Does Social Presence Increase Perceived Competence? Evaluating Conversational Agents in Advice Giving Through a Video-Based Survey

Schmid, Damaris ; Staehelin, Dario ; Bucher, Andreas ; Dolata, Mateusz ; Schwabe, Gerhard

Abstract: Conversational agents (CA) have drawn increasing interest from HCI research. They have become popular in different aspects of our lives, for example, in the form of chatbots as the primary point of contact when interacting with an insurance company online. Additionally, CA find their way into collaborative settings in education, at work, or financial advisory. Researchers and practitioners are searching for ways to enhance the customer's experience in service encounters by deploying CA. Since competence is an important treat of a financial advisor, they only accept CA in their interaction with clients if it does not harm their impression on the client. However, we do not know how the social presence of the CA affects this perceived competence. We explore this by evaluating three prototypes with different social presences. For this, we conducted a video-based online survey. In contrast to prior studies focusing on single human-computer interaction, our study explores CA in a dyadic setting of two humans and one CA. First, our results support the Computers-Are-Social-Actors paradigm as the CA with a strong social presence was perceived as more competent than the other two designs. Second, our data show a positive correlation between CA's and advisor's competence. This implies a positive impact of the CA on the service encounter as the CA and advisor can be seen as a competent team.

DOI: https://doi.org/10.1145/3492845

Does Social Presence Increase Perceived Competence? Evaluating Conversational Agents in Advice Giving Through a Video-Based Survey

DAMARIS SCHMID, University of Zurich, Switzerland
DARIO STAEHELIN, University of Zurich, Switzerland
ANDREAS BUCHER, University of Zurich, Switzerland
MATEUSZ DOLATA, University of Zurich, Switzerland
GERHARD SCHWABE, University of Zurich, Switzerland

Conversational agents (CA) have drawn increasing interest from HCI research. They have become popular in different aspects of our lives, for example, in the form of chatbots as the primary point of contact when interacting with an insurance company online. Additionally, CA find their way into collaborative settings in education, at work, or financial advisory. Practitioners and researchers are searching for ways to improve the customer's experience in service encounters by deploying CA. Since competence is an important treat of a financial advisor, they only accept CA in their interaction with clients if it does not harm their impression on the client. However, we do not know how the social presence of the CA affects this perceived competence. We explore this by evaluating three prototypes with different social presences. For this, we conducted a video-based online survey. In contrast to prior studies focusing on single human-computer interaction, our study explores CA in a dyadic setting of a CA and two humans. First, our results support the Computers-Are-Social-Actors paradigm as the CA with a strong social presence was perceived as more competent than the other two designs. Second, our data show a positive correlation between CA's and advisor's competence. This implies a positive impact of the CA on the service encounter as the CA and advisor can be seen as a competent team.

Additional Key Words and Phrases: Conversational Agent, Social Presence, Advisory Service Encounter, Human-AI Collaboration

ACM Reference Format:
Damaris Schmid, Dario Staehelin, Andreas Bucher, Mateusz Dolata, and Gerhard Schwabe. 20XX. Does Social Presence Increase Perceived Competence? Evaluating Conversational Agents in Advice Giving Through a Video-Based Survey. *Proc. ACM Hum.-Comput. Interact.* X, GROUP, Article XX (X 20XX), 22 pages. https://doi.org/XXX

1 INTRODUCTION

Artificial Intelligence-driven systems are becoming more widespread, especially with the establishment of speech-based personal agents such as Apple's Siri, Amazon's Alexa, or Alibaba's AliGenie. Due to its success in the private sector and the need for innovative solutions in business, intelligent

Authors' addresses: Damaris Schmid, damaris.schmid2@uzh.ch, University of Zurich, Binzmühlestrasse 14, Zurich, Switzerland, 8050; Dario Staehelin, University of Zurich, Binzmühlestrasse 14, Zurich, Switzerland, 8050, staehelin@ifi.uzh.ch; Andreas Bucher, University of Zurich, Binzmühlestrasse 14, Zurich, Switzerland, 8050, bucher@ifi.uzh.ch; Mateusz Dolata, University of Zurich, Binzmühlestrasse 14, Zurich, Switzerland, 8050, dolata@ifi.uzh.ch; Gerhard Schwabe, University of Zurich, Binzmühlestrasse 14, Zurich, Switzerland, 8050, schwabe@ifi.uzh.ch.

conversational agents (CA) find their way into the professional setting. Especially in times, where the world's industry is shifting more and more toward outcome-based business models, services are offered instead of classical products [62]. However, traditional service industries can equally benefit from such transformative innovations. For example, a face-to-face service encounter is a business format that can profit from introducing CA and compelling improvements they promise.

Financial advice giving is a highly fragile form of service encounters that requires careful considerations in the design of a CA. While recent research in HCI has investigated the potential of CA in the various sectors and generated valuable insights for design and application [3, 6, 27], we lack knowledge on how to design CA in multi human-CA interaction in service advisory. Research in this area, such as Seeber *et al.* [53] and Dolata *et al.* [11] is still scarce, and its contributions remain of theoretical nature. Design guidelines for such collaborations are of great importance as the dynamics in these settings is very different from a single human-CA interaction. Prior research has shown that IT, even in the form of a tablet, introduced in financial service encounters alters its fundamental structure from a dyadic to a triadic collaboration. These findings are in accordance with the Computers-Are-Social-Actors (CASA) paradigm that states that humans tend to interact with artificial systems as if they were interacting with other human-beings [42].

Furthermore, a CA that interacts to some degree autonomously might potentially affect the perception of the advisor as trustworthy and competent. These are two characteristics that are of most importance for financial advisors. They fear losing their face in front of the client if a CA is perceived as a competent actor in the service encounter. Consequently, researchers and designers of CA need to be careful regarding CA's humanness or social presence and its effect on the perceived competence. This study investigates the introduction of CA in service encounters and addresses the following research question: *Does the social presence of CA increase their perceived competence?*

The study is part of a larger research project that aims at developing a CA for financial service encounters in Switzerland. The requirements exploration phase, which took place prior to this study, yielded three prototypes of the CA diverging in their social presence. The design of this three prototypes based on exploratory interviews with advisors, workshops, and contextual inquiries as well as design principles from research. The prototypes were evaluated in a video-based online survey by potential bank clients. In this paper, we present three insights from the quantitative analysis of the survey. First, it confirms that the positive impact of a strong social presence on the perceived competence of the CA also holds in a triadic setting. Second, the study finds no negative impact of any level of the CA's social presence on the advisor. Third, it positively links the perceived competence of the CA with that of the advisor. These insights are discussed considering existing literature about social presence of CA and in the light of the "machines as teammates" movement. Lastly, we formulate practical implications derived from our results such as the advice to obey conversational rituals and put a focus on the coherence and adjustability of the CA.

2 RELATED WORK

We structure our related work section around the concepts of IT in advisory service and the question of how a CA is applied to advisory as competent collaborators. These two perspectives provide a rich background for the analysis of our data.

2.1 IT in Advisory Sessions

The topic of advisory service support attracted the attention of HCI and CSCW research, as well as beyond it. Supporting the participants is more important than ever because the advisor's everyday work has become more burdensome and complex. While the mass of regulations, especially in the financial sector, raises the need for more standardization, clients require increasing individuality of encounters and tailored products. IT found its way into advisory, promising more efficient advisors and more satisfied clients. Multiple attempts to support participants in such settings have been undertaken in different industries, such as: city councils [20], medical counseling [48], travel agencies [52], home security counseling [14], or financial institutions [24, 47].

However, while the research has shown that the integration of IT, such as PCs [24] or touch-enabled single display groupware [24] can help the advisors motivating clients [8] or inciting better overall perception of the service [10], the uptake of those systems is slow. On the one side, although IT may enhance advisory regarding the transparency [7, 46], documentation [18], persuasion [8], or data access [16], it creates physical and structural constraints due to static seating positions and stationary screens [9]. Furthermore, the integration of IT may bother the relationship and communication between the advisor and the client [22]. Especially in advisory service encounters, the interaction between advisor and client requires a high level of intimacy and sensitivity as personal or complex topics are frequently discussed. Therefore, the integration of IT in the service encounter should not interfere or interrupt the natural flow of the conversation.

Various studies not only examined the seating positions of the participants [22] but also contemplated the design of the involved IT system such as touch-based large displays [45], non-planar interfaces [23, 24], and pen-and-paper interfaces which unite the digital interaction and paper [37, 58, 59]. Nevertheless, as these systems still rely on proactive control and haptic input, the advisors still had to focus on operating the systems [28, 29, 52]. Despite the careful design of such IT systems, their introduction into dyadic settings, like advisory service encounters, is therefore still perceived as a disruption of the natural flow of conversation and might even impair the participants' self-disclosure [28, 38]. Given the limitations that previous research on IT-supported advisory sessions faced, we believe that CA can be well embedded into the natural flow of conversation as they potentially are able to understand the context of the conversation and thus provide a contextually relevant and correct input.

Following the research of Kilic *et al.* [28] and Dolata *et al.* [11], we not only assume that a CA can naturally fit into the conversational flow, but we also believe that a CA changes the conversation by extending the dyad between advisor and client to a triadic collaboration between advisor, client, and CA. In the conventional dyadic format, advisory service encounters are often seen as instances of institutional talk, in which the conversation partners act according to their institutional identity. Acting on their institutional identity, bank advisors adapt and control their vocabulary, their turn-taking behavior, or other conversational patterns, such as the distribution of physical space, in accordance with their identity of the host and subject matter expert [11]. Integrating a CA as a third actor would thus require considerations regarding the specific role and identity of the CA as well as a reformulation of the frames of institutional talk. An exemplary question that might arise in such context is whether a CA should act on behalf of the institution, *e.g.,* the bank, or take on the identity of an independent interlocutor. However, as the research in this area is still very nascent, it remains open what implications arise for the dynamics of the conversation when introducing a CA into the dyadic context of the advisory service encounter.

2.2 Conversational Agents as Competent Collaborator

Starting in 1966, when ELIZA, the first chatbot, was developed [63], research and economy have continuously discovered its potential. CA are dialogue systems that simulate human conversations [36]. The usage of CA reveals new opportunities, such as interactively handling personalized customer requests, that lead to an engaging customer experience [6]. Nowadays, such speech-based personal agents are no longer only innovations for private homes or smartphones but have also found their way into professional settings. Therein, CA, including text-based chatbots, are deployed in various industries ranging from finance and insurance to education [3]. While there is literature in HCI about the use of generic chatbots in the financial industry [27], the opportunities of speech-based CA in advisory service encounters are still little explored.

A sub-movement in HCI research is human-centered AI (HCAI) [55]. It aims at creating and implementing AI systems that enhance and empower humans rather than replacing them. HCAI researchers express the opinion that computers should support humans and masterfully augment people's performance at work, following the motto "Humans in the group; computers in the loop" [55, p. 58]. By adding another perspective, there is evidence that machines can act as teammates if the tasks of humans and machines are distributed according to their strengths [53]. For instance, in financial advisory service, a CA could enrich the service encounter by providing additional information about investment products that are either too recent or beyond the advisor's knowledge. Therefore, we conclude that CA could be teammates with specific competencies adjudged.

Another concept adds to the scarce discussion about human-machine teaming and addresses the issue of limited interaction in terms of scope and CA capabilities. If CA mimic human communication behavior [36], they must provide a rich set of social cues. Therefore, we consider them as social actors according to the CASA paradigm [42]. It was established by Nass *et al.* [49] and was derived from the media equation. CASA strongly relies on the observation that humans mindlessly apply social scripts related to human-human interaction when interacting with computers. When the CASA paradigm was established around the turn of the millennium, the interaction with computers was not as mature and proliferated as it is nowadays. Hence, social scripts and mental models for interacting with other people [25, 26], mainly existed for human-human interaction. Since then, technological abilities, users, and human-computer interaction have changed significantly. Therefore, it is reasonable to believe that people have extended their social scripts to the specific interaction with computers [17].

One design element of a CA that can contribute to the exposition of the CASA paradigm [30] and is related to the social presence of a CA is anthropomorphism. In this sense, social presence is the ability of a CA or any other communication medium or technology to create the feeling of "being with another" by transporting human warmth, sociability, or telepresence [4, 56]. Therefore, social presence makes a statement about the extent to which people perceive that they are interacting with another social actor [21]. Therein, specific criteria should be fulfilled such that a machine is perceived as a social actor according to CASA.

Nass *et al.* [42] described two important criteria for the CASA paradigm to be applied to human-computer interaction. First, the machine must have enough cues "to lead the person to categorize it as worthy of social responses" [42, p. 83]. Unfortunately, it is neither defined what kind of social cues are required nor what does *enough* mean exactly [17]. However, a limited set of verbal, visual, auditory, and invisible cues [15] still provides sufficient evidence for a human to apply social behavior towards a machine [5]. Secondly, the computer needs to be a source of information rather than just the transmitter of data [43]. Therefore, a CA which is used in the context of this research should, for example, gather, aggregate, analyze, or enrich existing data to be an additional source of information to enrich the advisory service encounter. Existing research on CA is primarily focused on one-on-one communication between a single human user and a single CA. Thus, only little is known about the considerations that arise when integrating a CA in a complex multi-user environment within an organizational context, such as financial advisory service.

The broad conception of social presence often includes that a socially more present and animate agent is perceived as more intelligent and competent [2]. Van Doorn *et al.* [61] proclaimed that a highly automated social presence should result in a stronger perception of competence when researching the interplay between anthropomorphism and social cognition. However, highly anthropomorphized machines may also "fall" into the uncanny valley: when a CA becomes more human-like, people first show a positive emotional response; however, when it resembles a human too much, users experience eery feelings [39]. Similar to the eeriness, the uncanny valley effect may occur for competence [40]. It remains open if this effect is transferable into triadic settings with two humans and involved CA as actors.

Several studies already contributed to the discussion about the perception of intelligence of AI-supported robots [30, 32, 57]. Other research investigates the correlation of animacy and intelligence using competence as a sub-measure for intelligence [2]. Kim *et al.* [32] suggest that an increase in anthropomorphism does not affect the level of perceived competence. Contrarily, comparing images with different anthropomorphism values, users assigned higher competence scores to more human-like facial representations of computers [20]. Competence can be measured by evaluating the ability to act as an expert [30]. The amount of an agent's perceived competence is influenced by different factors such as animacy, which can be considered a synonym for social presence [2]. Furthermore, compliance, responsiveness, effort transparency, and robustness [35] affect the perceived competence as well. Nevertheless, we still lack knowledge about the perceived competence of a social actor with a speech-based interaction interface in a collaborative setting. For this reason, we derive the first research question as follows:

RQ 1 Does social presence increase the perceived competence of the CA?

Closely related to the question about the influence of social presence on the perception of a CA, like its perceived competence, is the impact of social presence on the perception of the human users. Similar to human-human collaboration, one can assume that an artificial teammate influences the behavior and perception of the group or the individual team members. For instance, Kim *et al.* [31] observed that CA can facilitate group behavior by encouraging members to participate more actively in group discussions. Moreover, Schuetzler *et al.* [51] found that an increase of a CA's human-like traits may increase specific deception strategies of humans, making it more challenging to detect untruthful answers. Nonetheless, to the best of our knowledge, only little is known about the effect on the perception of human team members when introducing a CA into a dyadic setting, such as advisory service encounters. Therefore, we formulate the second research question as follows:

RQ 2 Does social presence increase the perceived competence of the advisor?

Advisory service is a generic example of the principal-agent issue and the associated missing transparency [19]. Existing research about CA as social actors in principal-agent settings [14] is underexplored as well. The conventional dyadic setting of a financial advisory, including the advisor and the client, is extended with a third party, the CA. Since there is no literature, this question about the correlation of the perceived competence of the CA and the human advisor has also not been answered yet. This leads us to derive the third research question:

RQ 3 Does the perceived competence of the CA correlate with the perceived competence of the advisor?

3 METHODOLOGY

The goal of the study is to examine if human qualities such as competence can be attributed to machines and how the social presence influences the perception of these qualities. We conducted a video-based online survey with three different CA to answer the previously stated research questions. A video-based online survey was favored for this study due to multiple reasons. We can achieve a higher standardization of the advisory sessions, and therefore, the results are better comparable. The videos also allow for increased control of the experiment environment. Another advantage of video-based online surveys is the lower cost per participant, and thus, an increased sample size having the same budget. Finally, the expectations regarding the technical maturity of the CA are lower compared to an on-site test such as a Wizard-of-Oz test concept.

3.1 Setup

This study is part of a Design Science Research (DSR) project, including two universities, two regional Swiss banks, and two technology partners. The goal of the project is to develop a CA for bank advisory services. We conducted this study at the end of the requirements exploration and analysis phase of the project, which aimed at defining a use case for the CA and developing prototypes to evaluate design choices. Between June 2020 and October 2020, the project team conducted ten exploratory interviews with advisors, six shadowings of advisors' daily routines, and three workshops with stakeholders to better understand the research data through familiarization with the research context. In addition, the researchers carried out two systematic literature reviews focusing on CA in collaborative settings and the service sector. The client's perspective is at the heart of this study in order to achieve our primary goal of understanding the influence of a CA's social presence on the perceived competence of the CA itself and the advisor.

To explore this interrelation, the project team developed three visions of CA with different levels of social presence. The visual parts such as the projections on the table of these prototypes, were fully designed and implemented. However, the capability of understanding the conversation thus, the intelligence, and the voice if there is any, was realized for this study. Therefore, all experiments recorded in the videos, were implemented in a Wizard-of-Oz setting. The CA were presented as having the same capabilities in terms of creating meeting notes, searching and displaying information (*e.g.,* chart of Apple stocks), and adapting visualizations according to the users' wishes (*e.g.,* exchanging Apple stocks with Microsoft in the displayed portfolio). Similarly, all three CA were presented as being able to understand the general context of the conversation to the point where they could present relevant output data. These capabilities were identified as desired in the requirements exploration phase of the project, before the survey (*i.e.,* exploratory interviews with advisors, workshops, and contextual inquiries). We used those initial findings to create the prototypes for the survey.

While not differing in their capabilities, the CA were designed with three different levels of social presence. In line with Feine *et al.* [15], the project team included a wider variety of social cues for the CA design. The different levels of social presence were derived from relevant literature [2, 15, 61] and requirements gathered in our exploration as the design aimed at being literature-grounded and context-adapted. Many experienced advisors expressed the wish for an assistant who is able to perform different kinds of supporting tasks but without the high cost of a human assistant. Therefore, this advisor-CA pair might be perceived as two social actors forming a counseling team. The three prototypes diverged not only in their visual representation but also in their communication behavior (*i.e.,* verbal, visual-auditory, invisible). For the sake of simplicity, we use three names (*Box*, *Smiley*, and *Pius*) for the different CA as described in the following:

The least socially present version was represented by a simple speaker (in the following referred to as *Box*, top right in Figure 1). The *Box* did not receive a name or voice output to avoid humanization. It was designed with output capabilities limited to visualizations on an output device (*e.g.,* touchscreen). The *Box* received inputs via voice commands and haptic information (for error handling). It cannot process natural language and only interacts when triggered by specific commands. The advisor had to address the agent, for example, by saying "Digital agent, display the Apple stock", similar to the *Hey Google*-command of Google's intelligent assistant.

The second prototype was still unnamed and represented on the screen as a smiley icon to resemble a more socially present CA (in the following referred to as *Smiley*, bottom left in Figure 1). *Smiley* would not be displayed permanently but appear when called upon and disappear after the end of the interaction. It also had slight facial expressions and features. In terms of input processing, *Smiley* did not differ from *Box* and would only interact on specific commands. In contrast to *Box*, it provided additional speech-based output besides the visualizations. The CA spoke German and answered to commands in a confirmatory manner (*e.g.*, "done"). It was not capable of providing spoken answers to a question.

The CA with the highest social presence is called *Pius* (bottom right in Figure 1). The design aimed at creating a full-featured conversation partner. For this, *Pius* had a name and human-like representation on a separate display. The avatar is a male with brown hair and wearing a shirt, a tie, and jeans. *Pius* was able of gestures such as waving and giving a thumbs up. Additionally, it did not need an explicit command to interact with the two human actors. Compared to *Smiley*, *Pius* provided spoken output in Swiss dialect not limited to confirmatory phrases. For example, it could provide additional spoken input on a visual representation that it showed to the client (*e.g.*, displaying a risk profile on the touchscreen and orally explain the differences between profiles). It was able to interrupt the conversation if necessary and to show emotions. For example, when it had to intervene in the discussion, it apologized and showed a sorrowful facial expression.

Table 1 presents an overview of the CA and their social cues referring to the taxonomy of Feine *et al.* [15].

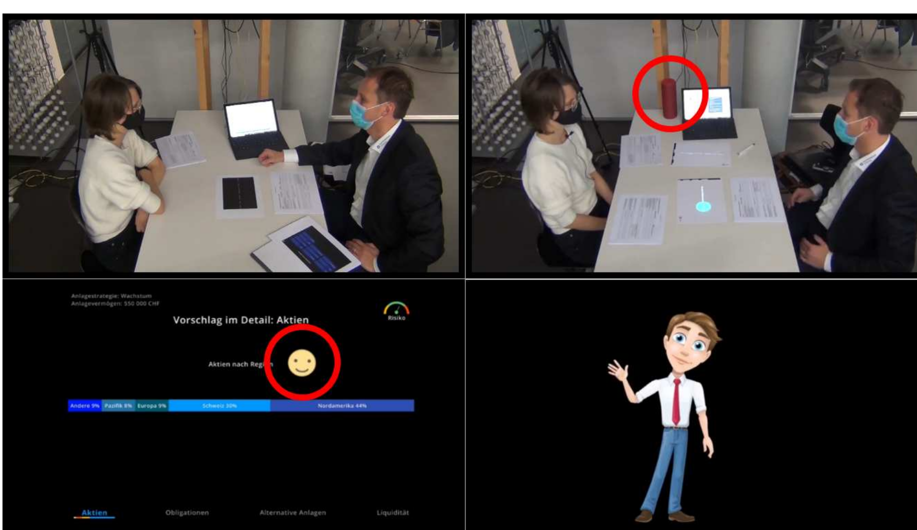

Fig. 1. Video keyframes of conventional advisory (t.l.), the *Box* (t.r.), the *Smiley* (b.l.), and *Pius* (b.r.)

Independently of their social presence, all versions recorded, gathered, and dynamically manipulated data. Only *Pius* explained facts and circumstances or introduced himself at the beginning of the advisory session. All three agents showed warnings during the portfolio elaboration when the modifications led to an extraordinarily risky asset allocation. However, *Box* only displayed a red border while *Smiley* and *Pius* both changed their facial expression. *Pius* even interrupted the ongoing conversation to call attention to the risk-shifting. The following section provides the reader a detailed description of the phases and actions performed in the encounter recordings.

			Box	***Smiley***	***Pius***
Verbal	**Content**	Greetings			x
		Farewells		x	x
		Excuse			x
		Tips and advice			x
Visual	**Kinesics**	Arm and hand gesture		x	x
		Eye movement		x	x
		Facial expression		x	x
		Head movement			x
		Posture shift			x
	Agent appearance	Agent visualization		x	x
		Age			x
		Clothing			x
		Color of agent			x
		Degree of human-likeness		x	x
		Facial feature		x	x
		Gender			x
		Name tags			x
Auditory	**Voice**	German		x	
		Dialect (Swiss)			x
		Gender of voice			x
Invisible	**Chronemics**	First turn			x
	Haptics	Tactile touch	x		x

Table 1. Social cues of the conversational agents based on the taxonomy of Feine *et al.* [15]

3.2 Scenario Description

After designing the prototypes, the project team created recordings of realistic financial service encounters, *i.e.,* one conventional one and three advisory sessions supported by one of the three CA, based on the workshops and interviews conducted with advisors in the requirements exploration and analysis. A male real-world advisor and female project member played the roles of the advisor and the client. In order to design viable advisory scenarios, we relied on expertise of the advisor participating in the videos as well as information from the requirements exploration and analysis phase. Each encounter was divided into five phases: *Welcome*, *Strategy*, *Interest*, *Investment*, and *Closure*. We integrated the prototypes based on contextual knowledge such that they provide added value and presumably minimize the IT-based disruption. The final scenarios were validated by multiple experienced advisors who found them to be realistic. Table 2 provides an overview of the five different phases, *i.e., Welcome*, *Strategy*, *Interest*, *Investment*, and *Closure*, the targeted outcome, and the CA's task for each phase.

To begin with, the advisor welcomes the client in the *Welcome*-Phase and breaks the ice by asking her how she arrived at the bank today. In the following, he introduces himself and the bank. Then, the advisor outlines the structure of the encounter. The CA is introduced either by the advisor or, in the case of the strongly anthropomorphic prototype *Pius*, by itself. Next, the advisor asks the client to tell him about herself and the reason for attending the advisory session (*e.g.,* date of birth, financial goals, and personal interests). The CA records the conversation and fills the corresponding fields on the screen with data. When the client asks not to record a specific answer, the recording is stopped and re-started by the advisor or, in the case of *Pius* automatically by the CA. In the end,

	Phase Goal	**CA's Tasks**
Welcome	Introduction of advisor and bank Setting appropriate ambiance	Introduction of itself Collection and display of client's personal data
Strategy	Understanding the client's risk capacity Definition of the appropriate investment strategy	Provision of information about investment strategies
Interest	Elicitation of the client's personal investment interests	Display of investment topics Adjustment of investment topics
Investment	Definition of a high-level investment strategy Selection of an individual investment portfolio	Display of investment portfolio Adjustment of investment portfolio Provision of information about investment strategies and recommendations
Closure	Summary of the discussion Agreement on the next steps	Sending of documentation Creation of meeting minutes

Table 2. Goal and CA's tasks ordered by the five phases of the service encounters

the advisor summarizes the collected data and asks the client if anything else is important to her. Once all data is collected, the advisor asks the CA to save the gathered data.

In the *Strategy*-Phase, the client fills out a regulatory required risk profile under the guidance of the advisor. The results from the risk profile are discussed by recommending a strategy for the investment. The client is informed about the strategy that fits her profile best, and the structure of an investment portfolio is explained (*i.e.,* different financial products like stocks, bonds). Additionally, the recommendation is put into perspective by providing information about riskier and more conservative strategies with different emphasis on the financial products (*e.g.,* more or fewer stocks). The CA displays the strategy and additional views (*e.g.,* diversification in regions), and *Pius* also explains the shown visualizations shortly. In contrast, the advisor explains everything in the encounter using *Box* or *Smiley*. Finally, the client and advisor decide on the appropriate strategy for the investment portfolio.

Next, in the *Interest*-Phase, the advisor wants to discuss investment topics that could be interesting for the client. The CA displays potential topics that are recognized from the discussion in the first two phases of the encounter (*e.g.,* sustainability and arts). The advisor explains the reasoning to the client and asks her if she would like to invest in those topics. She decides on the topics that are interesting for her, and the other topics are discarded. Again, *Pius* registers autonomously that a topic is not of interest and deletes it from the list. In the encounter with *Box* and *Smiley*, the advisor performs this task by manually clicking on the displayed topic. The advisor also encourages the client to raise topics that are interesting to her. These are added automatically by *Pius* or by request of the advisor. The advisor introduces topics of his own that might fit the client's interests. Then, the advisor asks the CA to save the investment topics.

In the following *Investment*-Phase, the advisor recalls the investment strategy and provides further information about the strategy like geographic diversification. *Pius* shows visualizations and explains what they mean. For example, *Pius* explains the allocation of different countries under the term “emerging markets”. *Box*, on the other side, only displays the visualization, and the advisor provides all the information. As a next step, the client would like to invest more in “emerging markets” which the CA then simulates. It intervenes following the simulation as the investment strategy would be too risky compared to the client’s risk profile. *Box* does that by displaying a red frame around the visualization. *Smiley* pulls down the corner of his mouth such that it should be perceived as unhappy. *Pius* voices his concerns verbally. The client and advisor decide against the simulation and undo the changes.

Next, the CA is asked to display an example of a portfolio. The advisor leads the client to the individual positions and provides some additional information. The client asks if the portfolio could be adjusted to her needs. *Pius* identifies that request and recommends a stock that is based on the client’s interests. He adds the stock to the portfolio as the client likes the recommendation. In the encounter with *Box* and *Smiley*, the advisor himself makes the recommendation and then instructs the assistant to add the stock to the portfolio. Once the client is satisfied with the portfolio, the advisor asks the CA to send the documentation as an e-mail. Finally, the advisor summarizes the discussion in the *Closure*-Phase. He thanks the client for her visit, and she notifies him that she will let him know about her decision regarding the investment portfolio. *Smiley* and *Pius* also say goodbye to the client.

After providing a textual description of the scenarios, we want to give the reader an even better understanding of the videos. The *Investment*-Phase provides a comprehensive overview over the design and interaction of each prototype[1].

- ***Box***: https://youtu.be/N3M-Iar1dU0
- ***Smiley***: https://youtu.be/LdAPo008nnM
- ***Pius***: https://youtu.be/RIzlsKMOq3I

3.3 Data Collection

To answer the research questions, we conducted a video-based survey. This format has two important advantages. First, clients and, especially, advisors are expensive research resources. By using videos instead of on-site experiments, we were able to distribute the survey to a larger audience. This, in turn, yielded broader support of our conclusions. Second, the experiment setting is more controlled using a video-based survey than a live advisory session. Furthermore, measuring social presence in a video-based setting has previously been successfully applied [1, 34].

Next, we identified suitable measures for social presence and competence. Competence and social presence cannot be objectively observed and measured with one single score. Swan *et al.* [60] developed the measurement to assess the buyer’s trust towards a salesperson, where competence is an important sub-measure. We use these measures to quantify the competence of the advisor and the CA since it is not only a commonly used approach to measure competence, but its original sales context also matches the financial advisory service context. Prior research has used these measures for bank advisory sessions, and it is established for measuring bank service quality properties of a CA [45, 50]. Since we assume the CA to be a social actor, we use the exact competence measurements for the human advisor and the CA.

Social presence is a multi-dimensional characteristic, and, therefore, it is necessary to use different approaches to capture this eclectic attribute. It is crucial to answering not only questions particularly related to perceived anthropomorphism since this would distort the results due to the obviousness of the purpose. To achieve this, we integrated two measures into the survey. First, we applied the 5-item anthropomorphism scale by Bartneck *et al.* [2] where the participants position the CA on a scale between two adjectives: fake/natural, machinelike/human-like, unconscious/conscious,

[1] All depicted persons gave their consent to publish the recorded videos on YouTube.

artificial/lifelike, and if it moves rigidly/elegantly. Second, we used a 5-item scale by Moussawi and Koufaris [41] which measures whether the agent can speak like a human, can be happy, funny, or caring, and is friendly and respectful.

	Measures by	**Likert-scale**	**Examples**
Social presence	Bartneck *et al.* [2]	5-item	"Do you perceive the conversational agent to be unconscious/conscious?"
	Moussawi and Koufaris [41]	5-item	"I think, the conversational agent can be happy. How much do you agree?"
Competence	Swan *et al.* [60]	5-item	"The conversational agent knows, what he is talking about. How much do you agree?"

Table 3. Measures used to quantify for social presence and competence

The survey was distributed among members of a university-wide database and the project members' personal network. It ran from December 2020 until February 2021. The participants watched the videos in a randomly assigned order (*e.g.,* First, *Box*; Second, *Pius*; Third, *Conventional*; Fourth, *Smiley*) and filled out the questionnaire after each service encounter.

The participants used their personal computers to access the survey. Besides other relevant information, the participants were instructed to empathize with the client's situation. They should imagine being present in the encounter and explore how they would perceive the interaction. The videos consisted of multiple views on the situation (*e.g.,* side or top view) to facilitate a comprehensive representation of the situation and interaction. Additionally, the video showed split screens of the encounter and the presented visualizations. Examples of the different views can be found in Figure 2.

The participants were incentivized with a voucher of CHF 30 (approx. $30) for a Swiss chocolatier, public transport, or an e-commerce platform. 75 participants filled out the survey completely. 42 (56%) were female and 33 (44%) were male. They were at the age of 19 to 68 years slightly skewed towards the younger ages (mean = 37 years, median = 34 years). 55% of the participants indicated to be inexperienced in bank advisory as a client. We sought to achieve a balanced ratio between experienced and inexperienced clients as financial advisors must be able to deal with both in their daily work. Often, they face inexperienced clients who request financial counseling for the first time in their life. We aim to measure the social presence and perceived competence which are two attributes of an advisor and a CA that require social competence. Consequently, our sample size reflects a realistic population of bank clients, which increases our results' validity and practical relevance.

3.4 Data Analysis

The data analysis was carried out in four main steps. First, we assessed whether the three prototypes achieved the intended social presence. Second, we analyzed the impact of the social presence on the CA's and advisor's competence. Third, we determined the correlation between the competence of the CA and the advisor. Fourth and last, we conducted a workshop among the co-authors to discuss the results concerning the relevant literature. The data analysis was carried out by the first author and checked by two co-authors.

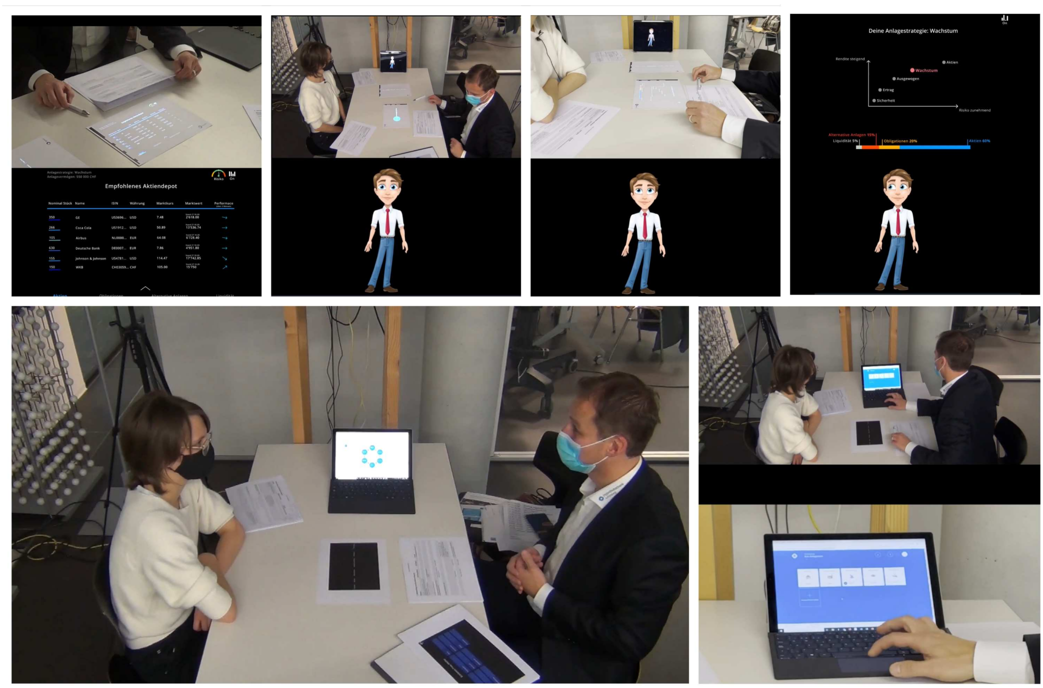

Fig. 2. Overview of different point of views used in the videos to provide the survey participant with an immersive experience despite the survey situation.

We did not perform any sort of data cleansing, such as excluding qualitatively poor or outlaying responses for the data analysis. We calculated the averages for the different measures and aggregated these again by averaging such that we used one value for competence and one value for social presence for the further data analysis. In a first step, we confirmed that the designed differences in social presence between the three prototypes were perceived by the study participants as intended. For this, we checked for significant differences between the prototypes' scores in the social presence measurements by carrying out an analysis of variance. As our data is not normally distributed but ordinal scaled, we performed the Kruskal-Wallis test with the level of social presence as the independent variable.

In a second step, we carried out a two-way analysis of variance (ANOVA) in order to answer **RQ 1** and **RQ 2**. This test allowed us to determine significant differences in perceived competence of the CA and the agent. The two independent variables were the prototypes (*i.e., Box, Smiley, Pius*) and the actor (*i.e.,* CA and Human Advisor). The perceived competence was the dependent variable. Our data set fulfilled the requirements for a two-way ANOVA. Shapiro-Wilk tests confirmed the normal distribution of the dependent variable. We also checked for the homogeneity of variance with Levene's test. In order to compare each treatment with the other two, we ran the Bonferroni adjustment.

In the third step, we calculated the effect of the CA's competence on the advisor's competence to answer **RQ 3**. We performed the Pearson correlation test to identify a potential positive or negative correlation between the perceived competence of both actors. Our data set fulfilled all requirements: variables are at least interval scaled, normally distributed, and the considered correlation is linear.

	Mean	Standard deviation
Box	2.69	0.56
Smiley	2.89	0.62
Pius	3.48	0.64

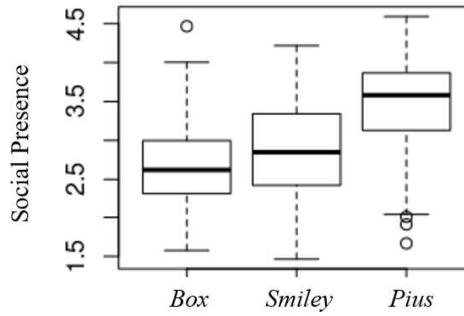

Fig. 3. Perceived social presence by treatment

4 RESULTS

In the first step, we aimed at verifying our experiment design. The Kruskal-Wallis test confirmedu that there are significant differences in social presence between some prototypes ($\tilde{\chi}^2 = 56.329$, $p < 0.001$, df = 2). On the one hand, the Bonferroni posthoc test revealed the significance between *Box* and *Pius* ($p < 0.001$) as well as between *Smiley* and *Pius* ($p < 0.001$). On the other hand, *Smiley* and *Box* are not perceived as significantly different in terms of social presence ($p > 0.05$). Figure 3 states the mean scores and standard deviations of the prototypes and visualizes them as a box-plot diagram. It shows that the difference between the less socially present prototypes is smaller than the difference to *Pius*.

Table 4 presents the descriptive statistics of the competence scores grouped by the treatment - *Box*, *Smiley*, and *Pius*.

		Mean	Standard deviation
Box	*Advisor*	3.45	0.522
	CA	3.2	0.477
Smiley	*Advisor*	3.44	0.564
	CA	3.04	0.578
Pius	*Advisor*	3.42	0.557
	CA	3.45	0.611
Conventional	*Advisor*	3.36	0.66
	none	na	na

Table 4. Perceived competence by treatment

We carried out the two-way ANOVA to reveal significant differences in the perceived competence of the CA and advisor between the three prototypes. Table 5 shows the mean score for the perceived competence by the advisor and CA per prototype as well as the standard deviation. There is no value for the CA in the conventional setting as this encounter is only conducted by the client and advisor. The results show that the perceived competence of the CA does not increase with a higher level of social presence. On the contrary, the perceived competence lowers with *Smiley* and reaches its highest score with *Pius*. This leads to a slightly V-shaped curve as shown in Figure 4. For the advisor, the perceived competence slightly decreases from a mean score of 3.45 with *Box* to 3.42 with *Pius*. With a mean score of 3.36, the conventional encounter was rated the lowest.

The ensuing Bonferroni posthoc test revealed that for the CA's competence there is only a statistically significant difference between *Box* and *Pius* ($p < 0.001$) as well as *Smiley* and *Pius* (p < 0.001). We could not find a significant difference between *Box* and *Smiley* (p > 0.05). Thus, *Pius* was significantly perceived as most competent. For the advisor, no significant difference could be determined between any of the settings. Consequently, the CA's social presence has neither a positive nor negative effect on the perceived competence of the advisor. Figure 4 sums up the results of the Bonferroni posthoc test with a significant two-way interaction between social presence and the CA's competence ($F(2, 456) = 10.7$, $p < 0.001$) as well as no interaction effect between social presence and the advisor's competence ($F(2, 456) = 0.08$, $p > 0.05$).

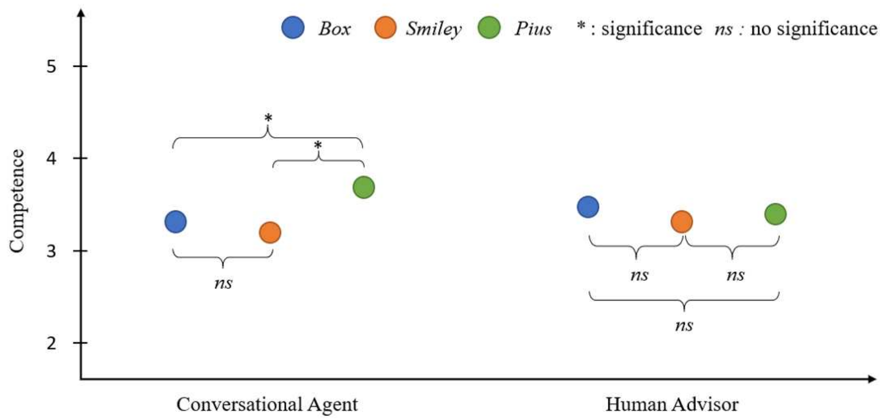

Fig. 4. Overview of differences in perceived competence between treatments

Finally, we carried out a Pearson correlation test to disclose the correlation between the perceived competence of the CA and advisor. For all prototypes, we found positive correlations with a medium to strong effect. This means that the advisor's perceived competence increases if the CA is perceived more competent and vice-versa. Table 5 shows the results of the Pearson correlation test between the advisor's competence and the competence of the CA per treatment.

	Correlation strength	**Significance** (p <...)	**Interpretation**
***Box*-Advisor**	0.48	0.001	medium to strong effects
***Smiley*-Advisor**	0.44	0.001	medium to strong effects
***Pius*-Advisor**	0.39	0.001	medium to strong effects

Table 5. Pearson correlation results between advisor's competence and the competence of the CA

5 DISCUSSION

We make three main theoretical contributions in accordance with the previously introduced research questions. First, machines are attributed competence depending on the intensity of their social presence. Second, the social presence of the CA does not affect the perceived competence of the advisor. Third, our results reveal a direct correlation between the perceived competence of the CA and the advisor.

5.1 Theoretical Contributions

RQ 1 asks whether the social presence of a CA increases its perceived competence. The results show that, indeed, social presence and humanness might play a role in the perception of competence.

Pius, designed to exhibit the highest level of humanness and attributed a significantly higher social presence by the participants, was perceived as significantly more competent than the other two prototypes. The other two prototypes did not significantly differ in terms of perceived social presence or competence. Still, both score relatively high on the social presence scale (approx. 3 on a 5-point scale).

Our results are in line with prior research attributing human qualities to computers [42] and more specific studies attributing such characteristics to CA [33, 41]. Subjects assigned some human characteristics to all prototypes. While previous research on the perception of CA was focusing on the interaction between a single user and a single CA, this research confirms that users might tend to attribute human characteristics to CA even if they observe a setting involving more than one human. Prior studies on the introduction of IT into service encounters also found that IT yields effects similar to introducing a third actor. Even if IT had the form of a presentation on a table without reactions on the human voice and was designed explicitly to be a tool rather than an agent, users attributed human characteristics [13].

This study confirms and extends those results. Even if the *Box* was designed to be non-intrusive and did not frequently engage in interaction, the subjects perceived it as exhibiting a social presence. This is comparable to *Smiley*, offering voice input and output, presenting oneself in the interface, and informing the users visually and verbally about its overall state. We hypothesize that the effect described as CASA [42, 44] might even improve in highly social situations like advisory service. The whole character of the situation might prime subjects to employ social rather than technical scripts, thus even enhancing the chance of treating simple technical devices as a potential third actor. Consequently, our study expands the knowledge about CA and their ascribed human qualities by investigating the application of CA in a collaborative setting with multiple human actors.

This extension has implications for the research on CA and advisory services. If we acknowledge that CA can be treated as a social being, then we need to recognize that what we observe in a service encounter is not a dyadic but rather a triadic interaction involving the client, the advisor, and the IT [13]. Service encounters, which were conventionally framed as a dyadic collaboration between a client and a frontline employee (*e.g.,* financial advisor or doctor), become a triad when a CA gets introduced into the setting. This reframing changes how one should think about designing a CA for a service encounter. It is not about designing a tool that integrates into an existing situation. Instead, it is about re-designing the social configuration of the whole situation. This means focusing entirely on the look and feel of the CA is too narrow.

Designers need to think about the whole social encounter they alter and what the ultimate social encounter should look like. This makes clear that research should broaden its view when attending to the design of CA. Whereas there have been many studies attending to specific design aspects of the CA, far too few studies try to understand the whole situation, especially in a context involving several human actors. Instead of referring to designing a CA, one should rather refer to the task of engineering a social encounter. At this point, it must be mentioned that the triadic setting should only be considered for the social engineering of an encounter. It is not contradictory to the observation that the advisor and the conversational agents are perceived as one team, but it is another perspective on the encounter. All prototypes, *i.e., Box*, *Smiley*, and *Pius*, did not differ in how well they understand conversation context during the encounter or how relevant their output is. Therefore, this characteristic does not have an influence on the results and the answer to **RQ 1** or **RQ 2**.

RQ 2 asks whether the human character of the CA has an impact on the perceived competence of the advisor. The answer is no. There was no significant difference between the treatments as to how competent the advisor appeared to the subjects. For all four treatments, including the three prototypes and the conventional advisory, we did not find any significant differences in the perceived advisor's competence. Even as the results do not show significant improvements to the advisor's competence, introducing a CA regardless of its level of social presence does apparently no harm to how the advisor is perceived. As advisory service encounters commonly cover complex or intimate topics and thus require close interaction and trust between the advisor and the advisee, it is highly relevant only to introduce IT, which does not disturb the interpersonal relationship between those two individuals.

Previous research has shown that the introduction of technology often created obstacles to establish a good working relationship between the advisor and the advisee [28, 38]. Additionally, many advisors are afraid of losing face when interacting with IT in front of a client [12, 46]. Therefore, we conclude that it is essential to show that CA do not generate unintended effects that might impair the advisor. Advisors do not have to fear a loss of perceived competence when being supported by a highly present CA that actively engages in the conversation by providing additional information or recommendations.

RQ 3 focuses on the interrelation between the perceived competence of the CA and the perceived competence of the advisor. Our results show a significant correlation with medium to strong effects. The CA and advisor are both seen as competent actors in the service encounter. Moreover, the participants' perception of CA's and advisor's competence stand in relation to each other. This suggests that they might be seen as a human-machine team that is generating an overall sense of competence or of having received competent advice. This insight is precious for the discourse of how to design agents as teammates for humans [53]. First, our study investigates the clients' perception of human-machine collaboration in a realistic scenario. Second, besides its design restrictions, the CA was able to interact in the encounter without any limitation, especially, *Box* and *Smiley*.

Our study shows that CA can be perceived as competent whilst positively affecting the perception of the human counterpart as a competent advisor. However, our data shows that the advisor was perceived as competent in all four scenarios (*Conventional*, *Box*, *Smiley*, and *Pius*). An advisor that is perceived less competent might not benefit from a CA with a strong social presence such as *Pius*. The correlation strength is lower for the *Pius*-Advisor combination as for the other two, as we can see in Table 5. This might indicate that a strongly present and competent CA could negatively impact less competent advisors as the CA potentially overpowers the advisor in collaboration with the client. For designers, this requires careful decisions regarding the social presence of a CA. While a competent advisor benefits from a competent machine-teammate such as *Pius*, the same design potentially harms the fragile relationship between the client and advisor. Consequently, designers need to consider multiple factors, such as the advisor's expertise in financial matters and the handling of technology.

Overall, the observation that the impression advisors make on the potential client is related to the impression CA make on the client is of particular importance. Modern IT tools involving voice and multimodal interfaces will likely spread more and more in commercial and non-commercial services, including various frontline interactions or doctor-patient encounters [10]. However, studies that assess the impact of such modern IT solutions on those highly ritualized encounters are rare [13]. Lacking clear evidence, frontline employees might decline the usage of such systems while pointing out that they destroy the typical configuration of institutional encounters, one in which the advisor acts as the moderator and the subject matter expert while controlling the material, and the thematic structure [12]. Much of what advisors do in their service encounters and how they do it is related to their need to present themselves as competent expert [12]. By providing them evidence that CA bear the potential to enhance the impression they make on the client might make them more likely to accept CA in their encounters.

5.2 Implications for Designs

This study has implications for designers of CA for service interactions involving multiple human and digital agents.

The core implication follows directly from the results of the survey. The perceived competence of the CA is related to its social presence. However, this is not a linear relation. Simply adding social cues like a face or a voice to the design does not guarantee that the CA will be perceived as more competent. In our case, only the highly social *Pius* was perceived as significantly more competent than the two other designs. Importantly, *Pius* is not simply a more social version of the *Smiley*, but rather a larger step towards personification of the CA. This might suggest that adding a sense of personality to an agent can lead to enhanced perceptions of its competence.

Additionally, the results make clear that the perception of the advisor's competence is positively correlated with the competence of the CA. This means that the designers should aim at improving the perceived competence of the CA. Since we used the measure of Swan*et al.* [60] to quantify competence, one possible starting point may be to improve a partial aspect related to a certain sub-measure. Assuming that the modification reduces none of the other sub-measures, we can consider the following example. A CA which is perceived as more knowledgeable what he is talking about results in an increase of the overall competence score.

Given that *Pius* was the CA with the highest level of perceived competence and outperformed the other two CA in terms of social presence and perceived competence, in the following, we summarize the specific design considerations we made when designing *Pius* and offer them in form of design guidance to be applied if a designer wants to achieve a highly social present and competent CA for use in triadic configuration.

Obey the conversational rituals. When designing *Pius*, we considered the mechanics of the conversation. For instance, we presented *Pius* as capable of Swiss German dialect as spoken by the participants and assured that its contributions fit the context of the ongoing conversation. This was inspired by earlier research on IT in advisory services [11, 28], which suggested that IT must be able to naturally fit into the service encounter, meaning that it needs to be aware of human rituals and scripts (*e.g.,* it should be able to identify the topic and phase of the encounter and act accordingly). We assert that the perception of conversational competence impacts the overall perception of competence, including the notion of professional expertise.

Create a coherent experience. When designing *Pius*, we were driven by the desire to make its verbal and non-verbal behavior synchronized to create a consistent experience. Earlier research on IT in advisory services suggests that the persuasive effect of an advice is more intense if the story presented verbally by an advisor is supported with convincing visual material [8]. *Pius* offered coherent experience in several aspects: the animation of mimics and face followed his verbal actions, he referred to what was shown on the screen, the presentation of the information was synchronized with the information provided verbally. We assert that this combination generates a coherent experience that is associated with competence.

Turn CA and advisor into a team with training. We put much effort into making sure that the advisor knows about the abilities of the CA to ensure smooth interaction and collaboration between advisor and CA. When thinking about human-teamwork, it is difficult to imagine a well-functioning team that lacks harmonic interaction and good communication between its members. We assume that the same holds for teaming a CA and an advisor. On the one hand, it might be necessary for the advisor to accommodate inputs from the CA in their explanations, *e.g.,* inputs regarding specific investment recommendations or portfolio modifications. On the other hand, the advisor should be able to engage with the CA in a way that the CA can easily and efficiently process and deal with. However, the interaction with AI systems, including CA, requires complex training of the human teammate as they must be able to control the AI system and understand potential outcomes and failures [54]. To enhance the interaction and communication between advisor and CA, and therein create the perception of a well-functioning advisor-CA team, we believe that the training should thus not be focused and limited to the technical capabilities of the CA. Rather, special emphasis should be placed on training the advisor on how to interact and communicate with the CA. Such training could exhibit cross-training and enactive learning, which presumably foster the advisor's understanding of the CA and teach them effective communication with it.

Make the CA adjustable. Regarding the results of the survey data, a CA in a dyadic interaction might be equipped with social cues (*e.g.,* facial expression) to appear more authoritative to persuade the user to carry out a specific task or behavior. The same behavior by the CA could probably have a very different effect in a service encounter. As an observant participant, the client might lose trust in the advisor's competence if the CA interacts very dominant. Specifically, if the CA offers information that contradicts the advisor's argumentation or to which the advisor cannot relate, the client might see the advisor as unprepared or simply not knowledgeable enough. Additionally, the CA could create an unacceptable level of intimacy with its presence in the service encounter. For instance, Kilic *et al.* [29] had discovered that the willingness of clients to disclose personal information was lowered when IT was used in a service encounter compared to a conventional encounter where the advisor took notes on a piece of paper. Misra *et al.* [38] found similar effects in social interactions between humans in the presence of mobile devices such as smartphones. Consequently, this implies that a CA's social presence potentially needs to be flexible to adjust to the changing context within a service encounter.

In summary, designers of CA can use multiple levers to manipulate the social presence of CA when designing them for service interactions with multiple human and digital agents. These levers are not only limited to technical considerations but also include organizational and socio-technical aspects. Keeping the mentioned design considerations in mind can help increase the social presence of a CA and might ultimately improve the perceived competence of a CA.

6 CONCLUSION, LIMITATIONS & FUTURE WORK

Conversational agents are spreading rapidly in all aspects of our daily lives. They also find their way slowly into professional settings that are more dynamic and less standardized such as financial service encounters. We developed three prototypes of CA in a DSR project together with two Swiss banks. This quantitative study tests these prototypes and their perception by potential bank clients through video recordings of CA-supported service encounters. Our results make three theoretical contributions when introducing CA into collaborative settings with more than one human actor. First, we confirm that humans perceive CA as competent actors also in settings where other humans are present and suggest not only focusing on the CA design but also the context it is applied in. Second, this study provides insights into the mechanics of autonomous teaming [47]. Third, we discovered a relationship between the client's perception of the competence of the CA and the advisor. This suggests that they can be seen as a human-machine team.

These findings provide two major implications for the design of CA for service encounters. On the one hand, designers need to consider the setting in which the CA is deployed. A service encounter has unique characteristics as the client as a third actor observes the interaction between the advisor and the CA. It is therefore important to account for the perception an observant might have of a CA and advisor. We also hypothesize that a socially present CA needs to be flexibly adaptable to different contexts and supports a positive perception of the advisor. On the other hand, the interaction between the advisor and the client is also influenced by the CA requiring new forms of institutional talk. For example, they need to learn how to incorporate the CA in a natural way, thus benefiting from its inputs.

This study does not come without limitations. First, the quantitative study reveals statistical effects that must be interpreted accordingly. Future research could conduct user testings with clients and advisors to gain deeper insights into the origin of the effects. Observations and interviews would provide rich data for further explaining the effects observed in the present study such that we can attribute a increase in competence or social presence to specific partial aspects of a CA. This would ease the way for future improvements of the CA's perception because designers and researcher can focus on the main attributes that provide the greatest marginal utility. Second, we cannot provide specific design requirements for CA in triadic collaboration due to the many design decisions the project team made in order to develop the prototypes. HCI research should gain deeper insights into the design of CA for settings with multiple human actors and try to adapt existing knowledge from dyadic settings. Third, our study focuses on one form of service encounters - financial advisory services. Future research could investigate differences in other triadic collaborations such as doctor-patient encounters supported by a CA. Additionally, studies on CA in settings with more than two human actors could yield further exciting results.

REFERENCES

[1] Stephanie A. Andel, Triparna de Vreede, Paul E. Spector, Balaji Padmanabhan, Vivek K. Singh, and Gert-Jan de Vreede. 2020. Do social features help in video-centric online learning platforms? A social presence perspective. *Computers in Human Behavior* 113 (2020), 106505. https://doi.org/10.1016/j.chb.2020.106505

[2] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1, 1 (2009), 71–81.

[3] Timothy W Bickmore, Kathryn Puskar, Elizabeth A Schlenk, Laura M Pfeifer, and Susan M Sereika. 2010. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers* 22, 4 (2010), 276–288.

[4] Frank Biocca, Chad Harms, and Judee K. Burgoon. 2003. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoper. Virtual Environ.* 12, 5 (Oct. 2003), 456–480. https://doi.org/10.1162/105474603322761270

[5] J. K. Burgoon, J. A. Bonito, B. Bengtsson, C. Cederberg, M. Lundeberg, and L. Allspach. 2000. Interactivity in human-computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior* 16, 6 (nov 2000), 553–574. https://doi.org/10.1016/S0747-5632(00)00029-7

[6] Minjee Chung, Eunju Ko, Heerim Joung, and Sang Jin Kim. 2020. Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research* 117 (2020), 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004

[7] Tino Comes and Gerhard Schwabe. 2016. From Fuzzy Exploration to Transparent Advice: Insights into Mobile Advisory Services. In *49th Hawaii International Conference on System Sciences (HICSS)*. University of Hawaii at Manoa, Association for Information Systems IEEE Computer Society Press, Hawaii, USA, 1256–1265. https://doi.org/10.1109/HICSS.2016.159

[8] Tino Comes and Gerhard Schwabe. 2016. How to Diminish Advice Discounting with Mobile Multimedia Interventions. In *European Conference on Information Systems (ECIS)*. s.n., Istanbul, Turkey. https://doi.org/10.5167/uzh-124598

[9] Mateusz Dolata, Doris Agotai, Simon Schubiger, and Gerhard Schwabe. 2019. Pen-and-paper rituals in service interaction: Combining high-touch and high-tech in financial advisory encounters. In *Proceedings of the ACM on Human-Computer Interaction*. ACM, New York, NY, USA, 1–24.

[10] Mateusz Dolata, Doris Agotai, Simon Schubiger, and Gerhard Schwabe. 2020. Advisory Service Support that Works: Enhancing Service Quality with a Mixed-reality System. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–22.

[11] Mateusz Dolata, Mehmet Kilic, and Gerhard Schwabe. 2019. When a computer speaks institutional talk: Exploring challenges and potentials of virtual assistants in face-to-face advisory services. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

[12] Mateusz Dolata and Gerhard Schwabe. 2017. Paper practices in institutional talk: How financial advisors impress their clients. *Computer Supported Cooperative Work (CSCW)* 26, 4 (2017), 769–805.

[13] Mateusz Dolata, Susanne Steigler, Fiona Nüesch, Ulrike Schock, Doris Agotai, Simon Schubiger, Mehmet Kilic, and Gerhard Schwabe. 2019. Welcome, Computer! How Do Participants Introduce a Collaborative Application During Face-to-Face Interaction?. In *IFIP Conference on Human-Computer Interaction*. Springer, 600–621.

[14] Kathleen M Eisenhardt. 1989. Agency theory: An assessment and review. *Academy of management review* 14, 1 (1989), 57–74.

[15] Jasper Feine, Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. 2019. A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human Computer Studies* 132 (2019), 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

[16] Joel E Fischer, Andy Crabtree, James A Colley, Tom Rodden, and Enrico Costanza. 2017. Data work: how energy advisors and clients make IoT data accountable. *Computer Supported Cooperative Work (CSCW)* 26, 4 (2017), 597–626.

[17] Andrew Gambino, Jesse Fox, and Rabindra A Ratan. 2020. Building a stronger CASA: extending the computers are social actors paradigm. *Human-Machine Communication* 1, 1 (2020), 5.

[18] Tobias Giesbrecht, Tino Comes, and Gerhard Schwabe. 2015. Back in Sight, Back in Mind: Picture-Centric Support for Mobile Counseling Sessions. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 486–495.

[19] Joseph H Golec. 1992. Empirical tests of a principal-agent model of the investor-investment advisor relationship. *Journal of Financial and Quantitative Analysis* 27, 1 (1992), 81–95.

[20] Li Gong. 2008. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior* 24, 4 (2008), 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007 Including the Special Issue: Integration of Human Factors in Networked Computing.

[21] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2010. Relating conversational expressiveness to social presence and acceptance of an assistive social robot. *Virtual Reality* 14, 1 (2010), 77–84.

[22] Peter Heinrich, Mehmet Kilic, Felix-Robinson Aschoff, and Gerhard Schwabe. 2014. Enabling Relationship Building in Tabletop-Supported Advisory Settings. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Baltimore, Maryland, USA) *(CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 171–183. https://doi.org/10.1145/2531602.2531697

[23] David Holman and Roel Vertegaal. 2008. Organic user interfaces: designing computers in any way, shape, or form. *Commun. ACM* 51, 6 (2008), 48–55.

[24] David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns. 2005. Paper Windows: Interaction Techniques for Digital Paper. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 591–599. https://doi.org/10.1145/1054972.1055054

[25] James M. Honeycutt and Suzette P. Bryan. 2011. *Scripts and communication for relationships*. Peter Lang, Bern, Switzerland. https://www.peterlang.com/view/title/64197

[26] Earl Hunt. 1979. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. *The American Journal of Psychology* 92, 1 (1979), 176–178. http://www.jstor.org/stable/1421499

[27] Moonkyoung Jang, Yoonhyuk Jung, and Seongcheol Kim. 2021. Investigating managers' understanding of chatbots in the Korean financial industry. *Computers in Human Behavior* 120 (2021), 106747.

[28] Mehmet Kilic, Mateusz Dolata, and Gerhard Schwabe. 2016. How IT-Artifacts Disturb Advice Giving–Insights from Analyzing Implicit Communication. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*. IEEE, University of Hawaii at Manoa, Association for Information Systems IEEE Computer Society Press, Hawaii, USA, 878–887.

[29] Mehmet Kilic, Mateusz Dolata, and Gerhard Schwabe. 2017. Why do you ask all those questions? Supporting client profiling in financial service encounters. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. University of Hawaii at Manoa, Association for Information Systems IEEE Computer Society Press, Hawaii, USA, 4695–4704.

[30] Ahyeon Kim, Minha Cho, Jungyong Ahn, and Yongjun Sung. 2019. Effects of Gender and Relationship Type on the Response to Artificial Intelligence. *Cyberpsychology, Behavior, and Social Networking* 22, 4 (2019), 249–253. https://doi.org/10.1089/cyber.2018.0581

[31] Soomin Kim, Jinsu Eun, Changhoon Oh, Bongwon Suh, and Joonhwan Lee. 2020. Bot in the Bunch: Facilitating Group Chat Discussion by Improving Efficiency and Participation with a Chatbot. (2020), 1–13. https://doi.org/10.1145/3313831.3376785

[32] Seo Young Kim, Bernd H. Schmitt, and Nadia M. Thalmann. 2019. Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing Letters* 30, 1 (2019), 1–12. https://doi.org/10.1007/s11002-019-09485-9

[33] Youjeong Kim and S Shyam Sundar. 2012. Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior* 28, 1 (2012), 241–250. https://doi.org/10.1016/j.chb.2011.09.006

[34] Karel Kreijns, Paul A Kirschner, Wim Jochems, and Hans Van Buuren. 2011. Measuring perceived social presence in distributed learning groups. *Education and Information Technologies* 16, 4 (2011), 365–381.

[35] Samantha Krening and Karen M Feigh. 2018. Characteristics that influence perceived intelligence in AI design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications, Los Angeles, CA, USA,

1637–1641.

[36] Liliana Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9 (2018), 1248–1258.

[37] Paul Luff, Karola Pitsch, Christian Heath, Peter Herdman, and Julian Wood. 2010. Swiping paper: the second hand, mundane artifacts, gesture and collaboration. *Personal and Ubiquitous Computing* 14, 3 (2010), 287–299.

[38] Shalini Misra, Lulu Cheng, Jamie Genevie, and Miao Yuan. 2014. The iPhone Effect: The Quality of In-Person Social Interactions in the Presence of Mobile Devices. *Environment and Behavior* 48 (07 2014). https://doi.org/10.1177/0013916514539755

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

[40] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100.

[41] Sara Moussawi and Marios Koufaris. 2019. Perceived intelligence and perceived anthropomorphism of personal intelligent agents: Scale development and validation. In *Proceedings of the 52nd Hawaii international conference on system sciences*. University of Hawaii at Manoa, Association for Information Systems IEEE Computer Society Press, Hawaii, USA, 115–124.

[42] Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of social issues* 56, 1 (2000), 81–103.

[43] Clifford Nass and Jonathan Steuer. 1993. Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research* 19, 4 (1993), 504–527.

[44] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 72–78.

[45] Jasminko Novak and Susanne Schmidt. 2009. When Joy Matters: The Importance of Hedonic Stimulation in Collocated Collaboration with Large-Displays. In *Human-Computer Interaction – INTERACT 2009*, Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, Raquel Oliveira Prates, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 618–629.

[46] Philipp Nussbaumer, Inu Matter, and Gerhard Schwabe. 2012. "Enforced" vs."Casual" Transparency–Findings from IT-Supported Financial Advisory Encounters. *ACM Transactions on Management Information Systems (TMIS)* 3, 2 (2012), 1–19.

[47] Rui Prada and Ana Paiva. 2009. Teaming up humans with autonomous synthetic characters. *Artificial Intelligence* 173, 1 (2009), 80–103. https://doi.org/10.1016/j.artint.2008.08.006

[48] Tamara Rakić, Melanie C Steffens, and Amélie Mummendey. 2011. Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of personality and social psychology* 100, 1 (2011), 16.

[49] Byron Reeves and Clifford Nass. 1996. *The media equation: How people treat computers, television, and new media like real people.* Cambridge university press, Cambridge, UK.

[50] Susanne Schmidt-Rauch and Gerhard Schwabe. 2010. From Telesales to Tele-Advisory in Travel Agencies: Business Problems, Generic Design Goals and Requirements. *ACM Trans. Manage. Inf. Syst.* 2, 3, Article 17 (Oct. 2010), 13 pages. https://doi.org/10.1145/2019618.2019623

[51] Ryan M. Schuetzler, G. Mark Grimes, and Justin Scott Giboney. 2019. The effect of conversational agent skill on user behavior during deception. *Computers in Human Behavior* 97 (2019), 250–259. https://doi.org/10.1016/j.chb.2019.03.033

[52] Gerhard Schwabe and Philipp Nussbaumer. 2009. Why information technology is not being used for financial advisory. In *ECIS 2009 Proceedings*.

[53] Isabella Seeber, Eva Bittner, Robert O. Briggs, Triparna de Vreede, Gert-Jan de Vreede, Aaron Elkins, Ronald Maier, Alexander B. Merz, Sarah Oeste-Reiß, Nils Randrup, Gerhard Schwabe, and Matthias Söllner. 2020. Machines as teammates: A research agenda on AI in team collaboration. *Information & Management* 57, 2 (2020), 103174. https://doi.org/10.1016/j.im.2019.103174

[54] Ben Shneiderman. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems. 10, 4 (2020). https://doi.org/10.1145/3419764

[55] Ben Shneiderman. 2021. Human-Centered AI. *Issues in Science and Technology* 37, 2 (2021), 56–61.

[56] John Short, Ederyn Williams, and Bruce Christie. 1976. *The social psychology of telecommunications.* Toronto; London; New York: Wiley.

[57] Nicolas Spatola, Nolwenn Anier, Sandrine Redersdorff, Ludovic Ferrand, Clément Belletier, Alice Normand, and Pascal Huguet. 2019. National Stereotypes and Robots' Perception: The "Made in" Effect. *Frontiers Robotics AI* 6 (2019), 21. https://doi.org/10.3389/frobt.2019.00021

[58] Jürgen Steimle. 2009. Designing pen-and-paper user interfaces for interaction with documents. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*. 197–204.

[59] Jürgen Steimle. 2012. *Pen-and-paper user interfaces: Integrating printed and digital documents.* Springer Science & Business Media.

[60] John E Swan, I Fred Trawick Jr, David R Rink, and Jenny J Roberts. 1988. Measuring dimensions of purchaser trust of industrial salespeople. *Journal of Personal Selling & Sales Management* 8, 1 (1988), 1–10.

[61] Jenny Van Doorn, Martin Mende, Stephanie M Noble, John Hulland, Amy L Ostrom, Dhruv Grewal, and J Andrew Petersen. 2017. Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of service research* 20, 1 (2017), 43–58.

[62] John van Wyk, Peter Brooke, and Bornstein Justine. 2018. Selling Industry 4.0 - The new sales mind-set for connected products. *Deloitte Insights* (2018). https://www2.deloitte.com/ch/en/pages/consumer-industrial-products/articles/transitioning-from-products-to-services-to-fuel-growth.html

[63] Joseph Weizenbaum. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (1966), 36–45.

Received July 2021; revised September 2021; accepted October 2021