent's identity would moderate the humanizing effect of corrected typos. To ensure that the conditions were identical, we created a dynamic version of the observation paradigm employed in Study 1 showing an online chat agent sending messages over time. We employed the observer paradigm—instead of a live chat—to keep the conversation identical and to only manipulate participants' knowledge of the agent's identity across conditions; otherwise, people would likely speak differently to a chatbot versus a human. When the agent was supposedly typing a message, an animated bubble with three dots appeared below the most recent message to indicate ongoing activity by the agent.

Procedure. Prior to the study, participants were randomly assigned to one of the three agent identity conditions (i.e., chatbot, ambiguous, and human). Depending on their condition assignment,

participants were either told that a cellular company was training “their new AI-powered chatbot that they call Angela” (chatbot condition) or “their new employee, named Angela” (human condition) to “use an interactive chat platform to answer customer service questions,” or that that the cellular company was “developing an interactive chat platform to answer customer service questions” (ambiguous condition). Participants in all conditions were then asked to watch and evaluate an online chat conducted on this platform. On the next page, participants saw the chat agent send greeting messages one by one. While all words were spelled correctly in the no-typo condition, the corrected-typo condition included a typo in one message (“esle”) and another message to correct it (“*else”). The forward button on this page was disabled for 25 seconds in all conditions, to ensure that participants spent at least that amount of time viewing the stimuli.

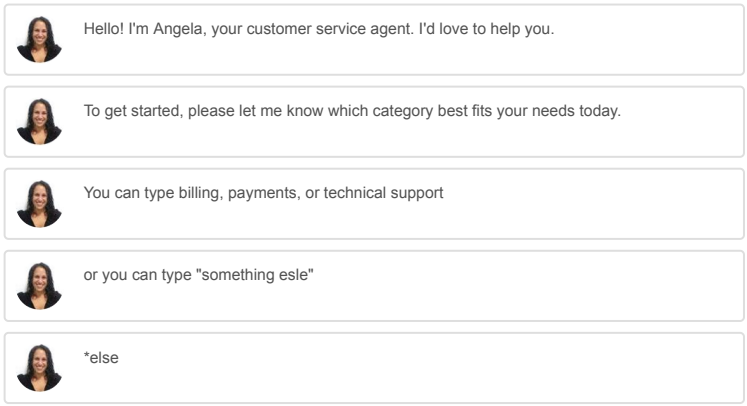

Figure 4. A screenshot of the chat interface in Study 5 (typo condition) with all messages displayed.

Measures. After seeing the chat page, participants responded to three *perceived humanness* questions identical to those in previous studies ($\alpha = .79$; we removed the item, “How likely is it that the customer service agent is a bot versus a human?” because the agent’s identity was already stated in some conditions). We counterbalanced whether perceived humanness was measured before or after other measures to ensure that the mediation results were robust to measurement order.

Next, participants reported to the same items as in Study 1 on their *perceptions of the agent's warmth* (α = .93) and *competence* (α = .89), *expectations of the agent's helpfulness* (r = .85)*, endorsement of a reward for the agent*, and their *impressions of the company* (including a new item on business ethics, "I expect this cellular company to treat its customers fairly"; α = .94). As secondary measures (reported in the Web Appendix), participants indicated how error-prone and careless the agent seemed to be (r = .76), to assess an additional potential downside of making typos, and they also reported their *information-disclosure intention* using the single-item measure in Study 1 and the six-item index in Study 2 (α = .69). Finally, participants completed manipulation checks on the agent's identity ("a human employee," "an AI-powered chatbot," or "it wasn't mentioned") and what typo the agent made (if any), before reporting their gender, age, and comments about the study.

Results

Manipulation checks. For the typo manipulation check, 191 out of 194 participants in the no-typo condition and 179 out of 192 in the corrected-typo condition correctly identified their condition (95.8% on average), and results remained the same when we included or excluded those who failed this manipulation check. For the identity manipulation check, all participants in the chatbot condition (127 out of 127) and the majority in the human condition (80.6%; 104 out of 129) correctly reported the identity of the agent; by contrast, only 66.2% in the ambiguous condition (86 out of 130) correctly indicated that the agent's identity was not mentioned, whereas others believed it was an AI-powered chatbot (18.5%; 24 out of 130) or a human employee (15.4%; 20 out of 130). Including or excluding those who failed the identity manipulation check led to largely similar results; below, we report results including all participants regardless of the manipulation checks, and we also included results excluding participants who failed the agent identity manipulation in the Web Appendix.

Perceived humanness. Supporting H1a, a two-way ANOVA on perceived humanness revealed a significant main effect of the corrected-typo manipulation ($F(1, 380) = 102.11$, $p < .001$, $\eta^2_p = .21$),

such that participants perceived the agent to be more human when it made a corrected typo (M = 4.49, SD = 1.56) than when it did not (M = 3.15, SD = 1.43). It also revealed a significant main effect of agent identity (F(2, 380) = 53.06, $p < .001$, $\eta^2_p = .22$) such that participants perceived a known human agent to be more humanlike (M = 4.57, SD = 1.57) than an ambiguous agent (M = 3.97, SD = 1.63; t(257) = 3.01, $p = .001$, d = .37), which was more human than a known chatbot agent (M = 2.90, SD = 1.24; t(255) = 5.86, $p < .001$, d = .73). Finally, there was a significant interaction effect (F(2, 380) = 6.01, $p = .003$, $\eta^2_p = .031$; see Figure 5). Examining the effect of the typo manipulation on each agent identity condition separately, we found that the humanizing effect of a corrected typo was larger when the agent's identity was ambiguous (corrected-typo vs. no typo: Ms = 4.97 vs. 3.14, SDs = 1.34, 1.38; t(128) = 7.62, $p < .001$, d = 1.34) or was known to be a human (Ms = 5.30 vs. 3.78, SDs = 1.31, 1.45; t(127) = 6.25, $p < .001$, d = 1.10) than when it was known to be a chatbot (Ms = 3.25 vs. 2.53, SDs = 1.18, 1.22; t(125) = 3.40, $p < .001$, d = 0.60).

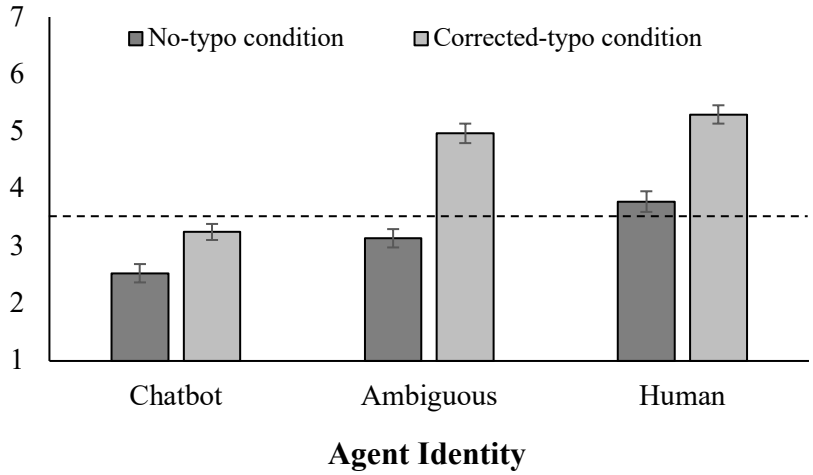

Figure 5. Results on humanness perception in Study 5. The dotted horizontal line represents the neutral rating of 4. Error bars represent ±1 standard error around the means.

Perceptions of agent warmth. A two-way ANOVA revealed a significant main effect of the typo manipulation (F(1, 380) = 8.96, $p = .003$, $\eta^2_p = .023$), such that participants perceived the agent to

be warmer when it made and corrected a typo (M = 5.01, SD = 1.26) than when it did not (M = 4.61, SD = 1.35). It also revealed a significant main effect of agent identity ($F(2, 380) = 3.27$, $p = .039$, $\eta^2_p = .017$), whereby participants perceived the agent to be warmer when the agent was a human (M = 5.05, SD = 1.24) compared to when its identity was ambiguous (M = 4.73, SD = 1.35; $t(257) = 2.98$, $p = .024$, d = .24) or a bot (M = 4.64, SD = 1.34; $t(254) = 2.55$, $p = .006$, d = .31). There was a non-significant interaction ($F(2, 380) = 1.48$, $p = .23$, $\eta^2_p = .008$. Examining the effect of the typo manipulation separately for each agent identity condition revealed a significant effect when the agent was ambiguous (Ms = 5.01 vs. 4.51, SDs = 1.35, 1.31; $t(128) = 2.14$, $p = .034$, d = .38) or was known to be human (Ms = 5.35 vs. 4.74, SDs = 1.06, 1.35; $t(127) = 2.86$, $p = .005$, d = .50), but not when it was known to be a chatbot (Ms = 4.68 vs. 4.60, SDs = 1.30, 1.41; $t(125) = .33$, $p = .37$, d = .05). Finally, supporting H2b, perceived humanness of the agent mediated the difference between the no-typo condition and the corrected-typo condition (b = .58, SE = .08; 95% CI = [.43, .75]).

By contrast, results on perceived competence showed no significant main effects of either the typo manipulation or agent identity, nor an interaction effect ($Fs < 1.08$, $p > .34$), again suggesting that correcting the typo did not harm competence perceptions. The effect of the typo manipulation on perceived competence was not significant for any of the three agent identity conditions when tested in separate pairwise comparison tests (ps > .33).

Expectations of agent helpfulness. A two-way ANOVA revealed a main effect of the typo manipulation ($F(1, 380) = 11.83$, $p < .001$, $\eta^2_p = .030$), such that participants perceived the agent to be more helpful when it made a corrected typo (M = 4.59, SD = 1.49) than when it did not (M = 4.10, SD = 1.60). It also revealed a main effect of agent identity ($F(2, 380) = 25.15$, $p < .001$, $\eta^2_p = .12$), such that participants believed that a human employee (M = 4.85, SD = 1.43) would be marginally more helpful than an ambiguous agent (M = 4.54, SD = 1.43; $t(257) = 1.75$, $p = .08$, d = .22), and an ambiguous agent would be more helpful (M = 4.54, SD = 1.43) than a chatbot (M = 3.61, SD = 1.58;

t(255) = 4.95, $p < .001$, d = .62). There was no interaction effect (F(2, 380) = .528, $p = .59$, $\eta^2_p = .003$). Examining the effect of the typo manipulation separately for each agent identity condition, the effect was significant when the agent was ambiguous (Ms = 4.90 vs. 4.25, SDs = 1.32, 1.45; t(128) = 2.65, $p = .004$, d = .47), and when it was a human employee (Ms = 5.13 vs. 4.55, SDs = 1.19, 1.60; t(127) = 2.36, $p = .010$, d = .41), but not when it was known to be a chatbot (Ms = 3.76 vs. 3.46, SDs = 1.56, 1.60; t(125) = 1.07, $p = .144$, d = .19). Finally, supporting H3b, perceived humanness of the agent mediated the difference between the no-typo condition and the corrected-typo condition (b = .84, SE = .11; 95% CI = [.63, 1.06]).

Endorsement of reward for agent. A two-way ANOVA revealed a significant main effect of agent identity (F(2, 380) = 38.10, $p < .001$, $\eta^2_p = .17$), such that participants were more inclined to endorse a reward for a human employee (M = 4.44, SD = 1.58) than for an ambiguous agent (M = 3.91, SD = 1.61; t(257) = 2.66, $p = .008$, d = .33), and for an ambiguous agent than for a chatbot (M = 2.73, SD = 1.64; t(255) = 5.83, $p < .001$, d = .73). There was no main effect of the typo manipulation nor an interaction effect (Fs < 1.84, $p > .16$). Examining the effect of the typo manipulation separately for each agent identity condition, participants were directionally more inclined to endorse the typo-correcting agent than the no-typo agent when the agent's identity was ambiguous (supporting H4; Ms = 4.15 vs. 3.72, SD = 1.52 vs. 1.67; t(128) = 1.54, $p = .13$, d = .27), whereas this pattern was completely absent when the agent was known to be a chatbot or a human (ps > .64). Finally, supporting H4b, perceived humanness of the agent significantly mediated the relationship between the typo manipulation and participants' endorsement for the agent (b = .73, SE = .11; 95% CI = [.53, .95]).

Impressions of the company. A two-way ANOVA showed a significant main effect of agent identity (F(2, 380) = 8.19, $p < .001$, $\eta^2_p = .041$), such that participants perceived a company using an ambiguous chat agent (M = 4.60, SD = 1.16) similarly positively to one with a human employee (M = 4.69, SD = 1.31; t(257) = .56, $p = .58$, d = .07) and more positively than one using a chatbot (M =

4.09, SD = 1.35; t(211) = 3.27, p = .001, d = .41). There was no main effect of the typo manipulation nor an interaction effect (Fs < .55, ps > .49). Examining the effect of the typo manipulation separately for each agent identity condition showed that the effect of the typo manipulation did not achieve statistical significance for any identity condition (ps > .35), including the key condition in which the agent was ambiguous (Ms = 4.69 vs. 4.53, SD = 1.30 vs. 1.04, t(128) = 0.75, p = .45, d = .13). Nonetheless, supporting H5b, perceived humanness showed a robust mediation effect on participants' impressions of the company, such that participants formed more favorable impressions of the company when they perceived the agent to be more human (b = .64, SE = .08; 95% CI = [.47, .81]).

Discussion

Study 5 provides at least four key findings regarding the humanizing effect of typos. First, consistent with our prediction, knowledge about a chat agent's identity moderated the humanizing effect of typos. In particular, the humanizing effect of the typo was larger when the agent's identity was ambiguous than when the agent was known to be a chatbot. Second, correcting a typo still humanized the communicator even when people knew that the chat agent was a chatbot or a human. Third, seeing an ambiguous agent or a human (but not a chatbot) make and correct a typo led people to perceive the agent to be warmer and expect the agent to be more helpful than seeing an agent that does not make a typo. Fourth, by directly manipulating the identity of the agent, Study 5 confirmed that people indeed perceived a human agent to be warmer and more helpful than a chatbot agent, were more inclined to endorse a reward to the agent, and perceived the company more favorably, thus providing direct, causal evidence of the impact of humanness on our downstream consequences.

Study 5 also yielded some unexpected results. Contrary to Studies 1 and 2, there was no impact of corrected typos on consumers' endorsement of a reward for the agent nor impressions of the company the agent represented—even in the ambiguous agent condition. However, the *agent's identity* directly influenced the endorsement of a reward and company perceptions: Participants who observed

a human agent or an ambiguous agent were more likely to endorse a reward and perceived the company more favorably than those who observed a chatbot agent. Furthermore, consistent with recent methodological advances suggesting that mediation effects may exist despite a non-significant association between the independent variable and the dependent variable (e.g., Memon et al. 2018; Zhao, Lynch, and Chen, 2010), we found that perceived humanness was a significant mediator for those consequences. These results suggest that endorsement of rewards and company impressions may be influenced by multiple determinants, and future research should continue to investigate how a chat agent's identity and behavior in a conversation can influence consumers' attitude toward the agent and company.

General Discussion

Due to the lack of diagnostic cues such as voice, appearance, and social identity in text-based online communications, consumers must rely on other aspects of the chat to ascertain the humanness of a service agent who may be either a human or a chatbot. How do consumers discern the identity of an ambiguous chat agent? And what consequences does that have on consumers' perceptions and behaviors? In this paper, we explored a common yet rarely researched phenomenon: making and correcting typological errors in a text-based chat. Across five experiments using a variety of stimuli and paradigms—from static or dynamic message displays (Studies 1 and 5) to live chats with a chatbot (Studies 2 and 4) or a human agent (Study 3)—our research consistently demonstrates that making and correcting one's typographical errors chat can humanize a customer service agent. Crucially, we found that it is correcting of one's typo, rather than merely making a typo, that humanizes the agent (Study 4). Moreover, the humanizing effect of typos persists even when the agent is known to be a chatbot, but it creates a greater impact when the agent's identity is kept ambiguous or when the agent is known to be a human (Study 5).

Even though errors are often perceived negatively, our research further shows that making and correcting typos can elicit favorable impressions of the agent, such as greater interpersonal warmth and higher expectations of its ability to understand and help customers solve complex issues. These positive perceptions are driven by increases in perceived humanness—consumers attribute desirable human traits and characteristics to a chat agent that appears more human. Finally, consumers even form more positive impressions of the company as a result of the perceived humanness of the chat agent.

Theoretical Contributions

Taken together, our findings contribute to prior research on the antecedents, processes, and consequences of anthropomorphism and mind perception. Different from the majority of prior work that focuses on how nonverbal cues such as face and voice can humanize a communicator, our research examines how people perceive the humanness of chatbots based on dynamic features communicated over text. Relatedly, our work also highlights the close relationship between the fundamental dimensions of interpersonal perceptions (warmth and competence) and perception of humanness. In particular, the fact that humanness is more closely related to perceptions of warmth (e.g., friendly, sincere) than perceptions of competence (e.g., intelligent, confident) extends previous research on the primacy of warmth over competence in driving social evaluations (Fiske et al. 2007) and consumer satisfaction (Güntürkün et al. 2020).

Moreover, our research expands the empirical literature on the interpersonal consequences of making mistakes. Specifically, our research helps to reconcile an intriguing discrepancy between research suggesting that small errors can increase the perceived attractiveness and likeability of people and products (e.g., Aronson et al. 1966; Ein-Gar et al. 2011) and a separate literature showing negative effects of errors (e.g., Figueredo and Varnhagen 2005). We show that acknowledging and correcting one's error, as opposed to simply making an error, can increase perceived humanness. We theorize

that this is because addressing one's error signals an engaged and caring mind—it requires some degree of intentionality (e.g., desire to avoid being seen as careless) and meta-cognitive capacities (e.g., self-awareness), which are generally considered to be lacking in machines, algorithms, and other artificial intelligence. In addition, the fact that perceived competence did not show consistent results across studies suggests that errors do not always lower competence perceptions, at least when they are less fundamental to the product or service *per se* (Mende et al. 2019).

Practical Implications

By examining the inner working of the human mind in online chats, our research provides important practical implications to companies, service representatives, policy-makers, and consumers. For companies that deploy online chat in customer service, our research suggests that humanizing chat agents can create a variety of benefits, even boosting consumers' general impressions of the company. Regardless of whether the company employs a human representative or a chatbot, finding ways to humanize the company's service agents can signal the company's dedication to connecting with customers "on a human level," which can offset the impersonal and dehumanizing nature of text-based interactions (while still reaping the benefits of their efficiency).

Our research further shows that one easy and cost-efficient way to humanize a chat agent is to allow, if not to actively encourage, making and correcting occasional typos. This is likely reassuring news to customer service agents, since our findings show that making typos is not a deal-breaker when they interact with customers over text; rather, humanizing errors may even increase their perceived warmth and their perceived helpfulness in the customers' eyes. However, it is also important to recognize that humanizing a chat agent (especially a chatbot) can be a double-edged sword. As we will discuss in the next section, companies should also exercise caution and recognize potential pitfalls when humanizing their chatbots.

Our work also has important implications for policymakers on key issues related to transparency and online security. Given that humanizing a chatbot and/or keeping its identity ambiguous can lead to a variety of benefits, policymakers should step in to provide more guidance on what should and should not be allowed when companies deploy increasingly humanlike chatbots. Indeed, recent regulation in California has made it unlawful for a bot to fail to disclose its identity when interacting with consumers or voters ("Bolstering Online Transparency" Act, California Business and Professional Code, Division 7, Part 3, § 17940). The fact that the home state of many tech giants passed such a law highlights the need for public policy to regulate rapidly developing yet potentially controversial technologies.

Finally, for everyday consumers, our research highlights a cognitive heuristic of theirs that could be deliberately exploited by chatbot developers. This sheds light on the various forces affecting the way consumers perceive communicators, which they might not always be aware of. In the new age of digital communication, where consumers often struggle to discern the identity of their online interaction partners, providing them with knowledge and tools to alleviate these concerns may help them gain a greater sense of control over their digital environment.

Limitations and Future Directions

Beyond expectations of the agent's helpfulness, intention to endorse a reward for the agent, and favorable impressions of the company, in all of our studies, we also examined the disclosure of personal information as another potential behavioral consequence of perceived humanness. While we did not find a consistent main effect on this dependent measure, we did observe a robust mediation effect of humanness perception across all studies, such that participants were more willing to share personal information with a chat agent when they perceived the agent to be more human (see Web Appendix). In light of these early and inconclusive results, future researchers should investigate how humanness perception can impact self-disclosure and other conversation behaviors, as well as the

underlying psychological mechanisms. For example, a human(like) communicator may highlight the possibility of social evaluation and social surveillance (Epley and Waytz 2010), thus promoting compliance and socially desirable behaviors (e.g., Oliveira et al. 2021), while also discouraging information disclosure that one deems embarrassing or socially inappropriate. Such findings will have important implications for online privacy in numerous contexts where consumers interact with chatbots, from online counseling to conversations on social media.

Future research should also investigate whether the number and type of errors may moderate their humanizing effect. While we showed that correcting one or two typos can humanize a chat agent without creating negative consequences for their perceived competence, it is possible that a large number of typos—even when corrected right away—could cause negative impressions (e.g., Kreiner et al. 2002). Furthermore, typos are harmless errors that everyone makes and can easily relate to, which might be why they were not perceived as diagnostic for the agent's competence. It is unclear, however, whether other types of errors, such as grammatical errors or factual inconsistency, would yield similar positive (and lack of negative) outcomes. When a customer service agent provides a piece of incorrect information, we suspect that consumers may be *less* forgiving toward the agent than when having merely seen a typo. Furthermore, given that text-based customer service—either powered by human agents or chatbots—is a common way for consumers to interact with companies nowadays and can strongly impact consumer experiences, this work highlights an important avenue for future research in customer service.

Finally, while humanizing artificial agents can create a variety of positive outcomes, overly humanizing an agent may introduce unintended "side effects." Below, we identify three potential pitfalls that are also worthy of future research: setting too-high expectations beyond a chatbot's actual capabilities, creating an eerily humanlike agent, and misleading consumers about the agent's identity.

First, reaping positive effects of perceived humanness can be tempting (see also Svenningsson and Faraon 2019), yet more humanness perception can also mean higher expectations for the agent's performance (Malle et al. 2021). As discussed in the introduction, people expect human customer service agents to be reliable, empathic, and capable of understanding complex issues (LivePerson 2017), traits that are not necessarily possessed by current conversational AI systems. To avoid setting too high expectations for a human(like) service agent and later disappointing customers due to limited capabilities, companies may consider humanizing chatbots only for tasks where their performance is proven to be reliable, and avoid making chatbots appear more human than is warranted by their actual cognitive and emotional capacities.

Second, while correcting errors can humanize a machine, it is also possible that a machine can be perceived as *overly* humanlike. For instance, prior research has shown that highly humanlike robots often veer into the so-called "uncanny valley," evoking feelings of repulsion in human observers (Gray and Wegner 2012; Mori 1970) and leading to decreased liking, trust, and cooperation (Mathur and Reichling 2016). To date, it is unclear if a text-based communicator would be subject to similar negative effects, so further research is needed to understand when—if ever—a chatbot could appear overly human in consumers' eyes.

Third, misleading people about the identity of a chat agent could create ethical concerns and even backfire. When consumers discover a company's acts of commission (i.e., actively lying about a chatbot's identity) or omission (i.e., keeping the identity of a chatbot ambiguous), they may feel angry and see the company as manipulative or untrustworthy. Future research should investigate how people react to and perceive a convincingly humanlike chatbot when its true identity is revealed *after* the interaction. It is possible that the more convincing the human façade a chatbot puts on, the more consumers rely on increasingly sophisticated tools to identify robot imposters (e.g., Botometer;

https://botometer.iuni.iu.edu/), and the more negatively they may react when the inauthentic imposter is exposed.

Conclusion

In the current digital world, consumers commonly use text-based communication channels to interact with companies. Although texting with a customer service agent can be efficient, the inherent lack of nonverbal cues present in face-to-face conversations also makes it difficult for customers to experience a "human connection" with such an agent. Our research suggests that making and correcting typos can humanize an online chat agent, which in turn leads consumers to form more positive impressions of the agent and the company it represents. Evidently, perceptions of humanness come not only from what is written but also from the process of writing itself.

References

Adams, Tim (2018), "The charge of the chatbots: How do you tell who's human online?" *The Guardian* (November 18), https://www.theguardian.com/technology/2018/nov/18/how-can-you-tell-who-is-human-online-chatbots.

Adam, Martin, Michael Wessel, and Alexander Benlian (2020), "AI-based chatbots in customer service and their effects on user compliance," *Electronic Markets*, 1-19.

Aggarwal, Pankaj and Ann L. McGill (2007), "Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products," *Journal of Consumer Research*, 34 (4), 468-479.

Appel, Jana, Astrid von der Pütten, Nicole C. Krämer, and Jonathan Gratch (2012), "Does Humanity Matter? Analyzing the Importance of Social Cues and Perceived Agency of a Computer System for the Emergence of Social Reactions during Human-Computer Interaction," *Advances in Human-Computer Interaction*, 2012, 1–10.

Araujo, Theo (2018), "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions," *Computers in Human Behavior*, 85, 183–189.

Arechar, Antonio A., and David G. Rand (2121), "Turking in the time of COVID." *Behavior research methods*: 1-5.

Aronson, Elliot, Ben Willerman, and Joanne Floyd (1966), "The effect of a pratfall on increasing interpersonal attractiveness," *Psychonomic Science*, 4, 227-228.

Boden, Margaret (2018), "The robots won't take over because they couldn't care less," *Aeon Essays*, (August 13), https://aeon.co/essays/the-robots-wont-take-over-because-they-couldnt-care-less.

Broadbent, Elizabeth, Vinayak Kumar, Xingyan Li, John Sollers, Rebecca Q. Stafford, Bruce A. MacDonald, and Daniel M. Wegner (2013), "Robots with Display Screens: A Robot with a More Humanlike Face Display Is Perceived To Have More Mind and a Better Personality," *PLoS ONE*, *8*(8), 1–10.

California Bots: disclosure bill, Cal. S. B. 1001 (2017-2018), Chapter 892 (Cal. Stat. 2018).

Candello, Heloisa, Claudio Pinhanez, and Flavio Figueriredo (2017), "Typefaces and the perception of humanness in natural language chatbots," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,* New York, NY: Association for Computing Machinery, 3476–3487.

Chafe, Wallace L. (1982), "Integration and involvement in speaking, writing, and oral literature," in *Spoken and written language: Exploring orality and literacy,* Deborah Tannen, ed. Norwood, NJ: Ablex, 35-53.

Christian, Brian (2012), *The Most Human Human: What artificial intelligence teaches us about being alive*, New York, NY: Anchor Books.

Dal Porto, Lindsay (2017), "Chatbot chatter," *Quality Progress*, 50 (7), 6-9.

DiSalvo, Carl F., Francine Gemperle, Jodi Forlizzi, and Sara Kiesler (2002), "All robots are not created equal: the design and perception of humanoid robot heads," in *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques,* New York, NY: Association for Computing Machinery, 321-326.

Ein-Gar, Danit, Baba Shiv, and Zakary L. Tormala (2011), "When blemishing leads to blossoming: The positive effect of negative information," *Journal of Consumer Research,* 38, 846-859.

Epley, Nicolas and Juliana Schroeder (2014), "Mistakenly seeking solitude," *Journal of Experimental Psychology: General*, 143 (5), 1980.

Epley, Nicolas and Adam Waytz (2010), "Mind perception" in *Handbook of Social Psychology,* S. T. Fiske, D. T. Gilbert, and G. Lindzey, eds. Hoboken, NJ: John Wiley & Sons Inc, 498-541.

Epley, Nicolas, Adam Waytz, and John T. Cacioppo (2007), "On seeing human: a three-factor theory of anthropomorphism" *Psychological Review*, 114 (4), 864.

Fiske, Susan T., Amy J. Cuddy, and Peter Glick (2007), "Universal dimensions of social cognition: Warmth and competence," *Trends in Cognitive Sciences*, 11, 77-83.

Figueredo, Lauren and Connie K. Varnhagen (2005), "Didn't you run the spell checker? Effects of type of spelling error and use of a spell checker on perceptions of the author," *Reading Psychology*, 26, 441-458.

Fulmer, Russ, Angela Joerin, Breanna Gentile, Lysanne Lakerink and Michiel Rauws (2018), "Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial," *JMIR Mental Health*, 5 (4), e64.

Gottsegen, Gordon (2016), "Are you there, chatbot? It's me, human," *CNET*, (December 21), https://www.cnet.com/news/youre-talking-to-ai-chatbots-and-its-just-the-beginning/.

Grand View Research (2021), "Chatbot Market Size, Share & Growth Report, 2021-2028," *Grand View Research, https://www.grandviewresearch.com/industry-analysis/chatbot-market.*

Gray, Heather M., Kurt Gray and Daniel M. Wegner (2007), "Dimensions of mind perception," *Science*, 315, 619.

Gray, Kurt and Daniel M. Wegner (2012), "Feeling robots and human zombies: Mind perception and the uncanny valley," *Cognition*, 125, 125–30.

Haslam, Nick (2006), "Dehumanization: An integrative review," *Personality and Social Psychology Review*, 10, 252-264.

Heider, Fritz and Marianne Simmel (1944), "An experimental study of apparent behavior," *The American Journal of Psychology*, 57, 243-259.

Helmreich, Robert, Eliot Aronson and Jim LeFan (1970), "To err is humanizing sometimes: Effects of self-esteem, competence, and a pratfall on interpersonal attraction" *Journal of Personality and Social Psychology,* 16, 259.

Hoffmann, Laura, Nicole C. Krämer, Anh D. Lam-chi, and Stefan Kopp (2009), "Media Equation revisited: Do users show polite reactions towards an embodied agent?" in *Proceedings of the 9th international conference on Intelligent Virtual Agents IVA 2009,* Amsterdam, The Netherlands: Springer, 159–165.

Huang, Karen, Michael Yeomans, Alison W. Brooks, Julia Minson and Francesca Gino (2017), "It doesn't hurt to ask: Question-asking increases liking," *Journal of personality and social psychology*, 113 (3), 430.

Kemp, Simon (2020), "Report: Most important data on digital audiences during coronavirus," *Growth Quarters—The Next Web*. (April 24), https://thenextweb.com/growth-quarters/2020/04/24/report-most-important-data-on-digital-audiences-during-coronavirus/.

Kiesler, Sara, Jane Siegel and Timothy W. McGuire (1984), "Social psychological aspects of computer-mediated communication," *American Psychologist*, 39, 1123.

Kniffin, Kevin M., Jayanth Narayanan, Frederik Anseel, John Antonakis, Susan P. Ashford, Arnold B. Bakker and Mark van Vugt, (2021), "COVID-19 and the workplace: Implications, issues, and insights for future research and action," *American Psychologist*, 76 (1), 63.

Kim, Seo Y., Bernd H. Schmitt and Nadia M. Thalmann (2019), "Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking," *Marketing Letters*, 1-12.

Kreiner, David S., Summer D. Schnakenberg, Angela G. Green, Michael J. Costello and Anis F. McClin (2002), "Effects of spelling errors on the perception of writers," *The Journal of General Psychology*, 129, 5-17.

Kruger, Justin, Nicolas Epley, Jason Parker and Zhi-Wen Ng (2005), "Egocentrism over e-mail: Can we communicate as well as we think?" *Journal of personality and social psychology*, 89 (6), 925.

Kuchenbrandt, Dieta, Friederike Eyssel, Simon Bobinger and Maria Neufeld (2013), "When a robot's group membership matters: Anthropomorphization of robots as a function of social categorization," *International Journal of Social Robotics*, 5, 409-417.

Larivière, B., Bowen, D., Andreassen, T.W., Kunz, W., Sirianni, N.J., Voss, C., Wünderlich, N.V. and De Keyser, A. (2017), "Service Encounter 2.0": An investigation into the roles of technology, employees and customers. *Journal of Business Research*, 79, 238-246.

Lea, Martin and Russell Spears (1992), "Paralanguage and social perception in computer-mediated communication," *Journal of Organizational Computing and Electronic Commerce*, 2, 321-341.

Lee, Howard (2014), "Paging Dr. Watson: IBM's Watson supercomputer now being used in healthcare," *Journal of AHIMA*, 85 (5), 44-47.

Lee, Min K. (2018), "Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management," *Big Data & Society*, 5 (1), DOI: 10.1177/2053951718756684.

Lin, Mariana (2019), "How we can bring AI personalities to life," *TEDSummit,* https://www.ted.com/talks/mariana_lin_how_we_can_bring_ai_personalities_to_life.

LivePerson (2017), "How consumer view bots in customer care," https://www.liveperson.com/resources-reports/bots-in-customer-care

LivePerson (2019), "How consumers view conversational commerce in 2019," https://www.liveperson.com/resources-reports/consumer-preferences-2019.

Logg, Jennifer M., Julia A. Minson and Don A. Moore (2019), "Algorithm appreciation: People prefer algorithmic to human judgment," *Organizational Behavior and Human Decision Processes,* 151, 90-103.

Looser, Christine E. and Thalia P. Wheatley (2010), "The tipping point of animacy: How, when, and where we perceive life in a face," *Psychological Science*, 21, 1854-1862.

Lortie, Catherine L. and Matthieu J. Guitton (2011), "Judgment of the humanness of an interlocutor is in the eye of the beholder," *PLOS One*, 6 (9), e0025085.

Luangrath, Andrea W., Joann Peck and Victor A. Barger (2017), "Textual paralanguage and its implications for marketing communications," *Journal of Consumer Psychology*, 27, 98-107.

Lucas, Gale M., Albert Rizzo, Jonathan Gratch, Stefan Scherer, Giota Stratou, Jill Boberg and Louis-Philippe Morency (2017), "Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers," *Frontiers in Robotics and AI*, 4, 51.

Malhotra, Naresh K., Sung S. Kim and James Agarwal (2004), "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model," *Information Systems Research,* 15 (4), 336-355.

Malle, Bertram F., Kerstin Fischer, James E. Young, Ajung Moon and Emily C. Collins (2020), "Trust and the discrepancy between expectations and actual capabilities of social robots," in *Human-robot interaction: Control, analysis, and design,* D. Zhang and B. Wei, eds. New York, NY: Cambridge Scholars Publishing.

McCoy, John P. and Tomer D. Ullman (2018), "A minimal Turing Test," *Journal of Experimental Social Psychology*, 79, 1-8.

Memon, Mumtaz A., Cheah J. Hwa, T. Ramayah, Hiram Ting and Francis Chuah (2018), "Mediation Analysis: Issues and Recommendations," *Journal of Applied Structural Equation Modeling*, 2 (1), i–ix.

Metz, Cade and Keith Collins (2018), "To give A.I. the gift of gab, Silicon Valley needs to offend you," *The New York Times* (February 21), https://www.nytimes.com/interactive/2018/02/21/technology/conversational-bots.html.

Mende, Martin, Maura L. Scott, Jenny van Doorn, Dhruv Grewal, and Ilana Shanks (2019), "Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses." *Journal of Marketing Research* 56 (4), 535-556.

Morewedge, Carey K., Jesse Preston and Daniel M. Wegner (2007), "Timescale bias in the attribution of mind," *Journal of Personality and Social Psychology*, 93, 1–11.

Mori, Masahiro (1970), "Bukimi no tani [the uncanny valley]," *Energy,* 7, 33-35.

Nass, Clifford and Scott Brave (2005), "*Wired for speech: How voice activates and advances the human-computer relationship*," Cambridge, MA: MIT Press.

Nass, Clifford and Youngme Moon (2000), "Machines and mindlessness: Social responses to computers," *Journal of Social Issues*, 56, 81-103.

Oliveira, Raquel, Patrícia Arriaga, Fernando P. Santos, Samuel Mascarenhas, and Ana Paiva (2020), "Towards prosocial design: A scoping review of the use of robots and virtual agents to trigger prosocial behaviour." *Computers in Human Behavior*, 106547.

Pierce, Jason R., Gavin J. Kilduff, Adam D. Galinsky and Niro Sivanathan (2013), "From glue to gasoline: How competition turns perspective takers unethical," *Psychological Science*, 24, 1986-1994.

Pinker, Steven and Paul Bloom (1990), "Natural language and natural selection," *Behavioral & Brain Sciences,* 13, 707-784.

Reddy, Trips. (2017), "How chatbots can help reduce customer service costs by 30%," *IBM*. https://www.ibm.com/blogs/watson/2017/10/how-chatbots-reduce-customer-service-costs-by-30-percent/

Schanke, Scott, Gordon Burtch, and Gautam Ray (2021). "Estimating the impact of "humanizing" customer service chatbots." *Information Systems Research*.

Schegloff, Emanuel A., Gail Jefferson and Harvey Sacks (1977), "The preference for self-correction in the organization of repair in conversation," *Language*, 53, 361-382.

Schroeder, Juliana and Nicholas Epley (2015), "The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal," *Psychological Science,* 26, 877-891.

Schroeder, Juliana and Nicholas Epley (2016), "Mistaking minds and machines: How speech affects dehumanization and anthropomorphism," *Journal of Experimental Psychology: General,* 145, 1427-1437.

Schroeder, Juliana, Michael Kardas and Nicholas Epley (2017), "The humanizing voice: Speech reveals, and text conceals, a more thoughtful mind in the midst of disagreement," *Psychological Science*, 28, 1745-1762.

Schroll, Roland, Benedikt Schnurr and Dhruv Grewal (2018), "Humanizing products with handwritten typefaces," *Journal of Consumer Research*, 45 (3), 648-672.

Sommer, Barbara A., and Robert Sommer (2003), "A virtual lab in research methods." *Teaching of Psychology* 30 (2), 171-173.

Turing, Alan (1950), "Computing machinery and intelligence," *Mind,* 59, 433-460.

Vignovic, Jane A. and Lori F. Thompson (2010), "Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation," *Journal of Applied Psychology*, 95, 265.

Waytz, Adam and Michael I. Norton (2014), "Botsourcing and outsourcing: Robot, British, Chinese, and German workers are for thinking—not feeling—jobs," *Emotion*, 14 (2), 434.

Westerman, David, Aaron C. Cross and Peter G. Lindmark (2018), "I believe in a thing called bot: Perceptions of the humanness of 'chatbots'," *Communication Studies*, 1-18.

Wheatley, Thalia, Olivia Kang, Carolyn Parkinson and Christine E. Looser (2012), "From mind perception to mental connection: Synchrony as a mechanism for social understanding," *Social and Personality Psychology Compass*, 6 (8), 589-606.

Wilson, Timothy D., David B. Centerbar, Deborah A. Kermer and Daniel T. Gilbert (2005), "The pleasures of uncertainty: Prolonging positive moods in ways people do not anticipate," *Journal of Personality and Social Psychology,* 88, 5–21.

Zhao, Xinshu, John G. Lynch Jr, and Qimei Chen (2010), "Reconsidering Baron and Kenny: Myths and truths about mediation analysis." *Journal of Consumer Research* 37(2), 197-206.

Zhao, Xuan, Elizabeth Phillips and Bertram F. Malle (2019), "How people infer a humanlike mind from a robot body," forthcoming, https://doi.org/10.31234/osf.io/w6r24

Złotowski, Jakub, Diane Proudfoot, Kumar Yogeeswaran and Christoph Bartneck (2015), "Anthropomorphism: opportunities and challenges in human–robot interaction," *International Journal of Social Robotics*, 7 (3), 347-360.