Implementing an Intelligent Collaborative Agent as Teammate in Collaborative Writing: toward a Synergy of Humans and AI

Christina Wiethof
University of Hamburg
wiethof@informatik.uni-hamburg.de

Navid Tavanapour
University of Hamburg
tavanapour@informatik.uni-hamburg.de

Eva A. C. Bittner
University of Hamburg
bittner@informatik.uni-hamburg.de

Abstract

This paper aims at implementing a hybrid form of group work through the incorporation of an intelligent collaborative agent into a Collaborative Writing process. With that it contributes to the overall research gap establishing acceptance of AI towards complementary hybrid work. To approach this aim, we follow a Design Science Research process. We identify requirements for the agent to be considered a teammate based on expert interviews in the light of Social Response Theory and the concept of the Uncanny Valley. Next, we derive design principles for the implementation of an agent as teammate from the collected requirements. For the evaluation of the design principles and the human teammates' perception of the agent, we instantiate a Collaborative Writing process via a web-application incorporating the agent. The evaluation reveals the partly successful implementation of the developed design principles. Additionally, the results show the potential of hybrid collaboration teams accepting non-human teammates.

1. Introduction

Research on Artificial Intelligence (AI) is increasingly progressing shown by many new evolving technologies. Here, researchers mainly work on questions of effectiveness and efficiency regarding their newest developments [1]. Especially in the field of Machine Learning (ML) researchers aim to create an AI, which resembles Human Intelligence and could consequently replace a human being [2]. Thereby, they focus on an automatic learning approach [3] resulting in intelligent, autonomous systems. In certain domains with a huge amount of training data, this approach has already been successfully recognized [1, 4, 5].

However, it is known that technology is not everything [6]. Researchers aim to achieve a synergy of both humans and AI, i.e. combining the benefits and advantages of both [2, 7–10]. Therefore, also the human users' social perspective [11] is required. Even the best state-of-the-art technology will be useless, if its human users refuse it [6]. This also applies to ML approaches themselves, especially when they involve human users in the training, e.g. Reinforcement Learning or Human-in-the-Loop [1, 5, 12]. Thus, to achieve that synergy of working together and complementing as well as learning from each other, the human needs to accept a collaborative agent willing to learn from its contributions as well as to make corrections and improve the agent [2]. It should be pointed out, that throughout the paper we use the term *agent* for any *collaborative agent* and intelligent *computer agent* respectively "[covering] the idea of creating machines that can accomplish complex goals [including] facets such as natural language processing, perceiving objects, storing of knowledge and applying it for solving problems" [8] in collaboration settings.

As there is an advantage of combining human and artificial intelligence to achieve better collaboration outcomes [2, 8], the research gap and need for designing and developing such socio-technological teams has been disclosed [8]. We therefore consider socio technical factors of agent teammates and exemplify the intended synergy by regarding hybrid teams involving humans and agents. To specifically contribute and extend the scope of this research, Dellermann et al. [8] call for more research on practical applications in different domains. For instance, Bittner et al. [13] developed a taxonomy for conversational agents in collaborative work. Epstein [9], on the other hand, investigated a collaborative intelligence sharing a task with a person to demonstrate the potential synergy of humans and agents. Eventually, "rather than re-design our world for computers or submit to their decisions, we should begin to share our tasks with them" [9]. As we found a study, which revealed that an agent is capable of replacing actual human journalists [14], for our research at hand, we specifically regard a Collaborative Writing (CW) scenario. After all, there could rather be an advantage in

HĭCSS

the collaboration of an agent and human writers. For one thing, an agent may have more memory space and a higher computation rate as well as challenges the writer and promotes the writing process. For another thing, agents do not reach humans' skills and knowledge yet. Thus, by co-writing, the skills of the agent as well as the writer affect both the outcome as well as each other complementary. Manjavacas et al. [15] addressed this by developing an intelligent text generation system, which produces sentences or paragraphs to enable co-creation and CW between an author and an agent [15]. By doing this, agents contribute with story fragments and ideas, which the human collaborator might not be aware of [16]. Additionally, computational creativity itself has already achieved several successes, e.g. Narrative Science [16], poetry [17], storytelling [18] or melodic accompaniments for lyrics [19].

Still, "as machines evolve from tools to teammates, one thing is clear: accepting them will be more than a matter of simply adopting new technology" [6]. By fostering co-creativity in CW within a hybrid team, we examine the possibility of perceiving an agent as teammate [15]. This will enable further research on implementing hybrid forms of group work covering mutual learning benefits and acceptance. With that we aim to contribute to the overall research gap establishing acceptance of computer agents toward complementary hybrid work [2, 8, 10]. Therefore, we are conducting design science research to implement an agent into a collaborative writing process as teammate [10]. By doing this, we address three research questions: Q1: What are the requirements to ensure acceptance of an agent as teammate in CW? Q2: How can an agent be designed and implemented as teammate contributing to the goal of the CW process? Q3: How do the human teammates perceive and accept the contributions of the agent?

To support CW, we develop a CW process with an agent teammate and implement it on a web platform.

2. Research Approach

The research aims to contribute prescriptive design knowledge to the knowledge base by connecting the research areas of Human-Computer-Interaction and Socio-Technical Systems to design a solution for the incorporation of an agent teammate into a CW process [20]. In coherence with the design science research (DSR) approach, the DSR process by Peffers et al. [21] is used to derive design principles (DPs), which are then implemented and evaluated with an instantiated CW process in form of a web-application (see Figure 1) [21]. The problem identification and motivation are covered in the introduction. To define the objectives of the solution, meta-requirements (MRs) for an agent teammate are identified. This includes any personality traits and skills, that need to be assigned to an agent to be considered and accepted as teammate. To do so, we first consider related work from areas focusing on machines as teammates and hybrid teams as well as socio technical factors of agents. We then base our MRs on the Social Response Theory by Nass and Moon [22] aligned to the concept of the Uncanny Valley by Mori [23], and conduct expert interviews according to the approach of Meuser and Nagel [24]. For the design and development, the MRs are considered to derive DPs of an agent teammate, which are later on implemented. After the implementation of the agent in a CW process, a demonstration is carried out by instantiating the CW process in form of a prototypic web-application incorporating the agent [25]. Four groups of five participants took part in a test run [26] and in expert interviews [24] to evaluate the human teammates' perception and acceptance towards the contributions of the agent ex post [27]. Communication will be completed with this paper.
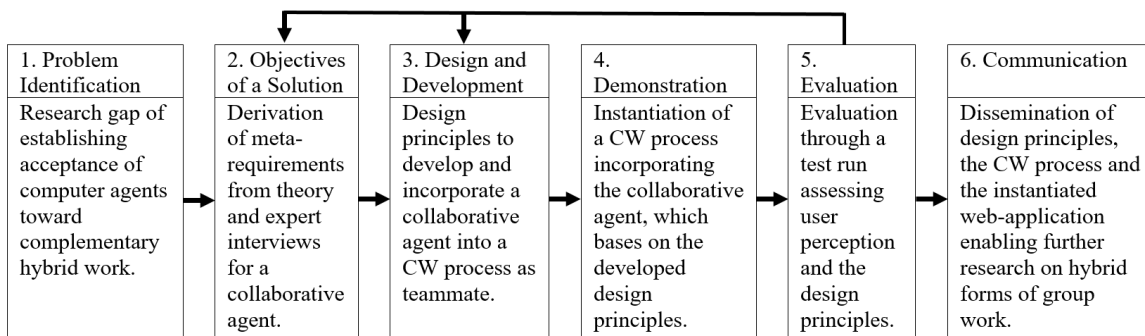
1. Problem Identification	2. Objectives of a Solution	3. Design and Development	4. Demonstration	5. Evaluation	6. Communication
Research gap of establishing acceptance of computer agents toward complementary hybrid work.	Derivation of meta-requirements from theory and expert interviews for a collaborative agent.	Design principles to develop and incorporate a collaborative agent into a CW process as teammate.	Instantiation of a CW process incorporating the collaborative agent, which bases on the developed design principles.	Evaluation through a test run assessing user perception and the design principles.	Dissemination of design principles, the CW process and the instantiated web-application enabling further research on hybrid forms of group work.

Figure 1. Structure along the DSR process

3. Related Work

Combining the strengths of humans and agents in collaborative work is not easy and neither is it enough to make a good collaboration team. Humans still think of technology as a tool, but need to consider and accept it as teammate of a hybrid group [6, 10]. Therefore, research considers social science findings about behaviors or attitudes toward humans and applies them for agents. The "computers are social actors" (CASA) paradigm introduces the relevance of assigning human characteristics, and social cues respectively, to agents [11] encouraging their acceptance [6, 13, 28, 29]. Considering the Uncanny Valley [23] and the balance of social cues and competence, researchers examined the least actual capabilities of an agent, which are to understand its teammates and to react appropriately with adequate length. Eventually, the outcome depends on the contributions of each member including the agent [30–33].

Next, researchers consider the aspect of transparency fostering the understanding of an agent, its behavior and purpose to accept it as a teammate [29, 34, 35]. This allows its human teammates to still critize and improve it [34], which eventually ensures a certain feeling of control as well as an enhancement of the group process and its outcomes [10, 29, 31, 36]. For instance, Gnewuch et al. [33] demand to include error-handling strategies considering potential misunderstandings. Also, Frick [6] suggests to give the human teammates the possibility to influence the computer algorithm's output. Still, due to the fact that much of today's technology including its technical details and mechanisms are very complex, humans cannot rely on a full system transparency and understanding to accept an agent teammate [37–39]. Therefore, it is recommended to establish trust and acceptance right at the beginning. Andras et al. [40] suggest making use of an explainable AI. An explainable AI will introduce itself in advance of starting the hybrid collaboration process. Thereby it will give its teammates insights into its behavior covering the how and the why [40].

At last, considering the enhancement of the process outcomes, researchers found out, that agents can contribute to group creativity effects and concurrently avoid negative effects including social loafing and free-riding, evaluation apprehension and production blocking by contributing with its own decisions [36]. However, the competence of an agent is not to be neglected. It involves the knowledge, abilities and skills of a teammate to satisfy the expectations of the other teammates. These expectations refer to the performance, specifically the contributions toward the team goal within the teamwork [35, 37, 38, 41].

As of our research at hand, we aim at the acceptance of an agent as teammate within a hybrid form of group work, specifically CW. Therefore, we consider these findings toward the acceptance of agents in human-computer-interactions, i.e. the application of social cues in terms of CASA and the Uncanny Valley.

4. Theoretical Background

In terms of accepting agents, many researchers and practitioners refer to human characteristcs, and social cues respectively [6, 13, 28, 29, 33]. Here, Nass and Moon [22] developed the **Social Response Theory** based on several previous studies, among others around the CASA paradigm, demonstrating the mindless application of social rules and expectations to computers. With that, they disclose the application of human social categories, social behaviors as well as premature cognive commitments to computers, and refute alternative explanations like anthropomorphism and intentional responses for their studies. They state that "inviduals are responding mindlessly to computers to the extent that they apply social scripts […] that are inappropriate for human-computer interaction". Therefore, "individuals must be presented with an object that has enough cues to lead the person to categorize it as worthy of social responses, while also permitting individuals who are sensitive to the entire situation to note that social behaviors were clearly not appropriate" [22]. Thus, social cues assigned to an agent trigger humans to apply social behaviors and rules towards the agent [11]. Such social cues could be a name, emotions [6] or also typing indicators [42]. The latter also addresses the concept of social presence [42], i.e. an agent is perceived as socially present, aware and conscious [32]. Still, next to the Social Response Theory, researchers also refer to the concept of the **Uncanny Valley** by Mori [23] reasoning the application of less social cues in order to match the human likeness with competence for maximum affinity [23, 33, 42]. As it is quite easy to generate a social relationship between humans and computers, it is recommended to make use of rudimentary but powerful cues instead of developing highly complex agents [11].

5. Objectives of the Solution

To derive MRs for an agent as teammate from theory and real-live problems, we conducted semi-structured qualitative expert interviews along the approach by Meuser and Nagel [24] and analyzed them in the light of the theoretical background [22, 23]. For the selection of experts (E1-E9), we considered nine diverse researchers from the fields of Information and

Knowledge Technology, Human-Computer-Interaction, Psychology and Sociology. The interview guideline included interdisciplinary open questions to reveal the experts' insights in a reliable and unbiased way. The questions asked covered 1) socio-technical factors within human-human- and human-machine-interaction, 2) agents influencing human-machine-interaction with socio-technical factors, 3) desires, demands and anxieties toward the application of agents, and 4) vision and future prospect about the interaction between agents and humans. To analyze the expert interviews, a thematic comparison was conducted along categories [24]. The categories were determined inductively after an initial scanning of the interview transcripts. Thus, the information from the interviews could be extracted and separated into the following categories: Competence, Social Cues and Feedback. Consequently, the experts' relevant remarks were extracted, merged and collocated along the established categories. Eventually, we connected the expert references to the Social Response Theory [22] and the concept of the Uncanny Valley to derive the MRs [23].

Competence: an agent is not expected to have a general human intelligence, but to have a certain expertise in the application area. As such, the agent should be able to enhance and contribute to the group process and its outcomes (MR1) with all the required skills (MR2) (E1, E2, E4-7). Accordingly, its interactions within the group should be transparent, easy to understand and intuitive through an intelligible display (MR3) (E3, E4, E7-9). Referring to the Uncanny Valley, this display does not have to be utterly human. In fact, too much human likeness might raise higher expectations toward the competence of the agent (E2, E4-7, E9). Eventually, the human teammates should have the right expectations and know that their agent teammate acts in their interest (MR4) (E2, E4-7, E9).

Social Cues: specifically regarding the appearance of an agent, referring to Social Response Theory, it is recommended to assign some humanness to the agent, e.g. a name, a face or an emoticon (MR5), as long as the complexity of the agent's functionality matches the complexity of its appearance (MR6) (E2, E4-7, E9). To further encourage social presence (MR7) in the light of Social Response Theory in terms of social cues, graphical typing indicators within the team interactions are useful (MR8). Additionally, as it is beneficial to initially establish trust and an emotional relationship between the humans and the agent in order to jointly work toward a common goal (E2, E4, E5), transparency about the agent's purpose and processes is required (MR9). Therefore, the schema of childlike characteristics might be of interest (E4). Hence, an agent introduces itself and asks for support within the collaboration (MR10) (E3, E7). As a result, the human

teammates do not expect the agent to not make any mistakes. This approach resembles self-deprecation, e.g. the agent knows, that it is an agent (E5). Considering Social Response Theory, making use of an explainable AI with self-depreciating and childlike characteristics aims at enocuraging social responses toward the agent leading to a closer and more emotional relationship.

Feedback: in the light of Social Response Theory, there are a few underlying characteristics of an agent, which trigger humans' social responses. The first aspect is interactivity covering responses based on inputs. Despite the writing process, human teammates should have the possibility to understand and control the situation (MR11), i.e. they are able to give feedback and influence (MR12) or even intervene, rectify and amend the agent's contributions at any time (MR13) (E2-5). Therefore, the agent's contributions need to be exposed for criticism and improvement (MR14). This feature is crucial for mutual learning benefits of both humans and agent. Following, an agent teammate should also show an interest in the human teammates. This is possible by giving it the same ability to give feedback (E1, E2) covering a second aspect for social responses: the filling of human roles. Therefore, the agent needs to react appropriately (MR15) by making the right decisions (MR16). With that, all teammates should be able to equally contribute to the process outcome (MR17).

Table 1 includes all identified MRs as objectives of the solution.

Table 1. MRs with description and expert reference

Meta-requirements	Expert reference
MR1: The agent enhances the group process and its outcomes.	E1, 2, 4-7
MR2: The agent has all skills to contribute to the team goal.	E1, 2, 4-7
MR3: The display of the agent is intelligible for its teammates.	E3, 4, 7-9
MR4: The human teammates have the right expectations and know that the agent teammate acts in their interest.	E1-7, 9
MR5: The agent is humanoid owning a name and lifelike characteristics.	E2, 4-7, 9
MR6: The agent remains a balance of social cues and competence.	E2, 4-7, 9
MR7: The agent is perceived as socially present.	E2, 4-7, 9
MR8: Graphical typing indicators are involved within the team interactions.	E2, 4-7, 9

MR9: The agent's purpose and processes are transparent and disclosed via an informative opening message.	E3-5, 7
MR10: The agent is an explainable AI introducing itself in advance.	E3-5, 7
MR11: The human teammates understand the situation and retain control.	E2-5
MR12: The human teammates are able to influence the agent teammate's output.	E2-5
MR13: The human teammates require error-handling strategies for interventions.	E2-5
MR14: The agent's contributions are exposed for criticism and improvement.	E2-5
MR15: The agent reacts appropriately.	E1, 2
MR16: The agent is able to make decisions.	E1, 2
MR17: All teammates equally contribute to the process outcome.	E1, 2

6. Artifact Design and Development

Based on the MRs, preliminary action oriented DPs toward the incorporation of an agent into a CW process as teammate were developed according to Chandra et al. [43] (Table 2).

Table 2. DPs with corresponding MRs

Design Principles (DP)	Source
DP1: Provide the agent with the capability of domain-specific natural language processing (NLP) in order for the human teammates to feel understood and obtain appropriate contextual contributions, given that its knowledge is trained, but limited to the context of the teamwork application.	MR15, MR16
DP2: Provide the agent with a controllability in order for the human teammates to have the opportunity to intervene and rectify its contributions, given that the modified new contribution of the machine teammate is qualitatively better and more suitable in regard to the group goal.	MR11, MR12, MR13, MR14
DP3: Provide the agent with the ability to react based on the human teammates' contributions by giving feedback to each individual contribution in order for the human teammates to perceive it as socially present, aware and conscious.	MR1, MR2, MR7, MR15, MR16, MR17
DP4: Provide the agent with explainable capabilities introducing itself including purpose and processes in advance in order for the human teammates to have the right expectations and to understand and accept the agent teammate, given that it is still not perfectly trained and may not make appropriate and useful contributions.	MR3, MR4, MR9, MR10, MR11
DP5: Provide the agent with a humanoid identity and social cues in order for the human teammates to perceive it as an equally social teammate, given a balance of social cues and competence.	MR3, MR5, MR6, MR7, MR8

To support CW, we developed a CW process and implemented it on a web platform. The process enables the participants to collaborate in writing a story. We use the process to design and implement an agent as teammate according to the DPs (Q2). The process steps and activities incorporating the agent are as follows.

1) Prepare

The agent introduces and presents itself right at the beginning to clarify its intended role as a teammate. It explains how it generally works for transparency. Next to its name it also has a picture (DP2, DP4, DP5).

2) Write Sentence

After that, the iterative part of the process starts: the first participant writes a sentence, which extends the story. Here, the agent is included in the order of the participants. When it is its turn, the agent processes the last written sentence to generate a new and contextual appropriate sentence contributing to the story like its human teammates. In doing so, it also takes some time to generate the next sentence. In this waiting period graphical typing indicators show up (DP1, DP5).

3) Extend Story with / without Reaction

There are then three exclusive activities to follow: either claiming, liking or not reacting to the contributed sentence. Exactly like its human teammates, the agent can react to the writer by showing that it likes the released sentence (DP3). The claim-functionality is only available for the human teammates: thereby, they can intervene and demand a rectification of the released sentence of the agent. The agent then generates a better and more suitable new contribution (DP2).

4) Completion

After the first participant's turn, the next participant in line has a turn and writes a sentence. The process ends

when the participants consider the story complete. Additionally, for the overall process, the agent has a picture and a name, which is used all along (DP5).

7. Demonstration

To assess the incorporation of the DPs (Q2) as well as to evaluate the user perception of the agent (Q3) an instantiation of the collaborative writing process was deployed in form of a web-application (Figure 2). This is done by means of Prototyping: developing and evaluating a standalone version, which is quickly available. It is typically limited to the functionalities, which are relevant for the research involving for one thing the feasibility and for another thing the user perception [25].
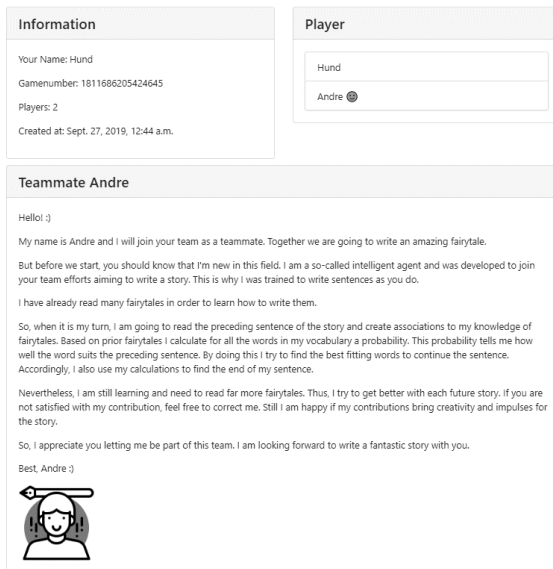
Figure 2. Lobby of the web-application: introduction of the collaborative agent

The DPs were accordingly implemented as follows:

Capability of domain-specific NLP (DP1): For the agent to generate contextually appropriate sentences for the story, it needs to refer to and process the preceding story fragment. The agent is therefore provided with the capability of NLP using Recurrent Neural Networks. Thereby, a word-level language model is developed to predict the probability of the next word in a sentence based on the previous words.

Claim Functionality (DP2): In case the human teammates are not satisfied with the contribution of the agent, they have the opportunity to intervene and claim. Then, the agent has a second chance to rectify its contribution by replacing its generated sentence with a new one. The human teammates therefore have an action panel. Here, they can choose a reaction to each

sentence contributed to the story. If they want to claim, they can choose the "Claim"-Button. For a qualitatively better and more suitable new contribution, the second output of the agent is strictly limited by hard-coding in terms of prototyping. In doing so, grammatically correct sentences giving neutral descriptions, which are likely to fit into any story, can be provided.

Like Functionality (DP3): Just like its human teammates, the agent is able to react by liking the contributed sentences. As soon as the agent liked a sentence, the human teammates receive a pop-up, which states "Andre like the sentence!". The decision on whether the agent likes a sentence or not is made randomly. In terms of prototyping this is a fast and effective way to implement the functionality for the test run in order to be evaluated.

Explainable AI (DP4): The introduction is used to set the right expectations and foster the acceptance of the agent. Here, the agent presents and explains itself. It reveals that it may not contribute appropriate sentences to the story as it is new in this field and still has to learn a lot. However, it is positive and motivated towards its human teammates (Figure 2).

Identity and Social Cues (DP5): To merge into the team as social teammate, the agent is assigned to an identity covering a name, which is Andre, and a picture, which is shown at the end of the introduction in the lobby of the web-application. Its name is used throughout the whole process within the web-application. Thus, the list of participants involved also contains its name. Furthermore, while waiting for the one who has a turn, three animated dots indicate that this person is still writing. In order to perceive the agent as equally social present, the graphical typing indicators also show up when it has a turn including a certain waiting period. As the sentence generation takes some time from approximately ten up to twenty seconds, we did not implement a fixed waiting period. Due to the fact that the NLP capability of the intelligent is still not perfect, only a few social cues are used to not generate disappointment, but to establish a level of trust and sympathy.

8. Evaluation

In order to assess the developed DPs and examine the human teammates' perception of the agent, we conduct a naturalistic ex post evaluation according to Venable et al. [27]. Therefore, four groups of five participants (P1-20) took part in a test run based on the instantiated web-application incorporating the agent. As the CW process does not specify a target group, the participants were selected based on availability, access to a computer and internet connection as well as the

ability to write. Eventually they cover both female and male participants with an age range from around twenty to sixty years. To ensure a smooth induction, each group forgathered at the same place, though the application enables distributed collaboration.

The test runs proceeded without any obstructive problems. Each test run lasted about forty up to sixty minutes including around ten minutes of preparation. After the test run, the participants were asked to reflect on their perception of the agent in qualitative semi-structured interviews. The interviews were aligned to the expert interview concept by Meuser and Nagel [24]. The guideline was designed to address the specific DPs as well as the user perception of the agent. Thus, each participant was asked about the specific instanatiation of each design principle covering their perception and overall satisfaction with the agent. All relevant remarks throughout the interviews have then been extracted, merged and collocated along the DPs and user satisfaction considering the agent and the overall process.

Capability of domain-specific Natural Language Processing (DP1): Most of the participants were not satisfied with the contributions of the agent generated by means of NLP, i.e. it was without context and confusing (P2-6, P8, P12-14, P16-19). However, some contributions were perceived as appropriate (P9, P10, P14, P17, P19) and as interesting (P8). Some participants appreciated that the agent remained in the abstract theme of the story (P7, P11, P20). Though the implementation of DP1 enabled the agent to generate sentences and make contributions to the story, there is much potential for improvement, e.g. it could be trained on a larger text corpus. The development of another language model is also an option.

Claim Functionality (DP2): The claim-functionality was perceived as very good, helpful and important (P1-5, P7-20). For most of the participants it was very easy to claim promptly, especially when sentences did not make any sense (P1, P2, P4, P6, P7, P10, P11, P13, P15-17, P19, P20). Additionally, some stated that it is easier to claim a sentence of an agent than of a human teammate. This is because they knew that it is a computer agent and did not perceive it as emotionally vulnerable (P17, P19, P20). Only one participant admitted to having felt sorry for the agent when claiming a sentence (P12). The number of claims additionally supports the low inhibition level. Only one out of 19 contributed sentences by the agent throughout the four groups was accepted without claiming. With that, the implementation of the claim-functionality can be partly confirmed. On the one hand, the button was accepted very well by the participants, but on the other hand, they might have overly relied on the second sentence.

Like Functionality (DP3): For one thing, the likes were perceived as funny (P10) and cute (P17), but for another thing also very random (P1, P7, P8, P12, P17). Also, three experts did not even recognize the likes (P3, P8, P11). So, while half of the participants did not perceive any difference on the social presence of the agent (P1, P4-8, P11, P12, P16, P17), half of them did perceive a positive effect on the social presence (P2, P7, P10, P13-15, P17-20). Two participants even stressed a humanization of the agent (P9, P15). For a successful implementation of DP3 the distribution of likes needs to be improved.

Explainable Artificial Intelligence (DP4): Due to the self-introduction of the agent, most of the participants had realistic expectations toward it (P4, P7, P9, P11-20). Thus, they were more likely to forgive mistakes of the agent (P12, P19). Even two of the participants stated that their expectations have been exceeded (P1, P10). Still three participants' expectations could not be met. Consequently, they were more disappointed by the agent (P2, P3, P8). For instance, one of them expected the agent to contribute useful complex sentences, which refer to the story and may even include sub-clauses (P8). Another one of them stated that the introduction was too well-formulated to lower any further expectations (P3). As the self-introduction of the agent achieved to set the correct expectations for almost all participants, the implementation of DP4 can be partly confirmed.

Identity and Social Cues (DP5): Regarding the identity of the agent, its name encouraged a more social and personal relation (P3, P6, P7, P9, P12-14, P16, P18-20). The participants within the groups also used its name when talking about the agent instead of calling it a bot. Thus, it could better merge into the group (P10). The picture was perceived as social by only a few of the participants (P3, P7, P14, P16, P19). In fact, the picture was considered impersonal (P2, P12, P15). Besides, there were several participants who did not even recognize nor care about its identity (P4, P5, P17). Though it was still obvious that the agent is not a real human, its identity, especially its name fostered the perception of a social artificial teammate. Thus, most of the participants accepted the agent in its entirety as a computer agent (P1-3, P5, P7, P8, P10-12, P14, P16, P17, P19, P20). Furthermore, as the graphical typing indicators during the waiting period were used for all participants, they had the same effect for the agent. Thus, several participants could better perceive it as a social present teammate thinking about its next contribution. In fact, without a waiting period and graphical typing indicators, the opposite effect would occur (P1, P2, P5, P7, P9, P11, P13, P14, P16, P18, P19, P20). However, two participants just considered the waiting period and typing indicators as loading time for

the agent, not as humanoid thinking time. Three other participants did not really recognize the graphical typing indicators and did not perceive any influence on the social presence (P4, P6, P10). Only one participant stated that the agent was expected to react promptly (P17). As most participants had the right expectations being met by the social cues and competence, DP5 was successfully implemented.

Asking the participants about their **overall perception**, half of the participants perceived and considered the agent a teammate (P5, P9, P10-14, P17, P19, P20). For instance, it contributed to the process like everyone else, i.e. it was part of the process and thereby part of the team (P11, P13). Even though the participants complained about some of the generated contributions (P1, P2, P4, P8, P10-13), it was appreciated that it at least tried to collaborate (P12). Furthermore, many of them enjoyed collaborating with the agent. They considered it fun (P2, P3), entertaining, amusing (P1, P4) and interesting (P4). Additionally, it sometimes diverted the topic by giving new ideas (P1). Still the other half did not consider it a real teammate (P1-4, P6-8, P15, P16, P18). This was mainly because the agent was perceived very inconspicuous (P1, P6, P8, P15, P18).

9. Discussion and Limitations

Overall, five formulated DPs rely on 17 MRs, that were identified through theory and expert interviews and eventually assessed by a test run and reflective interviews with the participants. Based on Social Response Theory [22] and the concept of the Uncanny Valley [23] we formulated DPs toward acceptance of an agent teammate and a complementary synergy of the agent and the human teammates (Q1). With the instantiation, test run and following interviews we could then evaluate the hybrid work (Q2) as well as the perception and acceptance of the agent and its contributions as teammate (Q3).

It was revealed that the five DPs could be partly successfully implemented within the instantiated CW process. As half of the participants in the test run perceived and considered the agent as teammate, the other half did eventually not consider it a real teammate. As DP1 is most criticized and shows much room for potential improvement, this might be the main influence for the overall perception of the teammate. This assumption might be further supported by the evaluation of DP2: most of the contributed sentences by the agent were claimed. Nevertheless, almost half of the participants appreciated the domain-specific contributions and ideas of the agent within the reflective interviews. Aiming at a synergy of both humans and computer agents, future research could define new strategies for dissatisfying contributions of an agent, e.g. grammatically correcting or adjusting them. This way, the human teammates could benefit from the agent's ideas and the agent could learn from the corrections and the adjustments made. However, within our research, we successfully revealed the positive acceptance of the claim-button showing a low inhibition level of the human teammates to easily help and intervene within hybrid work. Here, future research can further examine and elaborate on the right balance of trust and distrust, i.e. balancing the number of human interventions.

As of the Uncanny Valley and the balance of competence and social cues, DP5 was confirmed setting the right expectations for most of the participants. The Explainable AI element supported the right expectation setting for almost all participants, which is why DP4 can be partly confirmed.

Regarding Social Response Theory, we did not only give the agent a primitive identity successfully implemented with DP5, but it was further provided with the ability to like sentences as well as with graphical typing indicators. With the ability to like, DP3 was partly successfully implemented. Though it supported the social presence of the agent, the functionality was more perceived as random. For future research, instead of relying on a random 50 % probability, it could either be implemented by a rule-based-system or even by NLP. Thus, the participants might recognize real preferences of the agent and thereby perceive it as more socially present. Also, the agent might use the received likes to learn from them for future contributions.

All in all, the results show the potential toward a synergy of humans and computer agents in hybrid collaborative work. With a convenient competence and suitable appropriate social cues covering Social Response Theory and the Uncanny Valley, human teammates do not refuse, but accept working with an agent almost perceiving it as real teammate. What is more, a complementary synergy within the hybrid work can be easily achieved with further research work basing on the humans' willingness and low inhibition level to correct and improve the agent with their human intelligence.

Besides the promising results of this research, there are a few limitations to consider. First, the research at hand is only a small, qualitative study of collaborative agents in CW. It does not lead to general and solid conclusions about trust, performance or learning. Hence, it rather serves as a starting point showing the potential of hybrid teamwork. Thereby, it encourages to further conduct detailed studies and to generalize the findings toward a synergy of humans and computer agents in hybrid teams. Furthermore, during the test run the agent was the only teammate, which was not

physically present, i.e. the human teammates could talk and socialize outside the process recognizing voices and gestures. This could have affected the user perception of the agent. For further research, it would be interesting to have all participants at separated locations. Besides, the participants were selected convenience-based and did not have a connection to the practice of CW. With future research the DPs could be tested for their applicability to other CW practices, especially in work environments, where the practitioners' work ethic and job description involves CW. At last, with this research we did not aim to optimize the technological implementation of an agent's NLP capabilities, but to examine the general acceptance, perception and synergy of computer agents in hybrid teams. Still, we assume that with further focus on the development of the NLP capabilities, the utility of such a collaborative agent will probably increase.

10. Conclusion and Contribution

The findings of this paper serve as a starting point for further research in the field of Human-Computer-Collaboration. For this research we performed a test run via an implemented web-application focusing on CW. Thereby, we aim to contribute with prescriptive knowledge [20] towards a "theory of design and action" [44] with MRs and corresponding DPs. Based on Social Response Theory [22] and the concept of the Uncanny Valley [23], we examine related work and conduct expert interviews finding appropriate social cues and capabilities towards the acceptance of a collaborative agent and its contributions as well as the synergy of humans and computer agents in hybrid teams. With that we incorporated an intelligent collaborative agent into a CW process and evaluated its perception and acceptance within a hybrid group work to leverage the potentials of hybrid human-computer-collaboration teams. Eventually, five DPs were established and evaluated to foster a synergy within hybrid teams as well as the acceptance of a collaborative agent as teammate. The DPs should be further tested for their applicability to other hybrid collaborative processes. Additionally, in order to prove the quality of the system in detail, future research might conduct quantitative analyses comparing the design against other forms and test the DPs against control instances within an advanced experimental setting.

11. References

[1] Holzinger, A., "Interactive Machine Learning (iML)", Informatik Spektrum, 39, 2015, pp. 64–68.

[2] Dellermann, D., M. Söllner, P. Ebel, and J.M. Leimeister, "Hybrid Intelligence", Business & Information Systems Engineering, 2019.

[3] Amershi, S., M. Cakmak, W.B. Knox, and T. Kulesza, "Power to the People: The Role of Humans in Interactive Machine Learning", AI Magazine, 35, 2014, pp. 105–120.

[4] Holzinger, A., M. Plass, K. Holzinger, G.C. Crisan, C.-M. Pintea, and V. Palade, "A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop", arXiv:1708.01104.

[5] Holzinger, A., "Interactive machine learning for health informatics: when do we need the human-in-the-loop?", Brain Informatics, 3, 2016, pp. 119–131.

[6] Frick, W., "When Your Boss Wears Metal Pants", Harvard Business Review, June 2015.

[7] Bittner, E., S. Oeste-Reiß, P. Ebel, and M. Söllner, "Mensch-Maschine-Kollaboration: Grundlagen, Gestaltungsherausforderungen und Potenziale für verschiedene Anwendungsdomänen", HMD Praxis der Wirtschaftsinformatik, 56, 2019, pp. 34–49.

[8] Dellermann, D., A. Calma, N. Lipusch, T. Weber, S. Weigel, and P. Ebel, "The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems", in Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019: Maui, Hawaii, USA.

[9] Epstein, S.L., "Wanted: Collaborative intelligence", Artificial Intelligence, 221, 2015, pp. 36–45.

[10] Seeber, I., E. Bittner, R.O. Briggs, G.-J. de Vreede, T. de Vreede, D. Druckenmiller, R. Maier, A.B. Merz, S. Oeste-Reiß, N. Randrup, G. Schwabe, and M. Söllner, "Machines as Teammates: A Collaboration Research Agenda", in Proceedings of the 51st Hawaii International Conference on System Sciences. 2018: Waikoloa Village, Hawaii, USA.

[11] Nass, C., J. Steuer, and E.R. Tauber, "Computers are social actors", in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994.

[12] Martínez, M.A.M., M. Nadj, and A. Maedche, "Towards An Integrative Theoretical Framework Of Interactive Machine Learning Systems", in Proceedings of the 27th European Conference on Information Systems. 2019: Stockholm & Uppsala, Sweden.

[13] Bittner, E., S. Oeste-Reiß, and J.M. Leimeister, "Where is the Bot in our Team?: Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work", in Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019: Maui, Hawaii, USA.

[14] Adams, T., "And the Pulitzer goes to…a computer", The Guardian, 28.09.2015.

[15] Manjavacas, E., F. Karsdorp, B. Burtenshaw, and M. Kestemont, "Synthetic Literature: Writing Science Fiction in a Co-Creative Process", in Proceedings of the Workshop on Computational Creativity in Natural Language Generation. 2017: Santiago de Compostela, Spain.

[16] Ghuman, R. and R. Kumari, "Narrative Science: A Review", International Journal of Science and Research, 2(9), 2013.

[17] Colton, S., J. Goodwin, and T. Veale, "Full-FACE Poetry Generation", in Proceedings of the Third International Conference on Computational Creativity. 2012: Dublin, Ireland.

[18] Rank, S., S. Hoffmann, H.-G. Struck, U. Spierling, and P. Petta, "Creativity in Configuring Affective Agents for Interactive Storytelling", in Proceedings of the Third International Conference on Computational Creativity. 2012: Dublin, Ireland.

[19] Monteith, K., T. Martinez, and D. Ventura, "Automatic Generation of Melodic Accompaniments for Lyrics", in Proceedings of the Third International Conference on Computational Creativity. 2012: Dublin, Ireland.

[20] Gregor, S. and A. Hevner, "Positioning and Presenting Design Science Research for Maximum Impact", MIS Quarterly, 37(2), 2013, pp. 337–355.

[21] Peffers, K., T. Tuunanen, C.E. Gengler, M. Rossi, W. Hui, V. Virtanen, and J. Bragge, "The Design Science Research Process: A Model For Producing And Presenting Information Systems Research", in Proceedings of the International Conference on Design Science Research in Information Systems and Technology. 2006: Claremont, CA.

[22] Nass, C. and Y. Moon, "Machines and Mindlessness: Social Responses to Computers", Journal of Social Issues, 56(1), 2000, pp. 81–103.

[23] Mori, M., "The Uncanny Valley", IEEE Robotics & Automation Magazine, 2012.

[24] Meuser, M. and U. Nagel, ExpertInneninterviews - vielfach erprobt, wenig bedacht, VS Verlag für Sozialwissenschaften, Wiesbaden, 2002.

[25] Wilde, T. and T. Hess, "Forschungsmethoden der Wirtschaftsinformatik: Eine empirische Untersuchung", WIRTSCHAFTSINFORMATIK, 49, 2007, pp. 280–287.

[26] Leimeister, J.M., Collaboration Engineering: IT-gestützte Zusammenarbeitsprozesse systematisch entwickeln und durchführen, Springer-Verlag, Berlin Heidelberg New York, 2014.

[27] Venable, J., J. Pries-Heje, and R. Baskerville, "A Comprehensive Framework for Evaluation in Design Science Research", in Proceedings of the 7th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2012), Las Vegas, NV, USA, May 2012.

[28] Przybilla, L., L. Baar, M. Wiesche, and H. Krcmar, "Machines as Teammates in Creative Teams: Digital Facilitation of the Dual Pathways to Creativity", in Proceedings of SIGMIS-CPR. 2019: Nashville, TN, USA.

[29] Seeber, I., L. Waizenegger, S. Seidel, S. Morana, I. Benbasat, and P.B. Lowry, "Reinventing Collaboration with Autonomous Technology-Based Agents", in Proceedings of the 27th European Conference on Information Systems. 2019: Stockholm & Uppsala, Sweden.

[30] Kirchkamp, O. and C. Strobel, "Sharing responsibility with a machine", Journal of Behavioral and Experimental Economics, 80, 2019, pp. 25–33.

[31] Norman, D., "Design, business models, and human-technology teamwork", Research-Technology Management, 60(1), 2017, pp. 26–29.

[32] Robb, A. and B. Lok, "Social presence in mixed agency interactions", in IEEE Virtual Reality (VR). 2014: Minneapolis, MN.

[33] Gnewuch, U., A. Maedche, and S. Morana, "Towards Designing Cooperative and Social Conversational Agents for Customer Service", in Proceedings of the 38th International Conference on Information Systems. 2017: Seoul, South Korea.

[34] Trujillo, A.C., I.M. Gregory, and K.A. Ackerman, "Evolving Relationships between Humans and Machines", IFAC-PapersOnLine, 51(34), 2019, pp. 366–371.

[35] Gulati, S., S. Sousa, and D. Lamas, "Modelling trust in human-like technologies", in Proceedings of the 9th Indian Conference on Human Computer Interaction. 2018: Bangalore, India.

[36] Siemon, D., L. Eckardt, and S. Robra-Bissantz, "Tracking Down the Negative Group Creativity Effects with the Help of an Artificial Intelligence-like Support System", in Proceedings of the 48th Hawaii International Conference on System Sciences. 2015: Kauai, Hawaii, USA.

[37] Cheng, X., G. Yin, A. Azadegan, and G.L. Kolfschoten, "Trust Evolvement in Hybrid Team Collaboration: A Longitudinal Case Study", Group Decision and Negotiation, 25(2), 2015.

[38] Yu, K., S. Berkovsky, R. Taib, J. Zhou, and F. Chen, "Do I trust my machine teammate?: an investigation from perception to decision", in Proceedings of the 24th International Conference on Intelligent User Interfaces. 2019: Marina del Ray, California.

[39] Morris, A. and M. Ulieru, "FRIEND: A human-aware BDI agent architecture", in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. 2011: Anchorage, Alaska, USA.

[40] Andras, P., L. Esterle, T.A. Han, M. Guckert, and P.R. Lewis, "Trusting Intelligent Machines: Deepening Trust Within Socio-Technical Systems", IEEE Technology and Society Magazine, 37(4), 2018, pp. 76–83.

[41] Elson, J.S., G.S. Ligon, and D.C. Derrick, "Examining Trust and Reliance in Collaborations between Humans and Automated Agents", in Proceedings of the 51st Hawaii International Conference on System Sciences. 2018: Waikoloa Village, Hawaii, USA.

[42] Gnewuch, U., S. Morana, M.T.P. Adam, and A. Maedche, ""The Chatbot is typing …" - The Role of Typing Indicators in Human-Chatbot Interaction", in Proceedings of the 17th Annual Pre-ICIS Workshop on HCI Research in MIS. 2018: San Francisco, CA, USA.

[43] Chandra, L., S. Gregor, and S. Seidel, "Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions", in Proceedings of the 48th Hawaii International Conference on System Sciences. 2015: Kauai, Hawaii, USA.

[44] Gregor, S., "The Nature of Theory in Information Systems", MIS Quarterly, 30(3), 2006, pp. 611–642.