Human Communication Research 0360-3989

ORIGINAL RESEARCH

Persuasion in the Age of Artificial Intelligence (AI): Theories and Complications of AI-Based Persuasion

Marco Dehnert & Paul A. Mongeau

The Hugh Downs School of Human Communication, Arizona State University, Tempe, AZ, USA

Artificial intelligence (AI) has profound implications for both communication and persuasion. We consider how AI complicates and promotes rethinking of persuasion theory and research. We define AI-based persuasion as a symbolic process in which a communicative-AI entity generates, augments, or modifies a message—designed to convince people to shape, reinforce, or change their responses—that is transmitted to human receivers. We review theoretical perspectives useful for studying AI-based persuasion—the Computers Are Social Actors (CASA) paradigm, the Modality, Agency, Interactivity, and Navigability (MAIN) model, and the heuristic-systematic model of persuasion—to explicate how differences in AI complicate persuasion in two ways. First, thin AI exhibits few (if any) machinic (i.e., AI) cues, social cues might be available, and communication is limited and indirect. Second, thick AI exhibits ample machinic and social cues, AI presence is obvious, and communication is direct and interactive. We suggest avenues for future research in each case.

Keywords: AI-Based Persuasion, Artificial Intelligence, Human–Machine Communication, Computers Are Social Actors, MAIN Model, Machine Heuristic, Source Characteristics

https://doi.org/10.1093/hcr/hqac006

Artificial intelligence (AI) is an exciting, complex, controversial, and increasingly prevalent technology. Since AI entered the realm of persuasion, scholars have attempted to capture the interplay of technology and persuasion with perspectives including captology (Fogg, 2003), persuasive robotics (Siegel et al., 2009), and robot persuasion (Lee & Liang, 2018). We contend that AI has unique characteristics beyond these formulations that complicate the study of persuasion in important ways. As interactions with artificial agents (AAs) become increasingly complex and common, it is crucial to consider how AI complicates existing approaches to communication writ large and persuasion specifically.

Correspondence to Marco Dehnert; e-mail: marco.dehnert@asu.edu

AI and persuasion intersect in both practical and scholarly contexts ranging from political opinions (e.g., Cohen, 2021) to utilizing compliance-gaining strategies (e.g., Lee & Liang, 2019). Despite its wide use, we agree with Donath (2021) that "[t]he potential for technologies such as machine learning and artificial entities to be extraordinarily persuasive is immense—but not yet realized" (p. 162). This is true, in large part, because AI continues to develop at an accelerating rate and because the predominant theories provide only a partial perspective on persuasion. As AI capabilities advance, scholars, designers, and users will have to wrestle with increasingly complex relations between persuasion and AAs.

Gunkel (2020) argued that AI was, and remains, a communication science. By its nature, AI inherently involves communication as it is a defining condition of machine intelligence. Situated within human–machine communication (HMC), two goals underpin our work. First, we define *AI-based persuasion* and identify relevant theoretical perspectives. Second, we consider how AI-based persuasion prompts us to rethink existing persuasive communication concepts. In particular, we explicate two important ways that AI complicates persuasion.

Defining Persuasion, AI, and AI-Based Persuasion

Defining Persuasion

Working from Miller (1980, 2013), Perloff (2017), and Stiff and Mongeau (2016), we define persuasion as a *symbolic process in which a source intends to convince people to shape, reinforce, or change their responses (e.g., attitudes, behaviors, intentions, and source perceptions) through the transmission of a message*. Our definition clearly identifies persuasion as an intentional, message-based, and goal-directed process. For persuasion to occur, a source must intentionally send a message to a receiver(s) with a goal of influencing (i.e., shaping, maintaining, or changing) a response(s). Persuasive messages can come in myriad forms—e.g., verbal or nonverbal; face-to-face or mediated; include arguments, cues, or both—but must be utilized to influence specific responses (Perloff, 2017; Stiff & Mongeau, 2016).

Our definition widens persuasion's scope in two ways. First, we highlight the variety of outcomes targeted by persuasive attempts. Persuasion scholarship typically focuses on attitude change (Perloff, 2017; Stiff & Mongeau, 2016). Stiff and Mongeau, however, argued that several outcomes are appropriate targets for persuasive messaging. Second, our definition includes Cialdini's (2009) view of influence. Although persuasion research typically focuses on a single source and a large audience (e.g., Chaiken & Maheswaran, 1994), Cialdini focuses on persuasive communication in one-on-one contexts (e.g., foot-in-the door; Lee & Liang, 2019). Given their intentional and message-based nature, these phenomena fall within our definition of persuasion.

Defining AI

Defining AI is difficult as both *artificial* and *intelligence* are ambiguous (Gunkel, 2020). Scholars across disciplines struggle to define AI (e.g., Grewal, 2014, in computer engineering; Gunkel, 2020, in communication; Martinez, 2019, in legal studies; and Monett et al., 2020, in AI research). Broadly understood as a "learning algorithm used to approximate some form of intelligence operating within computing machines" (Ninness & Ninness, 2020, p. 100), AI is often divided into general and narrow (or strong and weak). Artificial general (or strong) intelligence represents fully agentic technologies (i.e., simultaneously performing multiple tasks) that emulate human reasoning and intelligence. Although common in science fiction (e.g., HAL in *2001: A Space Odyssey*), artificial general intelligence does not (yet) exist (Broussard, 2018)—if it ever will. Existing AI is narrow, or technology that performs "very specific computational tasks much faster and more accurately than humans" (Ninness & Ninness, 2020, p. 102).

Although definitions vary, Völkel et al. (2020) identified adaptation, automation, and interaction as core aspects of *intelligence*. As such, AI represents step-by-step procedures for solving problems and making decisions using software-driven systems including, but not limited to, algorithms, machine learning, artificial neural networks, computer vision, speech recognition, natural language processing, and robotics (Russell & Norvig, 2021). We find the European Commission's AI definition to be useful:

> Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve a given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. (EU AI HLEG, 2019, p. 6)

Given these understandings of AI, it is important to consider how scholars studied human–computer interactions. Initially, computer-mediated communication (CMC) described two or more humans communicating using a computer (e.g., email or video conferencing). Scholars recently expanded their focus to human–machine communication (HMC) or "the creation of meaning among humans and machines" (Guzman, 2018, p. 1). The term *machine* is intentionally broad as it encompasses a variety of AI- and non-AI-entities (e.g., voice assistants and chatbots). In AI contexts, HMC scholars focus on *communicative AI* (Guzman & Lewis, 2020), that is technologies "such as conversational agents, social robots, and automated-writing software—that are designed to function *as* communicators, rather than merely mediators of human communication" (Lewis et al., 2019, p. 673, emphasis in original). In short, HMC focuses on how humans communicate *with* machines whereas, in CMC, humans communicate *through* them (Fortunati & Edwards, 2020; Guzman, 2018).

Given our emphasis on HMC and communicative AI (i.e., narrow AI that can act as both channel and source; Guzman, 2018), one particularly useful approach is AI-mediated communication (AI-MC; Hancock et al., 2020). In AI-MC "messages are *modified, augmented, or even generated* by a computational agent to achieve communication goals" (p. 90, emphasis in original). Thus, coupling persuasion to AI is fitting given that both concepts are goal-directed.

Defining AI-Based Persuasion

Based on definitions of AI, communicative AI, and persuasion, we define AI-based persuasion as *a symbolic process in which a communicative-AI entity generates, augments, or modifies a message—designed to convince people to shape, reinforce, or change their responses—that is transmitted to human receivers*. Definitions of both persuasion and AI-based persuasion identify intentional, message-based, and goal-directed processes. For AI-based persuasion to occur, an artificial agent must develop all or part of a message, transmitted to a human receiver(s), with a goal of influencing (i.e., shaping, maintaining, or changing) a response(s). We call any such communicative AI technology *persuasive AI*.

Of course, the primary difference between persuasion and AI-based persuasion is the message source (human versus persuasive AI). In AI-based persuasion, persuasive AI must play some role in message construction, including composing the entire message, modifying a message shell for different receivers, or completing a message template (Hancock et al., 2020). The persuasive AI entity might, or may not, also transmit the messages to receivers. In the following section, we continue to contextualize AI-based persuasion by reviewing and rethinking theoretical perspectives typically used at the intersection of AI and persuasion. In particular, we argue that AI forces us to rethink persuasion in meaningful ways.

Explicating AI-Based Persuasion: Theoretical Perspectives

Given that communication scholars have approached computer and artificial technologies from multiple perspectives, we are more interested in using human–human communication (HHC), AI-MC, and HMC as bedrocks from which to explore AI-based persuasion than we are in clearly differentiating them. Our most basic assumption in explicating AI-based persuasion is that some form of AI *is* the message source, even though that may not be readily apparent to receivers (as we introduce below in considering thin and thick AI). Forms of AI are already involved in an array of persuasive contexts, utilizing communication and persuasion processes in novel ways (e.g., mining opinions from social media rather than polling;

Cohen, 2021). Below, we consider two theoretical HMC frames that are most relevant to persuasive AI.

Theoretical Frames and AI-Based Persuasion

In modern life, people are bombarded with information and messages (e.g., emails, news updates, tweets, text messages, and Facebook posts). To cope with this information overload, humans act as cognitive misers who primarily utilize simple decision rules (i.e., heuristics) to form, reinforce, or change responses (Cialdini, 2009; Fiske & Taylor, 2017). Mindfully evaluating messages does occur, but only under certain circumstances. Given its considerable use in HMC scholarship (e.g., Liao & Sundar, 2021), we focus on the heuristic-systematic model (i.e., HSM; Chaiken, 1987; Chaiken & Ledgerwood, 2012).

The HSM of Persuasion

The HSM describes two ways that persuasive messages influence responses: Heuristic and systematic (Chaiken, 1987; Chaiken & Ledgerwood, 2012). Heuristic processing describes the use of simple decision rules (i.e., cognitive heuristics) to base an individual's responses. Simple decision rules represent the default option because they require relatively little cognitive effort. In an AI-based persuasion context, use of the machine heuristic (i.e., computers are objective and free from bias) can generate a defendable attitude that has a high probability of being correct (Sundar 2008; Sundar & Kim, 2019). Three factors determine use of a heuristic: relevance, availability, and accessibility (Bellur & Sundar, 2014; Chen & Chaiken, 1999). Relevance is the extent to which a heuristic is applicable to a cue. Availability refers to the extent to which a heuristic exists in memory. Finally, accessibility represents the extent to which a heuristic can be accessed with minimal cognitive effort.

Systematic message processing, on the other hand, involves exerting considerable cognitive effort in evaluating messages (Chaiken, 1987; Chaiken & Ledgerwood, 2012). Given systematic processing, the HSM predicts that messages containing strong arguments will be more influential than messages containing weak arguments. In part because they involve greater cognitive effort, attitude changes created through systematic processing tend to last longer before decaying, and relate more strongly to behaviors, than changes created through heuristic processing (e.g., Chaiken, 1987; Petty & Cacioppo, 1986).

HMC and Persuasive Message Processing

The original formulations of CASA and MAIN and much, but by no means all, of the scholarship that followed, assumed that whatever persuasive impact computers, web sites, or AI might have, in and of themselves, is likely through the influence of heuristics. Following this trend, to clearly link HMC to the HSM, we consider three sets of cognitive variables: cues, social cues, and scripts.

Although we focus primarily on heuristic processing in this discussion, the use of heuristics is not limited to heuristic processing. Heuristics are learned through experience, so they can contain (or be linked to) troves of information about the message topic. In a political campaign, for example, one candidate could tweet an unflattering photo of the opponent that is relevant to a campaign issue. The photo could generate receivers' systematic processing of the opponent's position or characteristic just as effectively as a full message.

In addition, heuristics can also bias systematic processing. For example, when it is not clear whether messages are strong or weak, receivers will use heuristics to make argument strength (and attitude) judgments about an important topic. Thus, a person whose heuristic is consistent with message arguments will judge those arguments as being stronger (and change their attitude more) than individuals with heuristics that are inconsistent with the arguments (Chaiken & Maheswaran, 1994).

Cues are physical or behavioral features "salient to observers because of their potential as channels of useful information" (Fiore et al., 2013, p. 2). Technological affordances (i.e., action possibilities) can serve as cues, but are only one type of cue (Xu & Liao, 2020). Another important cue set in HMC and AI-MC are social cues, or "physical or behavioral features displayed by a social actor" (Lombard & Xu, 2021, p. 35). For example, facial expressions, eye gaze, gestures, and a human-like voice represent social cues that make an AA (appear) (more) human. (Our use of parentheses reflects competing perspectives regarding whether humans view AAs as actually human or as mimicking human behavior; cf. Fortunati & Edwards, 2020). Other social cues include self-disclosure from a chatbot (Ki et al., 2020; Meng & Dai, 2021).

Some AI technologies (e.g., AI-driven social robots) can provide strong social cues (e.g., human-like voice modulation, synchronized nonverbals, appropriate incorporation of slang). These social cues are particularly important in HMC interactions when they are translated by a receiver into social signals, or "an understanding of the social actor," such as their emotion, personality, empathy, and so on (Lombard & Xu, 2021, p. 32; Fiore et al., 2013). Social signals represent an important step in perceiving technology in human terms (Nass et al., 1994).

Finally, another outcome of processing cues is the initiation of communication scripts, which are cognitive representations of behavioral sequences enacted in particular contexts. Scripts include larger sets of rules that guide communication behavior (Kellermann, 1992). Scripts are learned, reflect heuristics, and as such, can be initiated and enacted with only enough cognitive effort to ensure behavioral appropriateness.

In summary, following Cialdini (2009), certain sets of social cues in HMC—when interpreted as social signals—can trigger a heuristic that, in turn, initiates a behavioral and/or cognitive script. Next, we describe two approaches for understanding AI-based persuasion that focus on cues, heuristics and scripts: The Computers Are Social Actors (CASA) paradigm and Sundar's (2008) MAIN (Modality, Agency, Interactivity, and Navigability) model.

Computers Are Social Actors

Although it predated contemporary AI, the CASA paradigm (Nass & Moon, 2000; Nass et al., 1994) generated much HMC and AI-MC scholarship. Put simply, CASA postulates that humans treat machines as if they are people (Nass & Moon, 2000). Thus, CASA claims that people's interactions with technology are mediated, for the most part, through heuristics and scripts. Specifically, social cues act as triggers that initiate the same heuristics utilized when communicating with a person (Nass et al., 1994; see also Nass & Moon, 2000).

Although scripts and heuristics have considerable value for AI-based persuasive outcomes, Gambino et al. (2020) claimed that CASA does not apply to all interactions with all technologies. Instead, they claimed that CASA applies only to media agents, or "any technological artifact that demonstrates sufficient social cues to indicate the potential to be a source of social interaction" (p. 73). Recently, Lombard and Xu (2021) proposed a hierarchy among social cues. People are evolutionarily sensitive to primary cues (e.g., eye gaze, gestures, or a human-sounding voice) because of their role in social perception. Secondary cues such as variations in language, motion, or robot size "have less power in activating users' social perception and responses" (Lombard & Xu, 2021, p. 34), and are less powerful in activating scripts.

CASA shifts attention from a communicator's (actual or perceived) ontological class (machine or human) toward the interaction (i.e., social cues, heuristics, and scripts). Thus, "what matters most is not the partner's humanness, but what occurs in the interaction itself" (Ho et al., 2018, p. 726). The focus shifts from who is (not) human to what happens between communicators, including which social cues act as triggers for which script(s) in a given context.

Scholars of AI-based persuasion can use CASA to investigate the capacity of AI-based sources, and the social cues they generate (in addition to persuasive messages) to trigger heuristics and scripts. Such a goal is consistent with the CMC literature where foci shifted from individual channels (e.g., Facebook or TikTok) to their affordances (Evans et al., 2017) and impacts. Affordances are the focal point of another important theoretical perspective, Sundar's (2008) MAIN model.

The MAIN Model

Sundar's (2008) MAIN model focuses on technologies' effects on message processing. Originally focused on source credibility judgments (Sundar, 2008), Sundar et al. (2019) recently expanded the MAIN model to online persuasion writ large (including AI-based persuasion) "by simply changing the outcome variable from credibility to attitudes and behaviors" (p. 82). Specifically, the model focused on "identifying cues in the technology of the interface that can impact user cognitions and attitudes, regardless of the content of the persuasive appeals" (Sundar et al., 2019, p. 83). Consistent with the HSM, MAIN asserts that technological affordances, as cues, trigger heuristics that influence source credibility (and other) judgments.

Sundar (2008) divided an impressive array of heuristics in human–machine interactions into four categories: Modality (M), Agency (A), Interactivity (I), and Navigability (N). Both the agency and interactivity affordances are particularly relevant to AI-based persuasion. The agency affordance capitalizes on confusion about the communication source in online contexts (see Complication #1, below). The interactivity affordance focuses on the social cues that facilitate seamless human–AI conversation (see Complication #2, below).

With the advent of HMC, machines can extend beyond mediating HHC by taking active roles in augmenting or modifying HHC (Hancock et al., 2020) and autonomously generating entire messages (Guzman, 2018). For example, chatbots' self-disclosure during conversation may have emotional, relational, and psychological effects (Ho et al., 2018), whereas communicative AI may take on a news-writing role in automated journalism (Lewis et al, 2019). Similarly, studies in advertising (e.g., Rodgers, 2021) and machine-generated artwork in cross-cultural settings (Xu et al., 2020) have explored the effect of source in online and machine contexts, offering insights as to how AI communicators are perceived, especially when compared to human communicators.

In AI-based persuasion contexts, the machine heuristic is particularly important (Sundar, 2008; Sundar & Kim, 2019). Specifically, the machine heuristic

> is activated when (a) agents deliver machinic cues that are (b) peripherally [i.e., heuristically] processed when MH [machine heuristic]-related beliefs are accessible, (c) resulting in attributions of systematicity, randomness, and/or objectivity; activation of that heuristic (d) influences judgments about an interaction or about the information created, selected, or validated by that machine. (Banks et al., 2021, p. 325)

The machine heuristic has profound implications not only for how humans judge AAs, but also for the effects of AAs on source credibility and overall persuasive effectiveness. The influence of the machine heuristic, however, is likely to vary depending on individual differences and contextual factors (Lombard & Xu, 2021). Sundar (2020), for example, claims that "The machine heuristic . . . may result in positive or negative expectations and experiences depending upon the appropriateness of applying machine attributes to the activity at hand" (p. 80). For example, Xu et al. (2020) differentiated human tasks (associated with creativity or emotion, e.g., producing art) from mechanical tasks (associated with instrumentality and rationality, e.g., processing credit-card information; Sundar & Kim, 2019). Moreover, Spence et al. (2019) reported that a robot news anchor was judged as less credible when suspicion was primed, which, in turn, decreased behavioral intention. Finally, people differ in their attitudes toward AI (e.g., Kieslich et al., 2021; Liang & Lee, 2017; Lobera et al., 2020).

Given the conceptual muddiness of persuasion, AI, and AI-based persuasion, summarizing the nature of AI-based persuasion requires nuanced theoretical perspectives and investigations. Based on these theoretical formulations, we devote the remainder of our essay to explore how AI complicates, and prompts us to rethink, the study of persuasion.

AI Complicates Persuasion

Given its nature, AI complicates persuasion, in part, because it plays different roles and performs different tasks across contexts. We focus on two complications that highlight how these differences influence message processing. The first complication focuses on *thin AI*, where few (if any) machinic (i.e., AI) cues exist, where social cues point toward a human or corporate (rather than an AI) source, and where HMC is limited and indirect. In the second complication, we focus on *thick AI*. In this case, AI presence is obvious (e.g., an AI-imbued social robot), where ample (i.e., quantity and quality; Lombard & Xu, 2021) machinic and social cues are displayed, and where HMC is direct and interactive. In line with previous work on the machine heuristic (Banks et al., 2021), we use the terms machinic and AI-based cues synonymously for our purposes. Both thick and thin AI complicate persuasion, but in different ways. We consider each case next.

Complication #1: Thin Persuasive AI Obfuscates the Persuasive Source

A persuasive message source is often obscured: Politicians use speechwriters and advertising agencies produce messages for corporations. In online settings, confusion concerning message sources is also likely as information can travel a sinuous route. Sundar (2008) provides an example focusing on a newsgroup message forwarded from one user to another that also included a third newsgroup user, an article from a newspaper website, and a wire service report, leading "to a confusing multiplicity of sources of varying levels of perceived credibility" (Sundar, 2008, p. 73). Adding AI to the amalgam of sources jumbles matters even further. Each source is a potential cue that, depending on availability and relevance, could initiate a heuristic(s) that may influence source and other persuasive judgments (Banks et al., 2021; Sundar, 2020).

Determining the role of AI in message production is difficult, stemming (in part) from the technology's nature. For example, given the autonomy affordance (Sundar, 2008), once given a goal, AI can independently collect huge amounts of data (both topic- and audience-related), look for patterns in those data, construct a persuasive message (or messages), and transmit it. Our notion of *thin AI* parallels AI-MC, which refers to "mediated communication between people where a computational agent operates on behalf of a [human] communicator by modifying, augmenting, or generating messages to accomplish communication or interpersonal [or persuasive] goals" (Hancock et al., 2020, p. 90). In thin AI, receivers are likely unaware of AI's role in creating, augmenting, or modifying the message. Receivers, instead, are likely to conclude that the source was another human (Hancock et al., 2020). This misattribution likely occurs because messages generated by thin AI contain few machinic cues but can contain ample social cues. These social cues, however, point to a person or organization rather than the AA that developed the message. Thus, social cues are misleading, obfuscating the true source.

Directions for Future Research

Future research on Complication #1 should investigate the influence of machinic cues, social cues, and message arguments in AI-based persuasive messages. For example, future studies could explore perceptions and influences of AI sources in combination with positively-, versus negatively-valenced organizational sources. This research could also consider participants' attitudes toward organizational sources, message topic, and AI.

Finally, scholars could explore how users' expectations related to AI involvement in communicative interactions impact message processing. For example, in the online dating algorithm context, Sharabi (2021) introduced *placebo AI* as people's expectations for successful matching algorithms that may be more important than how well these technologies actually work. Research could explore this phenomenon in AI-based persuasion contexts (and beyond) as AI-based expectations may influence message processing and, potentially, the effectiveness of persuasive AAs.

Complication #2: Thick Persuasive AI and Human–AI Relationships

Our second complication focuses on what we call *thick AI*; i.e., contexts where AI-imbued technologies are apparent (perhaps unmistakable) to receivers. We focus specifically on media agents (Gambino et al., 2020) where AI technologies generate ample primary and secondary social cues (e.g., human-sounding voice, appropriate language use, supportive verbal responses, as well as appropriate and synchronized gestures) during repeated conversations with a human. Unlike thin AI, social cues in thick AI focus directly on the entity itself (rather than some person or organization). These repeated interactions, in turn, can facilitate relationship development between humans and AAs. Relationship development complicates persuasion because it is typically strongly associated with perceptions of trust, which are key elements of persuasive interactions (Perloff, 2017; Stiff & Mongeau, 2016).

For human–AI relationships to develop, two important changes will likely have to occur. First, the human will likely treat the AA as if it were human. Second, the human must evaluate the communicative AI partner positively enough, and enjoy interactions with them enough, to consider the other as friend or companion. We discuss both changes and then tie them to persuasion.

From Technology to Human(like)

For a relationship between a human and AA to develop, the former is likely to anthropomorphize the latter. For example, despite the realization that virtual assistants are software-based, some users consider Siri "like a little person in [their] phone" (Guzman, 2019, p. 347). Anthropomorphizing occurs, in part, because the human translates social cues provided by the AA, across multiple interactions, into social signals (Lombard & Xu, 2021). Social signals, in turn, enliven technology with human characteristics (e.g., emotions and empathy). Translating cues into signals likely involves heuristics (based on human–human relational development). Translation also likely activates both heuristic (e.g., scripts) and systematic processes (mindful evaluation of the conversation content) also gleaned from human–human persuasion experiences (Banks et al., 2021; Gambino et al., 2020; Lombard & Xu, 2021). Szczuka et al. (2019) developed a similar explanation in their sexual interaction illusion model.[1] Specifically, Szczuka et al. claim that a subjective sensation (or illusion) of interacting

> with a real [social] partner includes the subjective perception among users that the other is really present (in the here and now), physically embodied and alive (rather than just inanimate or lifeless technology), and human (if the other is meant to display a human character). (p. 5)

Consistent with CASA, over time, people consider and treat virtual assistants (and other AI-based technologies) *as if* they are human.

Social cues, the social signals they generate, and systematic processes (e.g., elaboration on conversation) also likely create and maintain the MAIN model's notion of flow and Szczuka et al.'s (2019) willing suspension of disbelief, where humans forgive minor social-cue faux pas to maintain an illusion of HHC. Consistent with both CASA and MAIN, however, serious errors (reboots and language changes) interrupt flow.

Transition from Virtual Assistant to Friend

The social cues, heuristics, scripts, and systematic processing that facilitate the treatment of an AA as if they are human (Lombard & Xu, 2021) also relate to friendship formation (Croes & Antheunis, 2021; Ki et al., 2020). Thus, human–AI friendship development seems possible (although, perhaps, in a different form than human–human counterparts; Ryland, 2021). In short, Siri sounds like a human (and a human friend), responds like a human (and a human friend), such that it is possible to consider Siri a friend.

Considering an AA as human is likely a necessary, but not a sufficient, condition for relational development. Similar to human-to-human relationships, human-to-AA relationships likely develop from repeated, supportive, and mutually involving interactions between partners that include cues, heuristics, social signals, and mindful consideration of conversation that likely generate both heuristic and systematic processing. For example, much communication in relationships is everyday talk (Duck, 1994). On the other hand, relational transitions (such as first dates; e.g., Laner & Ventrone, 2000) are likely associated with considerable systematic processing of communicative, personal, and relational information.

Fehr (2008) describes situational, individual, and dyadic factors associated with human–human friendship development that are also relevant to human–AI relationships. Situational factors, including anticipation of future interaction, familiarity, and availability are likely AI's strength. Thus, these characteristics likely facilitate human-AA relationships as interactive AI technologies are constantly available and never tired of interaction (Sundar, 2020).

It is in Fehr's (2008) individual (social skills and responsiveness) and dyadic (e.g., humor) relational development factors where current AAs fall short. Although Croes and Antheunis (2021) claimed that human–AI friendships are not yet possible, limiting factors were primarily technological: i.e., the AA's lack of memory, humor, empathy, and the superficiality of interaction. Thus, although human–AI friendships are possible, technological improvements, and humans becoming more comfortable with human-technology relationships, will contribute significantly to their future development. In summary, whereas Complication #1 applies to current AI technology and processes, Complication #2 looks toward the future.

AI–Human Relationships and Persuasion

Treating AAs as both human and a relational other (see also Ki et al., 2020; Lombard & Xu, 2021) has important implications for source evaluations and persuasion. Anthropomorphism, for example, is linearly related to credibility perceptions (Gong, 2008) that, in turn, are important to acceptance of messages (Perloff, 2017; Stiff & Mongeau, 2016) and new technologies (Ghazali et al., 2020). Thus, as interactions with thick AI continue, the AA's communication skills improve, and AAs provide consistently useful and credible information, trust might be manifested in a number of positive heuristics (Siri is expert and trustworthy, Alexa is my friend) that can facilitate attitude and behavior change heuristically, systematically, or both.

As repeated interactions develop, human users will also likely focus on the AA (e.g., Siri or Alexa) themself/itself, as the conversation message source, rather than the corporation behind it (i.e., Apple or Amazon; Ischen et al., 2020). What is more, heuristics developed from these interactions could serve as a mental model for future interactions with other AAs. So, increased exposure to, and experience with, AAs will likely have an influence on trust and credibility towards AI writ large.

Finally, the persuasive role of thick AI will likely be extremely complex as there are many technological variables (e.g., voice only versus embodied; humanoid versus machine-like) that will likely interact with contextual and individual difference factors to influence persuasion. In particular, an AA's embodiment is important to human's immersion in, and reactions to, an interaction: A review of over 30 experimental studies found that physically present and embodied robots (when compared with entities digitally displayed on a screen) were more persuasive and generally perceived more positively (Li, 2015). Thus, physical presence may provide thicker social cues (Li, 2015)—particularly primary cues (Lombard & Xu, 2021)—highlighting the importance of affordances in the MAIN model (Sundar, 2008). Therefore, future research should investigate how the persuasive influence of a technology's modality, agency, interactivity, and navigability interacts with social cues to influence persuasive effectiveness and perceptions of presence.

Directions for Future Research

Research addressing Complication #2 can proceed in a variety of ways. Development of trust in different AI–human relational forms (Ryland, 2021) has implications for CASA, the MAIN model, the machine heuristic, and persuasion. Scholars could also investigate the relevance and applicability of human–human relationship scripts (e.g., friendship, intimacy, trust) to human–machine relationships (Ryland, 2021; Westerman et al., 2020). In particular, the role of trust in the AA has important implications for persuasion given these judgments can generate (or bias) systematic processing (Banks et al., 2021; Chaiken & Maheswaran, 1994) and/or generate systematic processing (Chaiken & Ledgerwood, 2012). Under what conditions will humans be most open to AI-based persuasion, heuristically, systematically, or both? The answer to this question will likely depend on the "thickness" of AI, where thicker AI is more likely to generate systematic processing. In addition, the effectiveness of AI-based persuasion will likely be influenced (i.e., moderated and/or mediated) by a wide variety of variables including receivers' individual differences, contextual variables, and experience with AAs. Finally, research should explore the role of time on relationship development and persuasion. Although exploratory, Croes and Antheunis (2021) offer a useful longitudinal template on human–bot relationship formation (i.e., participants interacted with a chatbot every three days for three weeks) that is useful for future investigations.

Conclusions

Although persuasion is a complicated enterprise (e.g., Perloff, 2017; Stiff & Mongeau, 2016), when AI enters the conversation (quite literally), it becomes even more complex. Drawing from existing theory and research, we defined AI-based persuasion and explored how AI complicates persuasion. In doing so, we introduced the concepts of thick and thin AI. Thick and thin AI differ in the quantity and quality of both machinic (i.e., AI) cues and social cues, as well as the opportunity they provide for direct human–machine interaction. In our view, thin and thick AI are the extremes of a continuum between which most cases will fall. Future studies should identify factors important to a technology's placement on the continuum and how they are combined.

Given the complexity of AI-based persuasion, the influence of persuasive AAs will likely prove difficult to summarize. Perceptions and influence of persuasive AI (thin, thick, or in-between) will likely depend on an extensive array of mediators and moderators, including, among others, the technology utilized (e.g., its location on the thick–thin continuum), the organizational source (actual or claimed), HMC history, the receiver's HMC and/or AI-based heuristics and scripts, and receivers' attitude towards AI. Individual differences and contextual factors should also be considered when investigating AI-based persuasion (Lombard & Xu, 2021). Moreover, as interactions between humans and machines develop, HMC scripts will change as well (Gambino et al., 2020). Finally, one claim that we can confidently make is that AI will continue to evolve at a fast rate. What we know as AI in 2022 will likely be much different from the AI of 2025, 2030, or beyond. Thus, AI prompts us to rethink the role of individual differences, contextual factors, and socio-temporal dynamics in humans' perception of, and reactions to AI-based persuasion.

Therefore, future research should utilize multiple methodological and paradigmatic approaches and should extend beyond existing (human–human) persuasive communication-based questions. For instance, scholars could focus on social issues such as the role of different types of human–AI relationships on AI-based persuasion (Croes & Antheunis, 2021; Ki et al., 2020). Cultural questions could center on how cultural understandings of AI influence existing and developing AI-based persuasion heuristics and scripts (Broussard, 2018). Ethical questions could involve consideration of the extent to which disclosure of AI-sources is necessary, or whether (perceptions of) friendship between AAs and humans is ethically appropriate (Donath, 2021; Ryland, 2021). Legal questions could explore the role of AI-sources in issues of persuasive authorship, agency, and accountability (e.g., in the context of fake news) (Martinez, 2019). Political questions could center on how to define AI in the rapidly changing technical and cultural landscape (EU AI HLEG, 2019). Finally, psychological questions could explore individual characteristics and circumstances that may moderate the influence of AI cues on persuasive outcomes (Sundar, 2020; Szczuka et al., 2019). As AI technology advances, and as humans' interactions with actual AAs develop, we believe that research on AI-based persuasion—as the study of mediator and moderator variables—will benefit from addressing the issues and complications raised in this essay.

Note

1. We apply their work on sex robots to nonsexual human–AI relationships. Hence, we replaced the word *sexual* in the original quote with *social.*

References

Banks, J., Edwards, A. P., & Westerman, D. (2021). The space between: Nature and machine heuristics in evaluations of organisms, cyborgs, and robots. *Cyberpsychology, Behavior, and Social Networking, 24*(5), 324–331. https://doi.org/10.1089/cyber.2020.0165

Bellur, S., & Sundar, S. S. (2014). How can we tell when a heuristic has been used? Design and analysis strategies for capturing the operation of heuristics. *Communication Methods and Measures, 8*(2), 116–137. https://doi.org/10.1080/19312458.2014.903390

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. The MIT Press.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. *5*, pp. 3–39). Erlbaum.

Chaiken, S., & Ledgerwood, A. (2012). A theory of heuristic and systematic message processing. In P. A. M. van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *The handbook of theories of social psychology* (Vol. *1*, pp. 224–245). SAGE.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*(3), 460–473. https://doi.org/10.1037/0022-3514.66.3.460

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). The Guilford Press.

Cialdini, R. B. (2009). *Influence: Science and practice* (5th ed.). Pearson Education.

Cohen, M. D. (2021). *Modern political campaigns: How professionalism, technology, and speed have revolutionized elections*. Rowman & Littlefield.

Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships, 38*(1), 279–300. https://doi.org/10.1177%2F0265407520959463

Donath, J. (2021). Commentary: The ethical use of powerful words and persuasive machines. *Journal of Marketing, 85*(1), 160–162. https://doi.org/10.1177/0022242920973975

Duck, S. (1994). *Meaningful relationships: Talking, sense, and relating*. SAGE.

Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication, 22*(1), 35–52. https://doi.org/10.1111/jcc4.12180

Fehr, B. (2008). Friendship formation. In S. Sprecher, A. Wenzel, & J. Harvey (Eds.) *Handbook of relationship initiation* (pp. 29–54). Psychology Press.

Fiore, S. M., Wiltshire, T. J., Lobato, E. J. C., Jentsch, F. G., Huang, W. H., & Axelrod, B. (2013). Toward understanding social cues and signals in human-robot interaction: Effects of robot gaze and proxemic behavior. *Frontiers in Psychology, 4*, 859. https://doi.org/10.3389/fpsyg.2013.00859

Fiske, S. T., & Taylor, S. E. (2017). *Social cognition: From brains to culture* (3rd ed.). SAGE.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do.* Morgan Kaufmann Publishers.

Fortunati, L., & Edwards, A. (2020). Opening space for theoretical, methodological, and empirical issues in human-machine communication. *Human-Machine Communication*, *1*, 7–18. https://doi.org/10.30658/hmc.1.1

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication, 1*, 71–85. https://doi.org/10.30658/hmc.1.5

Ghazali, A. S., Ham, J., Barakova, E., & Markopoulos, P. (2020). Persuasive robots acceptance model (PRAM): Roles of social responses within the acceptance model of persuasive robots. *International Journal of Social Robotics, 12*, 1075–1092. https://doi.org/10.1007/s12369-019-00611-1

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, *24*(4), 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007

Grewal, D. S. (2014). A critical conceptual analysis of definitions of artificial intelligence as applicable to computer engineering. *IOSR Journal of Computer Engineering, 16*(2), 9–13.

Gunkel, D. J. (2020). *An introduction to communication and artificial intelligence*. Polity.

Guzman, A. L. (2018). Introduction: "What is human-machine communication, anyway?" In A. L. Guzman (Ed.), *Human-machine communication: Rethinking communication, technology, and ourselves* (pp. 1–28). Peter Lang.

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior, 90*, 343–350. https://doi.org/10.1016/j.chb.2018.08.009

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human-machine communication research agenda. *New Media & Society, 22*(1), 70–86. https://doi.org/10.1177%2F1461444819858691

Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication, 25*(1), 89–100. https://doi.org/10.1093/jcmc/zmz022

HLEG AI (2019). *A definition of AI: Main capabilities and scientific disciplines.* https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Ischen, C., Araujo, T., van Noort, G., Voorveld, H., & Smit, E. (2020). "I am here to assist you today": The role of entity, interactivity and experiential perceptions in chatbot persuasion. *Journal of Broadcasting & Electronic Media, 64*(4), 615–639. https://doi.org/10.1080/08838151.2020.1834297

Kellermann, K. (1992). Communication: Inherently strategic and primarily automatic. *Communication Monographs, 59*(3), 288–300. https://doi.org/10.1080/03637759209376270

Ki, C.-W., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior, 111*, 106412. https://doi.org/10.1016/j.chb.2020.106412

Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The threats of artificial intelligence scale (TAI): Development, measurement and test over three application domains. *International Journal of Social Robotics, 13*, 1563–1577. https://doi.org/10.1007/s12369-020-00734-w

Laner, M. R., & Ventrone, N. A. (2000). Dating scripts revisited. *Journal of Family Issues, 21*(4), 488–500. https://doi.org/10.1177/019251300021004004

Lee, S. A., & Liang, Y. (2018). Theorizing verbally persuasive robots. In A. L. Guzman (Ed.), *Human-machine communication: Rethinking communication, technology, and ourselves* (pp. 119–143). Peter Lang.

Lee, S. A., & Liang, Y. (2019). Robotic foot-in-the-door: Using sequential-request persuasive strategies in human-robot interaction. *Computers in Human Behavior, 90*, 351–356. https://doi.org/10.1016/j.chb.2018.08.026

Lewis, S. C., Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial intelligence and journalism. *Journalism & Mass Communication Quarterly, 96*(3), 673–695. https://doi.org/10.1177%2F1077699019859901

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies, 77*, 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

Liang, Y., & Lee, S. A. (2017). Fear of autonomous robots and artificial intelligence: Evidence from national representative data with probability sampling. *International Journal of Social Robotics, 9*, 379–384. https://doi.org/10.1007/s12369-017-0401-3

Liao, M., & Sundar, S. S. (2021). When e-commerce personalization systems show and tell: Investigating the relative persuasive appeal of content-based versus collaborative filtering. *Journal of Advertising*. https://doi.org/10.1080/00913367.2021.1887013

Lobera, J., Rodríguez, C. J. F., & Torres-Albero, C. (2020). Privacy, values and machines: Predicting opposition to artificial intelligence. *Communication Studies, 71*(3), 448–465. https://doi.org/10.1080/10510974.2020.1736114

Lombard, M., & Xu, K. (2021). Social responses to media technologies in the 21st century: The media are social actors paradigm. *Human-Machine Communication, 2*, 29–55. https://doi.org/10.30658/hmc.2.2

Martinez, R. (2019). Artificial intelligence: Distinguishing between types & definitions. *Nevada Law Journal, 19*(3), 1015–1042.

Meng, J., & Dai, Y. (2021). Emotional support from AI chatbots: Should a supportive partner self-disclose or not? *Journal of Computer-Mediated Communication, 26*(4), 207–222. https://doi.org/10.1093/jcmc/zmab005

Miller, G. R. (1980). Afterword. In D. P. Cushman & R. D. McPhee (Eds.), *Message–attitude–behavior relationship* (pp. 319–327). Academic Press.

Miller, G. R. (2013). On being persuaded: Some basic distinctions. In J. P. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 70–92). SAGE. (Reprinted from *Persuasion: New directions in theory and research*, pp. 11–28, by M. Roloff & G. R. Miller, Eds., 1980. SAGE).

Monett, D., Lewis, C. W. P., & Thórisson, K. R. (Eds.). (2020). Special issue "On defining artificial intelligence"—Commentaries and author's response [Special issue]. *Journal of Artificial General Intelligence, 11*(2), 1–100. https://doi.org/10.2478/jagi-2020-0003

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72–78).

Ninness, C., & Ninness, S. K. (2020). Emergent virtual analytics: Artificial intelligence and human-computer interactions. *Behavior and Social Issues, 29*, 100–118. https://doi.org/10.1007/s42822-020-00031-1

Perloff, R. (2017). *The dynamics of persuasion: Communication and attitudes in the 21st century* (6th ed.). Routledge.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer.

Rodgers, S. (2021). Themed issue introduction: Promises and perils of artificial intelligence and advertising. *Journal of Advertising, 50*(1), 1–10. https://doi.org/10.1080/00913367.2020.1868233

Russell, S., & Norvig, P. (Eds.). (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Ryland, H. (2021). It's friendship, Jim, but not as we know it: A degrees-of-friendship view of human–robot friendships. *Minds & Machines 31*, 377–393. https://doi.org/10.1007/s11023-021-09560-z

Sharabi, L. L. (2021). Exploring how beliefs about algorithms shape (offline) success in online dating: A two-wave longitudinal investigation. *Communication Research, 48*(7), 931–952. https://doi.org/10.1177/0093650219896936

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive robotics: The influence of robot gender on human behavior. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2563–2568). IEEE. https://doi.org/10.1109/IROS.2009.5354116

Spence, P. R., Edwards, C., Edwards, A., & Lin, X. (2019). Testing the machine heuristic: Robots and suspicion in news broadcasts. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 568–569). IEEE. https://doi.org/10.1109/HRI.2019.8673108

Stiff, J. B., & Mongeau, P. A. (2016). *Persuasive communication* (3rd ed.). Guilford.

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital Media, Youth, and Credibility* (pp. 73–100). The MIT Press.

Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human-AI interaction (HAII). *Journal of Computer-Mediated Communication, 25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., & Kim, J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–9). https://doi.org/10.1145/3290605.3300768

Sundar, S. S., Xu, Q., & Dou, X. (2019). The role of technology in online persuasion: A MAIN model perspective. In S. Rodgers & E. Thorson (Eds.), *Advertising Theory* (pp. 70–88). Routledge.

Szczuka, J. M., Hartmann, T., & Krämer, N. C. (2019). Negative and positive influences on the sensations evoked by artificial sex partners: A review of relevant theories, recent findings, and introduction of the sexual interaction illusion model. In Y. Zhou & M. H. Fischer (Eds.), *AI love you: Developments in human-robot intimate relationships* (pp. 3–19). Springer. https://doi.org/10.1007/978-3-030-19734-6_1

Völkel, S. T., Schneegass, C., Eiband, M., & Buschek, D. (2020). What is "intelligent" in intelligent user interfaces? A meta-analysis of 25 years of IUI. *IUI '20: Proceedings of the 25th International Conference on Intelligent User Interfaces*, 477–487. https://doi.org/10.1145/3377325.3377500

Westerman, D., Edwards, A. P., Edwards, C., Luo, Z., & Spence, P. R. (2020). I-it, I-thou, I-robot: The perceived humanness of AI in human-machine communication. *Communication Studies, 71*(3), 393–408. https://doi.org/10.1080/10510974.2020.1749683

Xu, K., & Liao, T. (2020). Explicating cues: A typology for understanding emerging media technologies. *Journal of Computer-Mediated Communication, 25*(1), 32–43. https://doi.org/10.1093/jcmc/zmz023

Xu, K., Liu, F., Mou, Y., Wu, Y., Zeng, J., & Schäfer, M. S. (2020). Using machine learning to learn machines: A cross-cultural study of users' responses to machine-generated artworks. *Journal of Broadcasting & Electronic Media, 64*(4), 566–591. https://doi.org/10.1080/08838151.2020.1835136