Please quote as: Tolzin, A. & Janson, A. (2026). Uncovering the mechanisms of common ground in human–agent interaction: review and future directions for conversational agent research. Internet Research, 36, 292–315. doi: 10.1108/INTR-06-2023-0514

Internet Research

Uncovering the mechanisms of common ground in human–agent interaction: review and future directions for conversational agent research

Received 30 June 2023
Revised 31 December 2023
27 May 2024
13 June 2024
18 November 2024
Accepted 19 November 2024

Antonia Tolzin
Research Center for Information System Design (ITeG), University of Kassel, Kassel, Germany, and
Andreas Janson
Institute of Information Systems and Digital Business, University of St Gallen, St Gallen, Switzerland

Abstract

Purpose – Human–agent interaction (HAI) is increasingly influencing our personal and work lives through the proliferation of conversational agents (CAs) in various domains. As such, these agents combine intuitive natural language interactions by also delivering personalization through artificial intelligence capabilities. However, research on CAs as well as practical failures indicates that CA interaction oftentimes fails miserably. To reduce these failures, this paper introduces the concept of building common ground for more successful HAIs.

Design/methodology/approach – Based on a systematic literature analysis, we identified 38 articles meeting the eligibility criteria. We critically reviewed this body of knowledge within a formal narrative synthesis structured around the use of common ground in the interaction with CAs.

Findings – Based on the systematic review, our analysis reveals five mechanisms for achieving common ground: embodiment, social features, joint action, knowledge base and mental model of conversational agent. We point out the relationships between these mechanisms as they are related to each other in directional and bidirectional ways.

Research limitations/implications – Our findings contribute to theory with several implications for CA research. First, we provide implications about the organization of common ground mechanisms for CAs. Second, we provide insights into the mechanisms and nomological network for achieving common ground when interacting with CAs. Third, we provide a broad research agenda for future CA research that centers around the important topic of common ground for HAI.

Originality/value – We offer novel insights into grounding mechanisms and highlight the potentials when considering common ground in different HAI processes. Consequently, we secure further understanding and deeper insights of possible mechanisms of common ground to shape future HAI processes.

Keywords Common ground, Conversational agent, Human–agent interaction, Systematic review

Paper type Literature review

The results presented were partially developed in the research projects: Komp-HI funded by the German Federal Ministry of Education and Research (BMBF, grant 16DHBKI073), and Managing the Algorithm: Prompt Engineering for AI-based Systems as an Emerging Business Skill by the Swiss National Science Foundation (SNSF, grant number: 221281). We thank the BMBF and SNSF for supporting our research. Furthermore, this research builds on the paper "Mechanisms of Common Ground in Human-Agent Interaction: A Systematic Review of Conversational Agent Research" that has been presented at Hawaii International Conference on System Sciences (HICSS), 2023, in Maui authored by Tolzin and Janson (2023). We thank the reviewers and attendees for their valuable feedback, which helped us improve our research and write this paper. Finally, we thank the Associate Editor for his guidance and the three anonymous reviewers for their constructive feedback and openness during the review process.

Statement of competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Internet Research

e-ISSN: 2054-5657
p-ISSN: 1066-2243
DOI 10.1108/INTR-06-2023-0514

INTR

1. Introduction

Today, artificial intelligence (AI) is pervasive as it influences a lot of different areas in our private and working lives, for instance through AI-based conversational agents (CAs). AI mimics our human natural intelligence, in this context, it can lead an intelligent conversation with a human counterpart (Elshan *et al.*, 2022). CAs frequently assist in our individual lives (Janson, 2023), with OpenAI's ChatGPT being a significant recent innovation in this field (Schöbel *et al.*, 2024). These CAs interact with humans to achieve goals, improve through learning and are recognized as a unique type of information system (IS) entity characterized by their intelligence and interaction capabilities (Elshan *et al.*, 2022). Consequently, the collaboration between humans and CAs is becoming crucial, highlighting the importance of investigating their conversational aspects (Seeber *et al.*, 2020a, b).

While these AI-based CAs emulate human-like behavior and interaction patterns (Vassallo *et al.*, 2010), their reliance on rule-based mechanisms or large language models (LLMs) distinguishes them from human counterparts. Despite their potential for effective communication, these distinctions create fragility in conversational interactions, leading to dialogues that can easily fail due to misunderstandings (Benner *et al.*, 2021). Such failures impact the effectiveness between humans and CAs (Luger and Sellen, 2016). CAs sometimes provide none, wrong, or incomprehensible responses, which leads to discomfort, annoyance, and questioning of the capabilities of the CA, which can end up in usage discontinuance and hinder the future success and spread of CAs (Weiler *et al.*, 2022; Chakrabarti and Luger, 2015). Thus, improving the communication between humans and CAs is an important goal for researchers and practitioners (Meredith, 2017) as dialogue understanding is an inherently interactive process.

Despite the ambiguity of language, human dialogue partners understand each other remarkably well. This is because they establish common ground (Clark and Brennan, 1991). The interactive process to achieve dialogue understanding is called conversational grounding and describes the coordinative process through which dialogue partners establish mutual knowledge, beliefs, and assumptions, collectively referred to as common ground (Koulouri *et al.*, 2016). Following, we define common ground as the mutual and individual knowledge, beliefs, assumptions, and expectations shared by communicative partners during an interaction (Clark, 1996). Nevertheless, identifying the needs and capabilities in the context of human-computer interaction (HCI) and developing presumptions about what the CA can do and understand is a great challenge. On the other way around, it is also difficult for system designers to guess how the human part of the dialogue will act (Koulouri *et al.*, 2016). Moreover, achieving common ground is an identified gap for human-machine teaming (Wray *et al.*, 2022), but despite its importance, there are several ambiguities concerning the definition and application of common ground, shared understanding, or grounding in human-agent communication (Kontogiorgos, 2022). This might be due to the connection of different research fields (common ground from psychology and linguistics; and CAs from HCI) and leads to confusion about the actual meaning and underlying mechanisms of common ground in human-agent interaction (HAI). Further, comprehending how to achieve common ground between two individuals and transferring these mechanisms to human-agent communication is critical for enhancing conversational interactions. Beyond purely perfect answers, it is important to gather a common ground between users and CAs because this understanding fosters a coherent mental model that aligns user expectations with the capabilities of CAs. Such alignment ensures more effective communication, enabling CAs to provide not just accurate but contextually relevant responses, ultimately enhancing user satisfaction and interaction quality. Critically, to date, there is no systematic review examining how common ground is achieved in HAI, highlighting a significant gap in understanding the mechanisms that foster common ground between humans and CAs.

The above-described lack of comprehensive investigation underscores the necessity for an in-depth review elucidating the intricacies of establishing common ground within human-agent communication, an area vital for advancing the effectiveness and quality of interactions

between users and CAs. Therefore, we review in this paper prior literature and contribute to the literature of common ground by answering the following research questions:

RQ1. How is common ground in communication between a human and a conversational agent achieved?

RQ2. What are the relationships between the mechanisms for achieving common ground in the interaction between a human and a conversational agent?

We answer the first research question by reviewing previous studies to distill the mechanisms of common ground. Then, to answer the second research question, we conducted a detailed analysis of how the mechanisms found relate to one another. By answering these questions, we provide an assessing overview of the relevant findings for common ground in CA research, identify research gaps, and provide an agenda for research avenues to engage the uprising stream of research that tackles the future of HAIs.

2. Theoretical background

2.1 Common ground

It is well investigated that senders adapt to the supposed needs and capabilities of the recipient in a dialogue (Clark and Brennan, 1991). That is why people speak differently to children, friends, foreigners, and colleagues. Dialogue understanding is an interactive process, aiming to resolve misunderstandings and build mutual understanding. Therefore, common ground is predominantly a communication theory. Psycholinguistic research has examined how dialogue partners achieve understanding and prevent misunderstandings and communicative breakdowns. The social and collaborative aspects of human conversation are well established. Theories in more linguistic areas focus on how individuals coordinate their linguistic behavior to promote understanding through shared language patterns and examine how individuals adapt their language to become more like their partners, thereby promoting psychological closeness [1]. To establish common ground, an interactive process to generate understanding in conversations is needed (Koulouri *et al.*, 2016). Common ground contains the background knowledge on which the communication planning of the conversational partners builds on and can be divided into the global knowledge (all knowledge about the conversational partner and their knowledge requirements) and the situational knowledge (knowledge about the mutual perception conditions and the communication protocol). Grounding processes lead to a coordination of the background knowledge of the conversational partners (Clark, 1996; Clark and Brennan, 1991).

Successful grounding results in a shared context, guided comprehension, instant feedback of actions, and enhanced processes in conveying intent (Brennan, 1998). Dialogue partners form shared representations of what they are talking about, and then jointly modify these representations (Brennan, 1991). Through adding mutually understood contents, a continuously updated shared knowledge base is created and is available for subsequent use in the conversation (Clark and Wilkes-Gibbs, 1986). This joint process of adding and adapting contents of common ground from conversational turn to conversational turn has been formalized in the contribution model (Clark and Schaefer, 1989). During the conversation, senders "check" that addressees understand what they are saying. In contrast, addressees provide different forms of evidence to let senders know about their understanding. For example, this could be implicit acknowledgments, like reacting to requests or active listening, or explicit feedback, like nodding, or saying "okay" or "Sorry . . . ?" (Clark and Schaefer, 1989; Clark and Brennan, 1991). This verbal grounding does not bring new information or arguments to the conversation. It is more like a semantic mechanism to check that both dialogue partners received and understood the sender's contribution. Responses can get interconnected and contingent on what has been said previously in the conversation by mutually grounding the conversational partner's input (Sundar *et al.*, 2010). This is a

fundamental element of human communication and gives the sender a signal that the dialogue partner is actively listening (Ghose and Barua, 2013). It is crucial to distinguish between disagreement and miscommunication. Achieving common ground is not about dialogue partners agreeing with each other but forming appropriate meta-perspectives in relation to the conversational context (Clark and Brennan, 1991).

As we focus on common ground in our study, we must acknowledge conceptual clarity regarding other concepts that share commonalities with common ground. Shared understanding (see Table 1 for a definition) is an important concept that shares certain aspects with common ground studied in this paper. Coming back to our initial definition, common ground refers to the mutual and individual knowledge, beliefs, assumptions, and expectations shared by dialogue partners. In contrast, shared understanding encompasses a broader cognitive alignment regarding the interpretation of concepts between dialog partners. Shared understanding is defined as a comprehensive and stable mutual agreement among individuals or groups regarding the interpretation of concepts, the value of properties, and mental models that transcend immediate interaction and foster effective collaboration toward common goals (Bittner and Leimeister, 2014). In consequence, the two concepts share a conceptual overlap in certain contexts. Following both definitions, the concepts have in common that they form a basis for effective communication and coordination among individuals or groups (Foundation in Communication). Moreover, they facilitate interaction by enabling individuals to understand each other better, fostering smoother communication (Enhancing Interaction) and they contribute to collaborative efforts by aligning knowledge and comprehension (Supporting Collaboration). Lastly, both concepts can evolve or change over time based on ongoing interactions or discussions (Dynamic).

With the paper focusing on common ground, it is also important that we look at the conceptual differences of both concepts. First, common ground primarily focuses on immediate context and interaction, whereas shared understanding extends beyond immediate interaction, encompassing broader domains or subjects (Scope of Application). Second, shared understanding tends to be relatively stable over time, while common ground is more dynamic and subject to change during interaction (Stability). Third, common ground evolves during interaction, adapting to the immediate communicative needs, while shared understanding remains more static (Evolution). Fourth, shared understanding implies a

Table 1. Contrasting common ground and shared understanding

Concept	Definition	Key characteristics
Common ground	Common ground refers to the mutual and individual knowledge, beliefs, and assumptions shared by communicative partners during a conversation. It involves the context-specific information necessary for effective communication and coordination (Clark, 1996; Clark and Brennan, 1991)	- Mutual knowledge shared by communicative partners - Context-specific information for effective communication - Dynamic and evolving during interaction (Clark and Brennan, 1991)
Shared understanding	Shared understanding involves a comprehensive and stable mutual agreement among individuals or groups regarding the interpretation of concepts, the value of properties, and mental models that transcends immediate interaction and fosters effective collaboration toward common goals (Bittner and Leimeister, 2014)	- Deeper and broader comprehension beyond immediate interaction - More comprehensive and relatively stable - May include broader concepts or domains (Bittner and Leimeister, 2014; Johnson and O'Connor, 2008)

Source(s): Authors' own work

more profound level of shared knowledge or comprehension, often incorporating complex or abstract concepts compared to common ground, which deals with more context-specific information (Depth of Knowledge).

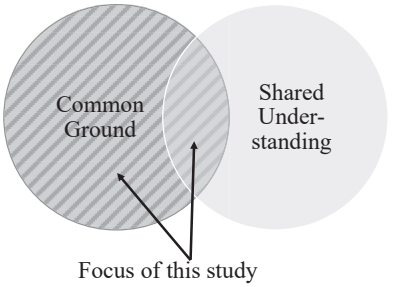
While our primary focus is on common ground, we also include shared understanding in the following review to ensure a comprehensive examination of the topic (see highlighted areas of Figure 1). Common ground, rooted in basic psychology, is not widely recognized in IS research. Consequently, researchers often use similar concepts under different names unknowingly. To address this, it is important to clearly delineate these concepts, even though this paper focuses on common ground.

Given the importance of common ground for effective communication and coordination, exploring the role of mental models is essential. Mental models play a significant role in contributing to common ground (Lee *et al.*, 2005; Johnson and O'Connor, 2008; Bittner and Leimeister, 2014). These shared, organized cognitive representations help individuals within a team to align their knowledge, beliefs, and assumptions about key elements of their environment, including equipment, tasks, and team member characteristics (Mohammed *et al.*, 2000). In HAI, mental models help users form accurate expectations about the CA's abilities and behavior, crucial for establishing and evolving common ground. Understanding mental models provides insights into mechanisms facilitating effective communication and coordination, contributing to the dynamic nature of common ground for mutual agreement. By shaping users' perceptions and expectations, mental models guide the interaction process, making it more efficient and reducing the likelihood of misunderstandings (Lee *et al.*, 2005). Incorporating mental models into this paper's framework enhances understanding of achieving and maintaining common ground in HAIs. This insight is essential for designing more effective CAs, ultimately improving user satisfaction and interaction quality. Overall, common ground, with its dynamic nature and focus on context-specific information, is key to understanding HAI processes more profoundly.

2.2 Conversational agents

Common ground is essential for both human-to-human and human-agent communication. Social Response Theory suggests people apply social rules and expectations to interactions with CAs, treating them like human counterparts (Nass and Moon, 2000). This theory suggests users expect CAs to understand and respond appropriately within a social context, requiring common ground. In HAI, common ground helps CAs interpret user input accurately and respond relevantly, creating a natural interaction experience. Thus, establishing common ground is crucial for effective dialogue.

CAs encompass software enabling conversation with computers, such as chatbots, virtual agents, and artificial conversational entities. They lead intelligent conversations and have a



Source(s): Authors' own work

Figure 1. Theoretical focus of this study

long history, with notable examples like ELIZA, ALICE, Claude, and HeX (Schöbel *et al.*, 2024). CAs are used for various reasons, in all areas of services, and in different contexts (Gnewuch *et al.*, 2017; Serban *et al.*, 2017). CAs possess certain natural language processing capabilities that enable on the input side a recognition of user inputs and understanding (Knote *et al.*, 2021; Schmitt *et al.*, 2023). Based on the user input, CAs that possess intent recognition capabilities react through a set of established rules or flows to the queries posed by users and provide an output (Schöbel *et al.*, 2024). In contrast, open-domain CAs based on (large) language models provide user outputs through next-token prediction based on user input. Besides that, CAs can be distinguished by their interaction interface and modality (Zierau *et al.*, 2023; Skantze and Doğruöz, 2023). CAs might include voice as an interaction modality (Schmitt *et al.*, 2023), e.g. Amazoń s Alexa or Google Assistant, and typically make use of natural language interfaces and machine learning techniques, which allow them to react to user inputs via voice-based interactions and synthesized speech (Budiu, 2018). In contrast, text-based CAs interact with the user via a text interface and can be technically distinguished into two types: messenger-like agents and embodied CAs (Hobert and Meyer von Wolff, 2019; Zierau *et al.*, 2020). The first type can be seen as interaction mode characteristics at a more mechanical dimension of CAs (Zierau *et al.*, 2020). Messenger-like agents contain regular chat interfaces, known from Facebook Messenger or WhatsApp, and can be deployed to a variety of use cases.

Examining the characteristics and capabilities of CAs helps us understand how to establish and maintain common ground with users. Aligning theoretical common ground mechanisms with practical CAs highlights its significance in effective HAIs, enhancing interaction dynamics across various modalities.

2.3 Bridging common ground and conversational agents

Common ground is crucial in both human-to-human and human-agent communication, especially for CAs. While its importance in human communication is well-documented, its application to CAs is less explored. Social Response Theory suggests humans apply social rules to interactions with CAs, expecting appropriate understanding and responses (Nass and Moon, 2000). Common ground mechanisms are crucial for the effectiveness and efficiency of human-agent dialogue. They shape interactions, promote understanding, and enhance communicative dynamics as users engage with CAs (Clark and Brennan, 1991). For example, CAs can adapt to specific language styles (Ahmad *et al.*, 2022), and these language styles and elements (e.g. mimicking intonation or specific words) can influence the establishment of common ground. This alignment in communication style facilitates smoother interactions and reduces misunderstandings, thereby enhancing the overall user experience.

This nuanced connection between common ground and CAs is a crucial, yet underexplored facet. CAs, which are often considered social actors due to their emulation of human-like behavior and interaction patterns (Feine *et al.*, 2019), inherently require the establishment of common ground for effective dialogue. This crucial link between common ground and CAs serves as a linchpin for communication effectiveness, guiding the coordination process between users and CAs to promote understanding and enhance conversation dynamics. Interactions with low common ground, though possible, often result in less satisfying, less successful, and more error-prone communications (Clark, 1996). These interactions can lead to misunderstandings and breakdowns, underscoring the necessity of common ground in achieving successful and meaningful HAIs.

The literature lacks a comprehensive exploration of how common ground mechanisms impact HAI effectiveness, necessitating a robust bridge between theory and practice. An in-depth investigation of common ground mechanisms in shaping effective communication is imperative. This study aims to enhance understanding of successful dialogue in CAs, advancing CA research and improving CA design and implementation.

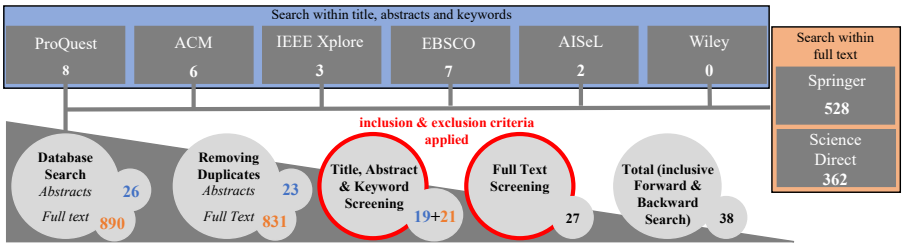
3. Method

In this section, we present our research approach and methodology. We performed a systematic literature analysis (SLA) as this facilitates a comprehensive analysis across diverse disciplines (Paré *et al.*, 2015), including IS, psychology, and HCI, given our research's interdisciplinary nature. Webster and Watson's (2002) framework underscores the structured approach of an SLA, ensuring our examination of common ground in HAI maintains methodological rigor. Rowe (2014) highlighted the significance of an SLA in identifying literature gaps and discerning patterns among studies, critical for our pursuit to unveil the mechanisms behind common ground. Furthermore, this review type not only summarizes current knowledge but also propels future research directions, in line with our aim to provide valuable insights for the future study of CA interaction (Leidner, 2018). We performed an extensive search of eight major bibliographic databases, which were, AISeL, IEEE Xplore, ACM Digital Library, ProQuest, EBSCO, Science Direct, Wiley, and Springer, to include perspectives from the domains IS, technical engineering, and communication. We selected search strings combining terms related to CAs and common ground, aiming to identify a wide range of literature. We used different keywords from recent publications to describe these concepts. Due to linguistic ambiguity and overlap between "common ground" and "shared understanding," we included studies investigating either concept for a comprehensive review. This led to the following search string:

(chatbot* OR "intelligent agent*" OR "intelligent assistant*" OR "intelligent personal assistant*" OR "virtual agent*" OR "virtual assistant*" OR "smart agent*" OR "smart assistant*" OR "conversational agent*" OR "conversational assistant*" OR "communicative agent*" OR "communicative AI" OR "social robot*") AND ("common ground" OR "shared understanding"))

The database-specific search strings were semantically equivalent but adapted to each search engine's syntax and capabilities. We included all peer-reviewed journal articles and conference papers in English, searching titles, abstracts, and keywords, except in Springer and ScienceDirect, where we searched full texts and then screened titles, abstracts, and keywords. No publication year limits were set to maximize relevant articles. The search included both empirical and non-empirical studies, as the review focuses on concepts rather than research outcomes. There are several inclusion criteria. First, the study must address the interaction, dialogue, or communication between CAs and humans. Second, the study must focus on common ground theory or theories of shared understanding in communication. Third, the paper must be available in English. Articles examining communication and interaction more holistic (e.g. human-robot interaction or full conversational avatars in virtual reality environments) were also included, but only verbal and conversational aspects were investigated in this review.

Following recent reviews and protocols, the literature review process involved four rounds of selecting relevant papers (Zhou *et al.*, 2023; Schöbel *et al.*, 2024). Each round was conducted independently and then matched for a joint result. In the first round, the database search yielded 26 records (23 after duplicate removal) for title, abstract, and keywords in six databases, and 890 records (831 after duplicates removal) for full text in the other two databases (Figure 2). In the next round, papers were screened for relevance by reviewing titles and abstracts based on inclusion and exclusion criteria. There are several exclusion criteria. First, studies that did not explicitly address interaction, dialogue, or communication between CAs and humans were excluded. Second, studies that did not focus on common ground theory, theories of shared understanding in communication, or did not emphasize verbal and conversational aspects were not selected. Third, papers not available in English were removed. Lastly, papers were excluded if they focused solely on non-verbal or non-conversational aspects of HAI without exploring verbal communication domains. The references were categorized as matching, maybe matching, and not matching the inclusion criteria. All papers that possibly meet inclusion criteria (matching or maybe matching) were retrieved as full text in the third round. A total of 40 full-text versions of studies were selected for further



Source(s): Authors' own work

Figure 2. Literature search process

investigation—19 from the screening in six databases based on titles, abstracts, and keywords, and an additional 21 identified through full-text searches in two other databases. This led to the inclusion of 27 studies. The fourth and last round identified additional backward and forward references to capture papers not covered through the database search. Through screening the references and applying forward searches using Google Scholar, eleven papers were added to the list, resulting in 38 papers for in-depth analysis. The final list of included papers can be found in Appendix 2.

The papers included were critically reviewed by a formal narrative synthesis structured around the use of common ground in the interaction with CAs, and pre-defined criteria for data extraction were used (Popay *et al.*, 2006).

To derive possible mechanisms of common ground in HAI different aspects of common ground implementation were coded. In the first step, we coded in which context common ground was embedded (e.g. "Common ground and common interest are necessary components of engagement" (McKeown, 2015) was coded as *engagement*) and the specific mechanism to reach common ground (e.g. "common ground was conceptualized as personalization, where information would be remembered by the agent to tailor their experience" (Clark *et al.*, 2019) was coded as *personalization: remembering information*). In the second step, studies with similar contexts and specific mechanisms were grouped together and generic terms were developed to describe these different sets to achieve common ground. The terms were iteratively derived from the different perspectives found in the literature. We extracted and categorized these terms based on the language and terminology used by the source authors themselves. Sometimes we used the same terms as the authors (e.g. knowledge base (Blache, 2017)) and sometimes we made minor changes (e.g. joint activity vs joint action (Frijns *et al.*, 2021)). Subsequently, the first author trained a graduate student coder who then independently applied and assigned the five codes to the paper, blinded to the coding of the first author. Inter-rater reliability (Cohen's kappa; $\kappa = 0.59$) showed moderate agreement for the common ground mechanisms. Both raters resolved any notable discrepancies through discussion until agreement was reached on a single consensus mechanism.

4. Results

4.1 General characteristics of included articles

We assess the state of the art concerning common ground in HAI so far. The 38 included articles were all published over the preceding 21 years (2003–2023) with 30 (78.95%) papers published in the last ten years (2014–2023) and twelve (31.58%) papers published in 2021, 2022, or 2023 (see also Appendix 1). In general, an increasing interest in common ground theories in the application of CAs can be observed in the last few years. Almost half of the papers examined had an empirical foundation (18, 47.37%), but there were also some theoretical (8, 21.05%) and some conceptual (12, 31.58%) articles. Looking more specifically at the different characteristics of the investigated CAs, we see that most CAs are voice-based

agents (12, 31.58%), some are text-based agents (7, 18.42%) and in few articles, CAs have both, voice- and text-based elements (7, 18.42%). In six papers (15.79%) no specific information about the CA was given or there were only theoretical considerations about CAs in general. As articles investigating human-robot communication were also included, six papers (15.79%) did not include a specific CA, but investigated a speech-based human-robot interaction.

Next, we analyzed the application context. Exploring the diverse application contexts of CAs illuminates the involved dynamics of HAIs, unveiling varied strategies employed by agents in, e.g. healthcare, education, customer service, and social interactions. The most common application context for CAs was a collaborative scenario or some kind of task performance (e.g. physical tasks, cooking, schedule change) with nine articles (23.68%) included. Three CAs were used in a medical context (7.89%) and three CAs in education (7.89%). Two CAs were applied in public spaces (5.26%), namely a guide in a museum providing background information and an orientation guide at an airport. Moreover, two CAs were used in the context of aeronautics and spacecraft operations (5.26%). One CA each was applied in the context of customer service, product complaints, house inspection, and marketing. Three papers (7.89%) contained more than one application context and tested or discussed the use of CAs in different domains. The last eleven articles (28.95%) did not name an application context or examine the usage of CAs more generally and only theoretically.

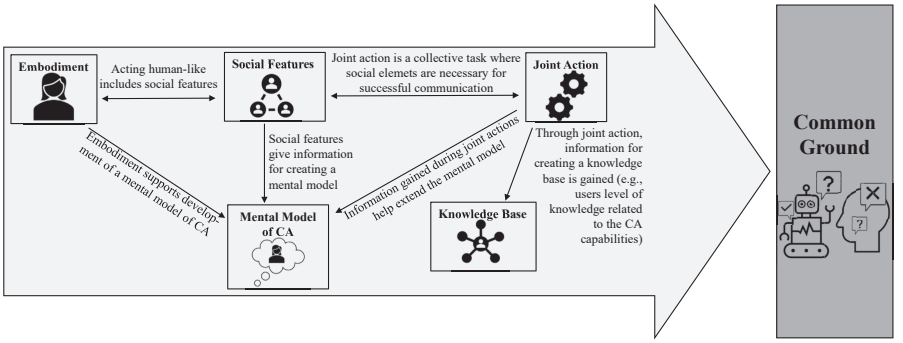
The theoretical foundation of common ground or shared understanding was broad, but most articles (30, 78.95%) were based on the common ground theory proposed by the psycholinguist Clark (1996) or on more than one theory including Clark's definition of common ground. Furthermore, articles were primary or secondary based on other theories, namely Wittgenstein (1967), Stalnaker (2002), Fusaroli *et al*. (2014), Kanda *et al*. (2004), and Pickering and Garrod (2004). Only two (5.26%) papers considered shared understanding and not common ground, with one paper theoretical based on Dillenbourg (2008) and one paper without a theoretical foundation (Hanna and Richards, 2014). The remaining six (15.79%) articles had their own definition or no theoretical foundation of common ground. We see a dominance of the common ground theory of Clark, which is favorable as the ambiguous and heterogeneous interpretations of common ground seem to have the same base, and a widespread and general definition of common ground distinguished from everyday use is possible in future research.

Lastly, we differentiated whether common ground was directly or indirectly addressed and investigated in the included studies. Indirect investigations contain all articles where common ground was not the main focus of the study, e.g. concepts contributing to common ground were studied without a main focus and direct assessment of common ground, or the indirect investigation was just a result of post-hoc examinations not addressed beforehand. The exploration shows a semi-balanced ratio of direct (22 paper, 57.89%) and indirect (16 paper, 42.11%) investigation of common ground.

4.2 Mechanisms of common ground

To answer the first research question, we reviewed past studies to distill the mechanisms of common ground. Based on various implementation strategies and contexts, we identified five underlying mechanisms of common ground. We use the term "mechanism" to encompass the diverse processes, cognitive elements, and socio-cognitive factors contributing to common ground in HAI. While it may imply causality, it signifies our intent to cover a comprehensive array of essential components shaping this phenomenon [2]. Figure 3 depicts these mechanisms as an overview of their relationship to achieve common ground in HAI.

Embodiment. This mechanism refers to presenting a CA as an identifiable counterpart with human-like appearances and behaviors, either physically or virtually. It includes visual appearance and non-verbal communication cues, enhancing user interaction by fostering rapport and believability (Kim *et al.*, 2018). Embodiment also includes voice-based



Source(s): Authors' own work

Figure 3. Relationships of the five mechanisms to achieve common ground

communication as this is also crucial in establishing human-like interactions, particularly in virtual environments where it significantly impacts the perceived presence and effectiveness of the CA. Ten papers were assigned to this mechanism. Embodiment involves endowing CAs with human-like appearances and social behaviors, paralleling face-to-face interaction mechanisms to establish rapport and mutual understanding (Pickering and Garrod, 2004). People are more willing to establish common ground if they interact with a CA under similar situational, social, and psychological conditions as in human interactions (Corti and Gillespie, 2016; Pustejovsky and Krishnaswamy, 2021). This can be supported by a human-like body and presenting the CA as an autonomously communicating person in either a physical or virtual environment. Integrating human-like characteristics warrants human-like treatment and presents an authentic entity. Embodiment extends beyond conversational common ground to include non-verbal communication, aligning minds by being interesting, creative, and humorous (see *Social Features* and *Mental Model of CA*). For example, Strohmann *et al.* (2023) analyzed Replika, a virtual companion service with human-like design, showing thoughtful and natural verbal and non-verbal design. It appears as a full-bodied avatar, and over time, it becomes a constant conversation partner, friend, and companion (Murphy, 2019). Instead of a blank chat window, showing an animated avatar could help users learn more about potential areas of agreement (Skantze and Doğruöz, 2023).

A special case of common ground through embodiment is the virtual simulation environment, embedding the interaction within a shared context. In such environments, human and agent share the same virtual space, facilitating shared situational and perceptual references. For instance, in a medical virtual 3D environment, the user acts as a patient, and the CA acts as a doctor, both represented with full-body avatars. They can refer to visible objects (e.g. stethoscope) because of shared situational references, creating shared perceptual and epistemic common ground. Here, voice-based communication is also important for achieving perfect human-like doctor-patient interactions, fostering embodiment and human-like behavior.

Embodiment serves as a foundational mechanism for HAI by encompassing both physical and behavioral attributes of the CA. Acting as an identifiable conversational counterpart, embodiment fosters familiarity and enables the incorporation of social features within the interaction. Integrating human-like attributes into CAs facilitates common ground, fostering connection and alignment between users and artificial conversational entities, ultimately enhancing the effectiveness of human-agent dialogue.

Social features. Social features refer to behaviors and attributes that facilitate effective social interaction, such as coherence, trust, humor, and active listening. These features create authenticity and understanding in conversations, essential for establishing common ground

between users and CAs. Nineteen papers were assigned to this mechanism. Social elements play a vital role in both human conversation and establishing common ground with CAs (Neururer *et al.*, 2018). Social features like facework, affective strategies, trust, humor, and active listening help manage interactions by facilitating understanding and reducing misunderstandings. To show authenticity, a CA needs to be transparent about its purpose, learn from experiences, and exhibit human-like behavior and coherence. Coherence involves social and contextual awareness and the ability to relate to common experiences to establish common ground (e.g. Clark *et al.*, 2019; Neururer *et al.*, 2018). Social features strongly facilitate establishing common ground between CAs and users, fostering authenticity and meeting social expectations.

Skantze and Doğruöz (2023) differentiated between three speech events with distinct social behaviors and expectations: Informal/superficial talk (e.g. small talk, gossip, sports talk, reminiscing), Involving talk (e.g. love talk, conflict, discussing problems, complaining), and Goal-oriented talk (e.g. group discussion, lecture, planning, requesting favors). Furthermore, studies highlight the importance of trust and coherence in creating social connections (e.g. Clark *et al.*, 2019; Neururer *et al.*, 2018). Social features are particularly important in task-oriented and shared situations between humans and CAs. Familiarity can be established through personal information disclosure or discussing common topics (e.g. weather, political news) (Cassell, 2000). These examples illustrate how social features are essential in creating common ground between CAs and users. By facilitating effective communication and fostering authenticity, social features help meet social expectations, enhancing the overall connection and interaction quality between users and CAs, demonstrating their crucial role in establishing common ground.

Joint action. Establishing and maintaining common ground is crucial in human joint action and HAI. Joint action involves both conversational partners working toward a shared goal and establishing common ground. Sixteen papers were assigned to this mechanism. Joint action parallels the interactive process essential for resolving misunderstandings and building common ground. It emphasizes the need for joint attention and coordinated actions, aligning with the idea of coordinated background knowledge of conversational partners (Clark and Brennan, 1991). The sender must ensure the recipient has the necessary context to interpret the message. Joint attention is crucial for successful communication (Vinciarelli *et al.*, 2015). In linguistic communication, joint attention to both senders' and recipients" intentions is essential. This collective task can only be achieved with conversational partners, representing the situation to make common ground salient. However, it remains challenging for CAs to account for senders' intentions and motivations through joint attention (e.g. Bernard and Arnold, 2019). For example, in Cho and Rader (2020) study, joint action occurred between participants and the Google Home voice assistant, bounded by a task instructing users to ask for information from Google Home. The conversation's goal was aligned with the task's goal. Common ground is transparently shared knowledge between partners.

In summary, joint action involves establishing and maintaining common ground and goals between conversational partners. It requires joint attention, transparent knowledge exchange, shared meaning negotiation, and coordinated actions to reduce coordination costs.

Knowledge base. This mechanism involves a task-oriented dialogue system utilizing its comprehensive knowledge, including contextual information, evolving interaction-based information, and stakeholders' beliefs, to establish and update shared understanding with conversational partners, adapting to their knowledge level and existing common ground. Five papers were assigned to this mechanism. The knowledge base mirrors the importance of contextual information and evolving interaction-based knowledge in establishing common ground, foundational for effective communication (Clark and Brennan, 1991). It aligns with the concept that common ground involves background knowledge essential for communication planning (Clark, 1996). This mechanism is crucial for task-oriented dialogue systems, as the CA has pre-access to all task-related knowledge.

For example, Lorini (2020) introduced new semantics for logics of explicit and implicit beliefs based on multi-agent belief bases, differentiating between belief and knowledge conceptually, mathematically, and computationally. The CA relies on three knowledge sources: context, evolving interaction information, and stakeholder beliefs (Blache, 2017). In the second source common ground plays a crucial role and describes the initial user knowledge combined with the information instantiated during the dialogue. Effective common ground requires all conversational partners to have access to the same information and knowledge. New information is added relative to existing common ground, organizing the knowledge base specifically (Buschmeier and Kopp, 2013). To achieve common ground, a CA must adapt to the user's knowledge level and the common ground between user and CA (e.g. Blache, 2017). These examples illustrate how the knowledge base mechanism supports establishing and updating common ground in HAIs. By utilizing comprehensive and evolving knowledge, the CA can adapt to the user's understanding and existing common ground, facilitating more effective and context-aware communication. Therefore, the knowledge base serves as a foundational mechanism for achieving common ground by continuously integrating and updating contextual information and interaction-based knowledge, ultimately enhancing human-agent dialogue effectiveness.

Mental model of conversational agent. This mechanism involves individuals understanding and forming expectations about the CAs abilities, intentions, and behavior. This influences their willingness to invest effort in establishing common ground and considering the CA's perspective, actions, and adaptability. Twelve papers were assigned to this mechanism. People try to understand the intent and meaning behind what the conversational partner is saying (Di Maro, 2021). This is important for exploring common interests, goals, pre-given information, and environmental perception. Users should develop an appropriate mental model of the CAs' abilities and intentions. Estimating the CA's role and knowledge. This mental model also includes any anthropomorphism (see *Embodiment*), leading to expectations of the CA's behavior.

For example, during conversations, people create a mental model of the CA (Frijns *et al.*, 2021). This mental model then shapes their expectations of the CA and influences the effort they invest in establishing common ground (e.g. Frijns *et al.*, 2021; Kiesler, 2005). Moreover, Strohmann *et al.* (2023) developed the design principle of adaptivity and adaptability to create a CA companion. This principle supports establishing common ground, as meaningful communication and mutual understanding in interpersonal relationships require it. It enhances the ability to take the mental perspective of the other conversational partner and predict their actions (Wray *et al.*, 2022).

This mechanism explores how individuals form expectations about a CA's capabilities and intentions, similar to understanding meta-perspectives in conversations (Clark and Brennan, 1991). Shared mental models influence team performance and HAI by guiding users' perceptions and expectations about the CA's behavior and adaptability (Mohammed *et al.*, 2000). These shared mental models affect context-specific information exchange, evolving common ground, and the stability and adaptation of communication between users and CAs (Johnson and O'Connor, 2008; Mohammed *et al.*, 2000). Understanding and developing appropriate mental models of CAs is crucial for enhancing human-agent dialogue effectiveness and ensuring successful interaction outcomes.

4.3 Relationships of the mechanisms

The relationships between the mechanisms of common ground are directional and bidirectionally linked. In order to answer the second research question, the following section provides an in-depth analysis of these interrelationships based on Figure 3.

Embodiment and social features. Embodiment significantly influences the formation of social features. When a CA is equipped with human-like characteristics, it can engage in behaviors such as coherence, trust-building, humor, and active listening. These social features,

in turn, enhance the embodiment mechanism by reinforcing the CA's human-like presence, making it more relatable and authentic to users. This bidirectional relationship indicates that the attributes of the CA (embodiment) inherently involve the use of social elements to enhance its relatability and authenticity (Tapie *et al.*, 2006).

Embodiment and mental model. The embodiment mechanism also impacts the formation of the mental model of the CA. Users interacting with a human-like CA develop expectations and perceptions that contribute to their mental models, influencing their willingness to invest effort in establishing common ground (Frijns *et al.*, 2021). As users interact with an embodied CA, they form better mental models based on the CA's human-like characteristics, which helps in setting appropriate expectations about the CA's capabilities and behaviors.

Social features and mental model. Social features play a crucial role in shaping the mental model of the CA. Social interactions involving trust, coherence, and active listening provide users with information that helps them form accurate expectations and perceptions about the CA. This unidirectional relationship means that social features influence the mental model by providing the necessary social context and behaviors that guide how users perceive and interact with the CA. Moreover, these social features can create common ground and thereby elicit an appropriate mental model of the CA. Information obtained through the CA's social behaviors aids users in forming more accurate expectations and perceptions about the CA (Kiesler, 2005; Wray *et al.*, 2022). It is important that the CA's response is appropriate to the social situation and meets the user's socially motivated expectation of building common ground.

Social features and joint action. The relationship between social features and joint action is bidirectional. Social features such as trust and coherence facilitate effective communication in joint action, enabling successful coordination and collaboration. Conversely, engaging in joint action can enhance social features through shared tasks and goals (Montemayor, 2021; Frijns *et al.*, 2021). Joint action fosters the development of social features by creating opportunities for trust-building and demonstrating coherence in achieving shared objectives. This interdependence highlights the collective nature of tasks in HAI, where integrating social elements is crucial for achieving common ground and successful joint action.

Joint action and knowledge base. Joint action directly influences the knowledge base of the CA. The information gained during joint action contributes to the development and enrichment of the knowledge base, providing the CA with valuable insights into users' knowledge, comprehension, and requirements. These collaborative interactions offer crucial input that enhances and refines the knowledge base of CAs, facilitating continuous improvement, adaptive learning, and personalized responses tailored to users' needs and preferences (Buschmeier and Kopp, 2013). This iterative process of leveraging user engagement in joint action ensures the evolution of the system, making it more proficient, accurate, and attuned to individual user requirements over time.

Joint action and mental model. Joint action extends the mental model of the CA by providing users with experiential insights into the CA's behavior and capabilities (Vinciarelli *et al.*, 2015; Wray *et al.*, 2022). Engaging in joint action allows users to form more detailed and accurate mental models, enhancing their understanding of what the CA can do and how it operates.

By integrating these mechanisms and their interrelationships, we provide a comprehensive understanding of how common ground is achieved and maintained in HAIs. This underscores the importance of considering the interconnected nature of these mechanisms to foster effective and meaningful interactions between users and conversational agents.

5. Discussion and implications

Common ground is crucial for human dialogue and essential for CAs using language. This study investigated how common ground is achieved in human-CA communication. We synthesized five mechanisms of common ground in HAI and examined their

relationships. Our contribution includes embedding these mechanisms into a conceptual model, providing a framework for studying effective HAIs that promote common ground. In the following, we present an agenda for future research, discuss theoretical and practical implications, and highlight possible study limitations.

5.1 Research agenda for future avenues

As this review and our conceptual model pave the way for research on CA interactions that establish common ground, we discuss important implications as a research agenda for *future avenues (FA)* below. We develop four FAs for common ground research in the domain of HAI that relate to the core constructs of our model. Future research endeavors should take our model and the proposed FAs into account for developing a diverse set of theory contributions to account for the theoretical richness of the common ground phenomenon in HAI. Each FA provides ground for distinct research directions for future studies (Table 2).

5.1.1 Constructs and their interplay to achieve common ground. The five mechanisms to achieve common ground are related to each other (Figure 3). Future research should focus on how these constructs jointly contribute to CA outcomes *(FA #1)*. The *Mental Model of the CA* is key, as *Embodiment, Social Features*, and *Joint Action* foster its development. Research in human-AI collaboration shows that forming mental models about AI capabilities is crucial for performance (Taudien *et al.*, 2022; Bansal *et al.*, 2019). The Mental Model of CA is the most important mechanism since apart from the "knowledge base", all other mechanisms promote the formation and alignment of mental models. Future research should explore aligning mental models in conversational interactions, considering our conceptual model's constructs. Additionally, assessing different types of embodiment, joint action, and social features can further understand their role in aligning mental models with CAs. Thus, we propose this direction for future CA research on common ground. Therefore, we propose the following direction (D):

> D1.1: Explore the role and significance of aligning mental models for achieving common ground when interacting with CAs.
>
> D1.2: Examine how Embodiment, Social Features, and Joint Action contribute to aligning mental models of users with CAs.

Moreover, the mechanisms *Embodiment* and *Social Features* maintain human-like behavior in CAs, aiding users in finding common ground. *Joint Action* requires social elements crucial for task-oriented interactions. Therefore, future research should explore how anthropomorphizing CAs contributes to common ground, possibly by developing specific mental models related to the CA. Feine *et al.* (2019) provided a starting point for experimentation on how to leverage anthropomorphic design features for achieving common ground. Thus, we propose:

> D1.3: Investigate how CAs can be designed through anthropomorphic design features that leverage common ground.

5.1.2 Evolving nature of common ground. As the user-CA relationship evolves during interactions (Skjuve *et al.*, 2022), it is crucial to examine how constructs in our model (and their interplay) change over time *(FA #2)*. Agency perceptions of novel CAs may evolve, influencing common ground (Schmitt *et al.*, 2023). Complex interactions may require re-calibrating common ground over time. For example, ChatGPT's re-calibration of common ground and capabilities was facilitated through strong signaling of versioning, a strategy previously underused by CA providers (Skantze and Doğruöz, 2023). Thus, we suggest the following directions:

> D2.1: Study how the relationship of common ground constructs evolves and contributes to common ground over time and during long-term interaction processes through longitudinal investigations.

Table 2. Agenda for common ground research in human-agent interaction

Research avenue	Directions for future CA studies
Avenue 1: constructs and their interplay to achieve common ground	*D1.1:* Explore the role and significance of aligning mental models for achieving common ground when interacting with CAs *D1.2:* Examine how Embodiment, Social Features, and Joint Action contribute to aligning mental models of users with CAs *D1.3:* Investigate how CAs can be designed through anthropomorphic design features that leverage common ground
Avenue 2: evolving nature of common ground	*D2.1:* Study how the relationship of common ground constructs evolves and contributes to common ground over time and during long-term interaction processes through longitudinal investigations *D2.2:* Analyze how re-calibration interventions such as versioning contribute to common ground during long-term interaction processes *D2.3:* Investigate how knowledge base capacity, such as context windows in LLMs, facilitates the building of common ground *D2.4:* Explore step-by-step interactions as a scaffold for human-agent interaction, e.g. with prompt engineering and chaining, when building and maintaining common ground over time
Avenue 3: domain-specificity, interdisciplinarity and boundary conditions of common ground	*D3.1:* Conduct interdisciplinary and domain-specific studies on how common ground is established, maintained, and repaired in human-agent interaction *D3.2:* Compare and contrast how common ground is achieved and utilized in different domains and contexts of human-agent interaction, and how it relates to the goals and expectations of the users and the agents while acknowledging domain-specific boundary conditions *D3.3:* Analyze the nature of common ground when interacting with open-domain CAs *D3.4:* Consider how individual user characteristics shape the process of building common ground with CAs
Avenue 4: technological embedding when investigating common ground	*D4.1:* Test the comprehensive framework of common ground in human-agent interaction with different degrees of embodiment *D4.2:* Explore the effects of in- and output modalities on building common ground with agents that are prompted by natural language inputs

Source(s): Authors' own work

D2.2: Analyze how re-calibration interventions such as versioning contribute to common ground during long-term interaction processes.

The mechanism *Knowledge Base* differs from the other mechanisms as it does not directly refer to the interaction but describes how information gained during the interaction are organized within the CA. Future research should therefore carefully take the evolving nature of the knowledge base of a CA into account when adjusting common ground during an interaction. Especially the context window of LLMs is increasing from OpenAI's GPT-3 with 2k tokens to 128k tokens within the current GPT4 versions (as of mid-2024). Thus, we propose for future research:

D2.3: Investigate how knowledge base capacity, such as context windows in LLMs, facilitates the building of common ground.

When considering a CA as a learning agent, it is important to question whether increasing context windows enhances interaction. Educational psychology principles like scaffolding and the zone of proximal development suggest that learning—and building common ground—are facilitated step-by-step (Vygotsky *et al.*, 1980). If we transfer these thoughts to CA research, instead of providing large initial information, careful prompt engineering could gradually build common ground. Chain-of-thought prompts, guiding the CA through logical steps, can enhance mutual understanding and interaction (Wei *et al.*, 2022). Breaking down complex interactions into manageable steps may improve the CA's alignment with users' mental models and foster adaptive common ground. Thus, we propose:

D2.4: Explore step-by-step interactions as a scaffold for HAI, e.g. with prompt engineering and chaining, when building and maintaining common ground over time.

5.1.3 Domain-specificity, interdisciplinarity, and boundary conditions of common ground. CA research in the incorporated studies spans various disciplines (e.g. informatics, philosophy, psychology, linguistic, IS, robotics), leading to diverse research streams due to different application contexts, which impedes generalizability. Appropriate implementation and underlying mechanisms of common ground may vary across domains. Therefore, future research should connect fundamental communication elements with interdisciplinary approaches to expand understanding of HAI *(FA #3)*. It is important to address the mechanisms of common ground explicitly and focus on domain-specific requirements for CA design in future studies. Thus, we propose:

D3.1: Conduct interdisciplinary and domain-specific studies on how common ground is established, maintained, and repaired in HAI.

Results showed that about half of the reviewed papers directly addressed common ground and it is necessary to increase this part and focus on grounding effects, supporting, or restraining variables and underlying mechanisms of common ground in HAI. For instance, customer service as a domain is oftentimes characterized by short and one-time interactions. Thus, it is necessary to build common ground very quickly, e.g. through building consensus about the CA mental model or organizing knowledge. Moreover, as common ground is oftentimes studied in educational research, CAs in digital learning environments could provide through building common ground a more productive scaffolding for learning processes to ultimately improve learning outcomes (Winkler *et al.*, 2020). Thus, we call on research that explicitly investigates common ground mechanisms in isolation and in combination with other boundary conditions.

D3.2: Compare and contrast how common ground is achieved and utilized in different domains and contexts of HAI, and how it relates to the goals and expectations of the users and the agents while acknowledging domain-specific boundary conditions.

Taking these aforementioned boundary conditions, we should also acknowledge the distinction between open-domain CAs as well as closed domain CAs. Oftentimes, CAs are developed for a single purpose with an intent-based framework and a well-defined purpose. However, when taking LLM-based CAs into account, dialogue openness is maximized while at the same time, paradoxically common ground is minimized (Skantze and Doğruöz, 2023). In practice, CAs such as ChatGPT provide for this purpose example dialogues to offer contextual knowledge about the CA capabilities, therefore, building common ground despite being an open-domain CA. Thus, we propose the following:

D3.3: Analyze the nature of common ground when interacting with open-domain CAs.

Finally, a key factor for common ground with CAs is the user's predispositions, particularly AI literacy, which can significantly enhance interaction (Knoth *et al.*, 2024). To move beyond

initial impressions and emotions surrounding CAs, it is essential to consider potential applications, the tasks the CA can perform, and areas where human skills remain indispensable. This shift in perspective aids in understanding the human role in a hybrid human-AI relationship (Baird and Maruping, 2021; Dellermann *et al.*, 2019). Maintaining such a relationship requires at least basic AI knowledge for informed decision-making aligned with personal goals (Vuorikari *et al.*, 2022). Therefore, AI Literacy is crucial for building a common ground.

D3.4: Consider how individual user characteristics shape the process of building common ground with CAs.

5.1.4 Technological embedding when investigating common ground. Furthermore, it is important to consider the fast-paced development of human-machine interaction regarding common ground *(FA #4).* Mechanisms like *Embodiment, Social Features,* and *Mental Model of CA* show the need for a holistic approach, investigating verbal elements, CA presentation, social competencies, environment, and interaction purpose. By widening our perspective of common ground, we should include human-robot interaction, the most extreme form of embodiment. While D1.3 focuses on the specific role of anthropomorphic design features in fostering common ground through human-like behaviors and characteristics, D4.1 examines a comprehensive framework across different embodiment levels and interaction contexts, aiming to validate this within various HAI scenarios, including human-robot interaction.

D4.1: Test the comprehensive framework of common ground in HAI with different degrees of embodiment.

Further, consider input and output modalities when interacting with CAs and building common ground. Speech, as an input and output modality, provides contextual cues about CA capabilities (Schmitt *et al.*, 2023), forming common ground through social features of voice. Joint action between a human and a CA are more comparable when both interact via voice, suggesting speech-based interaction has additional mechanisms for achieving common ground. Additionally, CA outputs may not always be text-based. Image generators prompted with text-input also require common ground to produce useable output, though interaction processes differ by output modality. Building on these aspects, we propose:

D4.2: Explore the effects of in- and output modalities on building common ground with agents that are prompted by natural language inputs.

5.2 Theoretical implications

Our paper also has several theoretical implications and implications for practitioners. We will first present the theoretical implications and then the practical implications. The first theoretical implication lies in our identification and synthesis of five fundamental mechanisms that facilitate common ground in HAI. These mechanisms— Embodiment, Social Features, Joint Action, Knowledge Base, and Mental Model of CA—form the theoretical backbone, presenting the state of the art of common ground in HAI, and enabling a nuanced understanding of how common ground is established and maintained. Successful grounding results in shared context and improved processes in conveying intent, emphasizing the joint adaptation of shared knowledge bases during the conversation. Interlocutors jointly modify shared representations during interaction (Brennan, 1991). Therefore, we provide an assessing overview of the relevant findings for common ground in CA research and fill the gap in understanding the mechanisms that foster common ground between humans and CAs. Moreover, Johnson and O'Connor (2008) and Bittner and Leimeister (2014) discussed how mental models contribute to common ground by offering a foundation for understanding between team members. This connection resonates with the theoretical implication that mental models play a significant role in contributing to common ground and shaping shared knowledge within interactions and collaborative efforts.

Second, by integrating these mechanisms into a conceptual model, we provide not only a theoretical foundation but also offer a structured framework. This framework contributes to existing theories by elucidating how these mechanisms interact and influence HAI dynamics. In the theoretical background, we mentioned the importance of establishing common ground in HAI, highlighting that common ground is crucial for effective communication in different modalities such as text-based, voice-based, and embodied CAs. This aligns with our paper's central theme, aiming to unravel the mechanisms of common ground within HAI across different interaction modalities.

Third, we drew connections between the mechanisms we identified and established research in HAI. This effort enriches the scholarly discourse by bridging the gap between our findings and existing theoretical frameworks, showing how our work extends and aligns with previous research. For example, linguistic theories (Giles and Powesland, 1997; Pickering and Garrod, 2004) highlight the social and collaborative aspects of human conversation. The implications of our review answer the research question of how to achieve common ground in communication between a human and a CA and add to this particular body of knowledge a nuanced understanding of how fruitful interactions between humans and CAs happen.

Fourth, our paper outlines research gaps and future research directions, emphasizing interdisciplinary and domain-specific studies. These proposed avenues extend theoretical exploration by advocating comprehensive investigations into the establishment, maintenance, and repair of common ground in diverse contexts, contributing to a more holistic understanding of HAI.

5.3 Practical implications

As a first practical implication, the identified mechanisms provide actionable guidelines for practitioners involved in CA design and development. For instance, the inclusion of anthropomorphic design features can increase the user's sense of connection and understanding with the CA. The social interaction or personality of the CA could facilitate the process of individuals understanding and forming expectations about the CA's capabilities, intentions, and behavior, for example, promoting the building of background knowledge and common ground. It is crucial to maintain human-like behavior, especially in customer service or educational contexts, to expedite the establishment of common ground.

Second, the paper distills insights into how these different grounding mechanisms can influence various HAI processes. Practitioners could use these insights from our model as a starting point to inform the design of novel CA designs. Especially the rise of LLMs and their proliferation into productive settings makes it very important for practitioners to build mental models about CA capabilities that are context-specific, and, at the same time, crucial for their performance. Understanding how different context windows in LLMs influence common ground and suggesting optimal strategies for their use in specific domains becomes imperative for successful implementation.

Third, the paper suggests an extensive agenda for future research that informs future CA research centering around common ground. In particular, encouraging practitioners to engage in design studies in authentic environments enables them to explore different grounding mechanisms. These studies include rigorous evaluations to measure how these mechanisms affect user experience and interaction outcomes. In this way, iterative improvements can be implemented to improve the quality of interaction designs for CAs. In particular, these design studies serve as valuable platforms to test and validate different grounding mechanisms in real-world contexts, facilitating the identification and resolution of challenges while capitalizing on opportunities to establish common ground in interactions with CAs.

5.4 Limitations

Despite following standard procedures, our review and analysis may have some limitations and our results should be interpreted with some caution. First, due to a broad

search strategy, the SLA tried to cover the full spectrum of application domains and foundation of the papers, rather than concentrating on a specific application context or specific empirical outcomes. However, even if the search strategy was quite broad, there may be other relevant terms that were not included in the search term specification. Furthermore, not all arXiv submissions have been reviewed (Serban *et al.*, 2017; Lee *et al.*, 2023). However, this should not have a meaningful impact on the results. Nevertheless, future work should closely monitor these sources, for which a multi-vocal review method for example could be appropriate. Second, due to our broad search strategy and our chosen type of review, we did not assess the methodological and overall quality of the articles included resulting in a less differentiated synthesis of our findings. Third, this paper builds on a cross-disciplinary, literature-based definition of common ground. Using a different understanding of common ground might lead to a different set of papers and potentially different results.

Notes

1. For further understanding, information can be found in Interactive Alignment Theory (Pickering and Garrod, 2004) and Accommodation Theory (Giles and Powesland, 1997).
2. To account for theoretical flexibility of future research, we refrained from strictly formulating theoretical propositions (see Cornelissen (2023) for the problems of the theoretical proposition grammar) that provide direct ground for testable hypotheses. Rather, future research should take this model and the presented mechanisms as a starting point for providing a rich and diverse theoretical view (Mithas *et al*., 2022; Burton-Jones *et al*., 2015) on the phenomenon of human-CA interaction and common ground.

References

Ahmad, R., Siemon, D., Gnewuch, U. and Robra-Bissantz, S. (2022), "A framework of personality cues for conversational agents", Bui, T. (Ed.), *Proceedings of the 55th Hawaii International Conference on System Sciences*.

Baird, A. and Maruping, L.M. (2021), "The next generation of research on IS use: a theoretical framework of delegation to and from agentic IS artifacts", *Management Information Systems Quarterly*, Vol. 45 No. 1, pp. 315-341, doi: 10.25300/misq/2021/15882.

Bansal, G., Nushi, B., Kamar, E., Lasecki, W.S., Weld, D.S. and Horvitz, E. (2019), "Beyond accuracy: the role of mental models in human-AI team performance", *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 7, pp. 2-11, doi: 10.1609/hcomp.v7i1.5285.

Benner, D., Elshan, E., Schöbel, S. and Janson, A. (2021), "What do you mean? A review on recovery strategies to overcome conversational breakdowns of conversational agents", *International Conference on Information Systems*.

Bernard, D. and Arnold, A. (2019), "Cognitive interaction with virtual assistants: from philosophical foundations to illustrative examples in aeronautics", *Computers in Industry*, Vol. 107, pp. 33-49, doi: 10.1016/j.compind.2019.01.010.

Bittner, E. and Leimeister, J.M. (2014), "Creating shared understanding in heterogeneous work groups: why it matters and how to achieve it", *Journal of Management Information Systems*, Vol. 31 No. 1, pp. 111-144, doi: 10.2753/mis0742-1222310106.

Blache, P. (2017), "Dialogue management in task-oriented dialogue systems", Chaminade, T. (Ed.), *Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents*, ACM, pp. 4-8.

Brennan, S.E. (1991), "Conversation with and through computers", *User Modeling and User-Adapted Interaction*, Vol. 1, pp. 67-86, doi: 10.1007/bf00158952.

Brennan, S.E. (1998), "The grounding problem in conversations with and through computers", in Fussell, S.R. and Kreuz, R.J. (Eds), *Social and Cognitive Approaches to Interpersonal Communication*, pp. 201-225.

Budiu, R. (2018), "The user experience of chatbots", available at: https://www.nngroup.com/articles/chatbots/ (accessed 16 September 2024).

Burton-Jones, A., McLean, E.R. and Monod, E. (2015), "Theoretical perspectives in IS research: from variance and process to conceptual latitude and conceptual fit", *European Journal of Information Systems*, Vol. 24 No. 6, pp. 664-679, doi: 10.1057/ejis.2014.31.

Buschmeier, H. and Kopp, S. (2013), "Co-Constructing grounded symbols—feedback and incremental adaptation in human–agent dialogue", *Künstliche Intelligenz*, Vol. 27 No. 2, pp. 137-143, doi: 10.1007/s13218-013-0241-8.

Cassell, J. (2000), "Embodied conversational interface agents", *Communications of the Association for Computing Machinery*, Vol. 43 No. 4, pp. 70-78, doi: 10.1145/332051.332075.

Chakrabarti, C. and Luger, G.F. (2015), "Artificial conversations for customer service chatter bots: architecture, algorithms, and evaluation metrics", *Expert Systems with Applications*, Vol. 42 No. 20, pp. 6878-6897, doi: 10.1016/j.eswa.2015.04.067.

Cho, J. and Rader, E. (2020), "The role of conversational grounding in supporting symbiosis between people and digital assistants", *Proceedings of the Association for Computing Machinery on Human-Computer Interaction*, pp. 1-28.

Clark, H.H. (1996), *Using Language*, Sixth printing, Cambridge University Press, Cambridge.

Clark, H.H. and Brennan, S.E. (1991), "Grounding in communication", in Resnick, L.B., Levine, J.M. and Teasley, S.D. (Eds), *Perspectives on Socially Shared Cognition*, American Psychological Association, Washington, pp. 127-149.

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V. and Cowan, B.R. (2019), "What makes a good conversation?", Brewster, S., Fitzpatrick, G., Cox, A. and Kostakos, V. (Ed.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland, ACM, pp. 1-12.

Clark, H.H. and Schaefer, E.F. (1989), "Contributing to discourse", *Cognitive Science*, Vol. 13 No. 2, pp. 259-294, doi: 10.1207/s15516709cog1302_7.

Clark, H.H. and Wilkes-Gibbs, D. (1986), "Referring as a collaborative process", *Cognition*, Vol. 22 No. 1, pp. 1-39, doi: 10.1016/0010-0277(86)90010-7.

Cornelissen, J.P. (2023), "The problem with propositions: theoretical triangulation to better explain phenomena in management research", *Academy of Management Review*. doi: 10.5465/amr.2022.0297.

Corti, K. and Gillespie, A. (2016), "Co-constructing intersubjectivity with artificial conversational agents: people are more likely to initiate repairs of misunderstandings with agents represented as human", *Computers in Human Behavior*, Vol. 58, pp. 431-442, doi: 10.1016/j.chb.2015.12.039.

Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M. (2019), "Hybrid intelligence", *Business and Information Systems Engineering*, Vol. 61 No. 5, pp. 637-643, doi: 10.1007/s12599-019-00595-2.

Di Maro, M. (2021), "Computational grounding: an overview of common ground applications in conversational agents", *Italian Journal of Computational Linguistics*, Vol. 7 No. 1, pp. 133-156, doi: 10.4000/ijcol.890.

Dillenbourg, P. (2008), "Collaborative learning: cognitive and computational approaches", *Advances in Learning and Instruction Series*, Emerald Group Pub.

Elshan, E., Zierau, N., Engel, C., Janson, A. and Leimeister, J.M. (2022), "Understanding the design elements affecting user acceptance of intelligent agents: past, present and future", *Information Systems Frontiers*, Vol. 24 No. 3, pp. 699-730, doi: 10.1007/s10796-021-10230-9.

Feine, J., Gnewuch, U., Morana, S. and Maedche, A. (2019), "A taxonomy of social cues for conversational agents", *International Journal of Human-Computer Studies*, Vol. 132, pp. 138-161, doi: 10.1016/j.ijhcs.2019.07.009.

Frijns, H.A., Schürer, O. and Koeszegi, S.T. (2021), "Communication models in human–robot interaction: an asymmetric MODel of ALterity in human–robot interaction (AMODAL-HRI)", *International Journal of Social Robotics*, Vol. 15 No. 3, pp. 473-500, doi: 10.1007/s12369-021-00785-7.

Fusaroli, R., Rączaszek-Leonardi, J. and Tylén, K. (2014), "Dialog as interpersonal synergy", *New Ideas in Psychology*, Vol. 32, pp. 147-157, doi: 10.1016/j.newideapsych.2013.03.005.

Ghose, S. and Barua, J.J. (2013), "Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor", *International Conference on Informatics, Electronics and Vision*, pp. 1-5, doi: 10.1109/iciev.2013.6572650.

Giles, H. and Powesland, P. (1997), "Accommodation theory", in Coupland, N. and Jaworski, A. (Eds), *Sociolinguistics: A Reader and Coursebook, Modern Linguistics Series*, Macmillan, Basingstoke, Hampshire, pp. 232-239.

Gnewuch, U., Morana, S. and Maedche, A. (2017), "Towards designing cooperative and social conversational agents for customer service", *Proceedings of the International Conference on Information Systems 2017*.

Hanna, N. and Richards, D. (2014), "Measuring the effect of personality on human-IVA shared understanding", Bazzan, A. (Ed.), *Proceedings of the 2014 International Conference on Autonomous Agents & Multiagent Systems, ACM Digital Library*, International Foundation for Autonomous Agents and Multiagent Systems, Richland County, South Carolina, pp. 1643-1644.

Hobert, S. and Meyer von Wolff, R. (2019), "Say hello to your new automated tutor – a structured literature review on pedagogical conversational agents", *International Conference on Wirtschaftsinformatik 2019*.

Janson, A. (2023), "How to leverage anthropomorphism for chatbot service interfaces: the interplay of communication style and personification", *Computers in Human Behavior*, Vol. 149, 107954, doi: 10.1016/j.chb.2023.107954.

Johnson, T.E. and O'Connor, D.L. (2008), "Measuring team shared understanding using the analysis-constructed shared mental model methodology", *Performance Improvement Quarterly*, Vol. 21 No. 3, pp. 113-134, 10.1002/piq.20034.

Kanda, T., Hirano, T., Eaton, D. and Ishiguro, H. (2004), "Interactive robots as social partners and peer tutors for children: a field trial", *Human-Computer Interaction*, Vol. 19 No. 1, pp. 61-84, doi: 10.1207/s15327051hci1901&2_4.

Kiesler, S. (2005), "Fostering common ground in human-robot interaction", *Robot and Human Interactive Communication*, pp. 729-734, doi: 10.1109/roman.2005.1513866.

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G. and Welch, G.F. (2018), "Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR", *International Symposium on Mixed and Augmented Reality (ISMAR 2018)*, pp. 105-114, doi: 10.1109/ismar.2018.00039.

Knote, R., Janson, A., Söllner, M. and Leimeister, J.M. (2021), "Value Co-creation in smart services: a functional affordances perspective on smart personal assistants", *Journal of the Association for Information Systems*, Vol. 22 No. 2, pp. 418-458, doi: 10.17705/1jais.00667.

Knoth, N., Tolzin, A., Janson, A. and Leimeister, J.M. (2024), "AI literacy and its implications for prompt engineering strategies", *Computers and Education: Artificial Intelligence*, Vol. 6, 100225.

Kontogiorgos, D. (2022), "Mutual understanding in situated interactions with conversational user interfaces: theory, studies, and computation", available at: https://api.semanticscholar.org/CorpusID:248556993 (accessed 16 September 2024).

Koulouri, T., Lauria, S. and Macredie, R.D. (2016), "Do (and say) as I say: linguistic adaptation in human–computer dialogs", *Human-Computer Interaction*, Vol. 31 No. 1, pp. 59-95, doi: 10.1080/07370024.2014.934180.

Lee, S., Lau, I.Y., Kiesler, S. and Chiu, C.-Y. (2005), "Human mental models of humanoid robots", *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 2767-2772.

Lee, M., Srivastava, M., Hardy, A., Thickstun, J., Durmus, E., Paranjape, A. and Gerard-Ursin, I. (2023), "Evaluating human-language model interaction", available at: http//doi.org/10.48550/arXiv.2212.09746 (accessed 8 August 2023)

Leidner, D. (2018), "Review and theory symbiosis: an introspective retrospective", *Journal of the Association for Information Systems*, Vol. 19 No. 06, pp. 552-567, doi: 10.17705/1jais.00501.

Lorini, E. (2020), "Rethinking epistemic logic with belief bases", *Artificial Intelligence*, Vol. 282, 103233, doi: 10.1016/j.artint.2020.103233.

Luger, E. and Sellen, A. (2016), "Like having a really bad PA", Kaye, J. (Ed.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, pp. 5286-5297.

McKeown, G. (2015), "Turing's menagerie: talking lions, virtual bats, electric sheep and analogical peacocks: common ground and common interest are necessary components of engagement", *International Conference on Affective Computing and Intelligent Interaction*, pp. 950-955, doi: 10.1109/acii.2015.7344689.

Meredith, J. (2017), "Analysing technological affordances of online interactions using conversation analysis", *Journal of Pragmatics*, Vol. 115, pp. 42-55, doi: 10.1016/j.pragma.2017.03.001.

Mithas, S., Xue, L., Huang, N. and Burton-Jones, A. (2022), "Editor's comments: causality meets diversity in information systems research", *Management Information Systems Quarterly*, Vol. 46 No. 3, pp. 3-18.

Mohammed, S., Klimoski, R. and Rentsch, J.R. (2000), "The measurement of team mental models: we have No shared schema", *Organizational Research Methods*, Vol. 3 No. 2, pp. 123-165, doi: 10.1177/109442810032001.

Montemayor, C. (2021), "Language and intelligence", *Minds and Machines*, Vol. 31 No. 4, pp. 471-486, doi: 10.1007/s11023-021-09568-5.

Murphy, M. (2019), "This app is trying to replicate you", available at: https://qz.com/1698337/replikathis-app-is-trying-to-replicate-you/ (accessed 25 June 2023).

Nass, C. and Moon, Y. (2000), "Machines and mindlessness: social responses to computers", *Journal of Social Issues*, Vol. 56 No. 1, pp. 81-103, doi: 10.1111/0022-4537.00153.

Neururer, M., Schlögl, S., Brinkschulte, L. and Groth, A. (2018), "Perceptions on authenticity in chat bots", *Multimodal Technologies and Interaction*, Vol. 2 No. 3, p. 60, doi: 10.3390/mti2030060.

Paré, G., Trudel, M.-C., Jaana, M. and Kitsiou, S. (2015), "Synthesizing information systems knowledge: a typology of literature reviews", *Information and Management*, Vol. 52 No. 2, pp. 183-199, doi: 10.1016/j.im.2014.08.008.

Pickering, M.J. and Garrod, S. (2004), "Toward a mechanistic psychology of dialogue", *The Behavioral and Brain Sciences*, Vol. 27 No. 2, pp. 169-190, doi: 10.1017/s0140525x04000056.

Popay, J., Roberts, H.M., Sowden, A.J., Petticrew, M., Arai, L., Rogers, M. and Britten, N. (2006), "Guidance on the conduct of narrative synthesis in systematic reviews", *A Product from the ESRC Methods Programme*, doi: 10.13140/2.1.1018.4643.

Pustejovsky, J. and Krishnaswamy, N. (2021), "Embodied human computer interaction", *Künstliche Intelligenz*, Vol. 35 No. 3, pp. 307-327, doi: 10.1007/s13218-021-00727-5.

Rowe, F. (2014), "What literature review is not: diversity, boundaries and recommendations", *European Journal of Information Systems*, Vol. 23 No. 3, pp. 241-255, doi: 10.1057/ejis.2014.7.

Schmitt, A., Zierau, N., Janson, A. and Leimeister, J.M. (2023), "The role of AI-based artifacts' voice capabilities for agency attribution", *Journal of the Association for Information Systems*, Vol. 24 No. 4, pp. 980-1004, doi: 10.17705/1jais.00827.

Schöbel, S., Schmitt, A., Benner, D., Saqr, M., Janson, A. and Leimeister, J.M. (2024), "Charting the evolution and future of conversational agents: a research agenda along five waves and new frontiers", *Information Systems Frontiers a Journal of Research and Innovation*, Vol. 26 No. 2, pp. 729-754, doi: 10.1007/s10796-023-10375-9.

Seeber, I., Bittner, E., Briggs, R.O., Vreede, T.D., Vreede, G.-J.D., Elkins, A., Maier, R., Merz, A.B., Oeste-Reiß, S., Randrup, N., Schwabe, G. and Söllner, M. (2020a), "Machines as teammates: a research agenda on AI in team collaboration", *Information and Management*, Vol. 57 No. 2, 103174, doi: 10.1016/j.im.2019.103174.

Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I. and Lowry, P.B. (2020b), "Collaborating with technology-based autonomous agents", *Internet Research*, Vol. 30 No. 1, pp. 1-18, doi: 10.1108/intr-12-2019-0503.

Serban, I.V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., Kim, T., Pieper, M., Chandar, S., Ke, N.R., Rajeshwar, S., Brebisson, A.D., Sotelo, J.M.R., Suhubdy, D., Michalski, V., Nguyen, A., Pineau, J. and Bengio, Y. (2017), "A deep reinforcement learning chatbot", available at: http//doi.org/10.48550/arXiv.1709.02349 (accessed 10 June 2022)

Skantze, G. and Doğruöz, A.S. (2023), "The open-domain paradox for chatbots: common ground as the basis for human-like dialogue", Stoyanchev, S., Joty, S., Schlangen, D., Dusek, O., Kennington, C. and Alikhani, M. (Ed.), *Proceedings of the 24th Meeting of the Special Interest Group on Discourse and Dialogue*, Association for Computational Linguistics, Stroudsburg, PA, pp. 605-614.

Skjuve, M., Følstad, A., Fostervold, K.I. and Brandtzaeg, P.B. (2022), "A longitudinal study of human–chatbot relationships", *International Journal of Human-Computer Studies*, Vol. 168, 102903, doi: 10.1016/j.ijhcs.2022.102903.

Stalnaker, R. (2002), "Common ground", *Linguistics and Philosophy*, Vol. 25 No. 5, pp. 701-721, doi: 10.1023/a:1020867916902.

Strohmann, T., Siemon, D., Khosrawi-Rad, B. and Robra-Bissantz, S. (2023), "Toward a design theory for virtual companionship", *Human-Computer Interaction*, Vol. 38 No. 3, pp. 194-234, doi: 10.1080/07370024.2022.2084620.

Sundar, S.S., Qian, X. and Bellur, S. (2010), "Designing interactivity in media interfaces: a communications perspective", *CHI 2010: Perspectives on Design*, pp. 2247-2256.

Tapie, J., Terrier, P., Perron, L. and Cellier, J.-M. (2006), "Should remote collaborators be represented by avatars? A matter of common ground for collective medical decision-making", *AI and Society*, Vol. 20 No. 3, pp. 331-350, doi: 10.1007/s00146-005-0028-z.

Taudien, A., Fuegener, A., Gupta, A. and Ketter, W. (2022), "Calibrating users' mental models for delegation to AI", *International Conference on Information Systems 2022*.

Tolzin, A. and Janson, A. (2023), "Mechanisms of common ground in human-agent interaction: a systematic review of conversational agent research", *Proceedings of the Hawaii International Conference on System Sciences*, Maui, Hawaii.

Vassallo, G., Pilato, G., Augello, A. and Gaglio, S. (2010), "Phase coherence in conceptual spaces for conversational agents", in Sheu, P.C.-Y. (Ed.), *Semantic Computing*, IEEE Press; John Wiley & Sons, Piscataway, NJ, Hoboken, NJ, pp. 357-371.

Vinciarelli, A., Esposito, A., André, E., Bonin, F., Chetouani, M., Cohn, J.F., Cristani, M., Fuhrmann, F., Gilmartin, E., Hammal, Z., Heylen, D., Kaiser, R., Koutsombogera, M., Potamianos, A., Renals, S., Riccardi, G. and Salah, A.A. (2015), "Open challenges in modelling, analysis and synthesis of human behaviour in human–human and human–machine interactions", *Cognitive Computation*, Vol. 7 No. 4, pp. 397-413, doi: 10.1007/s12559-015-9326-z.

Vuorikari, R., Kluzer, S. and Punie, Y. (2022), *DigComp 2.2, the Digital Competence Framework for Citizens: with New Examples of Knowledge, Skills and Attitudes*, Publications Office European Union, Luxembourg.

Vygotsky, L.S., Cole, M., Jolm-Steiner, V., Scribner, S. and Souberman, E. (1980), *Mind in Society*, Harvard University Press, New York.

Webster, J. and Watson, R.T. (2002), "Analyzing the past to prepare for the future: writing a literature review", *Management Information Systems Quarterly*, Vol. 26, pp. 13-23.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. and Zhou, D. (2022), "Chain-of-Thought prompting elicits reasoning in large language models", available at: https://doi.org/10.48550/arXiv.2201.11903 (accessed 10 June 2022).

Weiler, S., Matt, C. and Hess, T. (2022), "Immunizing with information – inoculation messages against conversational agents' response failures", *Electronic Markets*, Vol. 32 No. 1, pp. 239-258, doi: 10.1007/s12525-021-00509-9.

Winkler, R., Hobert, S., Salovaara, A., Söllner, M. and Leimeister, J.M. (2020), "Sara, the lecturer: improving learning in online education with a scaffolding-based conversational agent", Bernhaupt, R., Mueller, F., Verweij, D., Andres, J., McGrenere, J., Cockburn, A., Avellino, I., Goguey, A., Bjørn, P., Zhao, S., Samson, B.P. and Kocielnik, R. (Ed.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, pp. 1-14.

Wittgenstein, L. (1967), *Philosophical Investigations*, Blackwell, Oxford.

Wray, R., Kirk, J. and Folsom-Kovarik, J. (2022), "Improving common ground in human-machine teaming: dimensions, gaps, and priorities", *Artificial Intelligence and Social Computing*, AHFE International.

Zhou, S., Loiacono, E.T. and Kordzadeh, N. (2023), "Smart cities for people with disabilities: a systematic literature review and future research directions", *European Journal of Information Systems*, Vol. 33 No. 6, pp. 1-18, doi: 10.1080/0960085x.2023.2297974.

Zierau, N., Wambsganss, T., Janson, A., Schöbel, S. and Leimeister, J.M. (2020), "The anatomy of user experience with conversational agents: a taxonomy and propositions of service clues", *International Conference on Information Systems*.

Zierau, N., Hildebrand, C., Bergner, A., Busquet, F., Schmitt, A. and Leimeister, J.M. (2023), "Voice bots on the frontline: voice-based interfaces enhance flow-like consumer experiences & boost service outcomes", *Journal of the Academy of Marketing Science*, Vol. 51 No. 4, pp. 823-842, doi: 10.1007/s11747-022-00868-5.

Supplementary material
The supplementary material for this article can be found online.

Corresponding author
Antonia Tolzin can be contacted at: antonia.tolzin@uni-kassel.de