

Conversational Agents at Work: How Can HCI Help Facilitate Better Social Interaction in the Workplace?

Master Thesis Summary

The premise of this study stems from our pre-thesis, in which we conducted a systematic literature review on AI support systems in the labour market, within the HCI community. The findings from our pre-thesis highlighted a research gap, where the majority of research was conducted on intellectual tasks, few on social and none regarding physical. This research gap became the foundation and motivation for our master thesis. In this master thesis study, we investigate the barriers of adopting AI support systems for social interaction in group meetings. In addition, we aim to identify the needs and expectations of the end-users for incorporating a conversational agent into group meetings.

Through qualitative data collecting methods, we conducted initial interviews with master students from Aalborg University, and scenario-based evaluation with relevant employees from the labour market. The initial interviews provided us with insight and data, from which we created our scenarios. The scenarios were approximately one min long each, revolving around three overall themes that encapsulates different challenges and barriers for facilitating group meetings, in relation to our initial interviews. The scenarios were then evaluated with employees from a windmill manufacturing company that represents the end-users, to further generate data regarding conversational agents' impact on social interactions in group meetings. For analysing the collected data, we made use of the Reflexive Thematic Analysis for generating codes and themes, to systematically identify what challenges, needs and expectations.

In our findings, we discovered the usage of devices during meetings caused disruption. Reducing menial tasks could have a positive effect on efficiency, by making the conversational agent transcribe and send recap for meetings, through the role of a “secretary”. In addition, facilitating time management and agenda were highly sought after. The conversational agent received mixed feedback regarding its features for reducing miscommunication and inclusion of participants. Inclusion of participants were to be facilitated by the meeting holder and not the CA, unless mutual agreement. It was discovered that augmenting the facilitator would obtain a higher acceptance, without putting pressure on the participants. These findings were presented to eight companies that attended our stand at a design exhibition, facilitated by Aalborg University. These findings shared similarity to what all the eight companies had experienced in their own group meetings. They shared fondness and generally positive attitude towards incorporating a conversation agent to facilitate social interaction in their meetings, highlighting similar needs and expectations that were presented in our findings.

In our discussion, we reflect upon the findings that were discovered, in relation to the qualitative methods used. Through our systematic approach, we identified eight challenges regarding the facilitation of meetings, and six challenges when trying to adopt CAs in social interaction. Furthermore, we reflect on the diversity that the participants sample provided from different occupations, as well as the generated themes they produced.

Based on our findings and discussion, we propose seven recommendations to HCI practitioners and researchers that aim to highlight some of the challenges when adopting AI support systems for social interactions. For example, we highlight the importance of including participants during meetings, the impact that individual adaptability can have on adoption of conversational agents and the reduction of menial tasks to enable the end-users to focus on their main goal. Lastly, we hope to contribute to future research by accommodating the research gap of AI support systems in the labour market, within the HCI community.

Conversational Agents at Work: How Can HCI Help Facilitate Better Social Interaction in the Workplace?

THOMAS ILDSTRØM, Aalborg University, Denmark

ALEXANDER RATH HAKONSEN, Aalborg University, Denmark

In this master thesis, we present the findings from interviews and scenario-based evaluation on conversational agents, in relation to the social interaction, that occurs within group facilitation. We extend upon previous knowledge from conducting a literature review on published articles from the ACM conferences CHI and CSCW, regarding artificial intelligence and machine learning support systems, in the labour market. Based on this, we wanted to investigate the gap discovered in the research regarding social tasks and AI support systems. We identified the challenges and barriers of adopting AI in this context, by systematically analysing five interviews and four scenario-based evaluation from master thesis students and employees in the labour market using *Reflexive Thematic Analysis*. This resulted in fourteen challenges regarding the adoption of AI in the social context of group meetings. We provide seven recommendation for HCI researchers and practitioners, to help overcome these challenges and motivate contribution to the research gap within this area, in the future.

CCS Concepts: • **Human-centered computing** → **Artificial Intelligence**; **Labour Market**; *Conversational Agents*; HCI; • **Human-centered Computing** → Social Interaction.

Additional Key Words and Phrases: Conversational Agents, Machine Learning, Artificial Intelligence, Labour Market, AI Support Systems, Group Facilitation, Social Interaction

ACM Reference Format:

Thomas Ildstrøm and Alexander Rath Hakonsen. 2022. Conversational Agents at Work: How Can HCI Help Facilitate Better Social Interaction in the Workplace?. In *IXD Thesis*. ACM, New York, NY, USA, 23 pages.

1 INTRODUCTION

Traditionally, a dialogue is manifested through a conversation between two or more people. In this digital age, information technology has created opportunities for obtaining human-computer dialogues as well, namely dialogue systems. A dialogue system is a computerised system which facilitates interaction with humans in a natural language [3]. Furthermore, the advancement in Artificial Intelligence (AI) has given dialogue systems even more possibilities by enhancing the "conversation", resembling a more human-natural conversation.

Conversational AI systems can create human-AI communication through Natural Language Processing (NLP). Originating from their popular predecessor, chatbots, which only operates from a prefix set of possibilities, the conversational AI system's greatest potential is to mimic human language automatically through computational linguistics with the implementation of Machine Learning (ML) and NLP. Often depending on the level of embodiment, domain and capabilities, these conversational AI systems' goes by a variety of different names, such as virtual agent, intelligent personal assistant and conversational agent. In this paper, we will mainly refer to these systems as Conversational

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.

Manuscript submitted to ACM

Agents (CA) which is a general term used within research to encapsulate modern conversational AI systems, that uses ML and NLP to understand user input and facilitate natural language responses [9, 11, 24, 30].

In our pre-thesis study, we conducted a literature review on AI support systems in the labour market. The scope was within the field of Human-Computer Interaction HCI by extracting published papers from the Computer-Human Interaction (CHI) and Computer-Supported Cooperative Work (CSCW) conferences. We performed a search query that revealed 222 relevant papers from both conferences combined. To assess the papers, we used a comprehensive taxonomy developed by Fernández-Macías and Bisello (2021) that provides detailed categorisation for different labour tasks [13]. The taxonomy consist of three main groupings; *physical*, *intellectual*, and *social tasks*, each consisting of subcategories that correlates to different settings of labour tasks. We systematically reviewed and identified 54 relevant papers, published between 1981 and 2021, and assessed their labour tasks, participant sample, method of work, geographical location, and the adaptation of technologies. In relation to the labour tasks, we identified 78 studies on intellectual tasks, two on social tasks and none regarding physical tasks. The findings from our study therefore highlight a rather unequal distribution of research on AI support systems in the labour market, within the HCI community [15].

While we can see a rise in the usage and development of digital assistants, the research is lacking in the area of AI support systems in the labour market [26]. This lack of research created questions regarding the adoption of AI support systems in relation to physical and social tasks. In our field of expertise as interaction designer, we knew that facilitating and testing physical tasks would be too demanding and out of our grasp, hence why we dismissed this area of work tasks. In contrast, social tasks were more approachable based on the already existing AI support systems that have been adopted commercially, but also more manageable in relation to time, technology, test, and evaluation. Furthermore, social interaction is important in today's society, as Adler quoted that; "[...] *each individual needs to find a place among others that allows them to cooperate*.". It is important that we have healthy work relationships, and that we feel heard in our environment, to ensure job satisfaction and increase motivation [28].

To narrow our investigation, we have adjusted the scope to group meetings, since it requires a certain level of social interaction between people. Fernández-Macías and Bisello (2021) defines social tasks as the tasks aimed at the interaction with other people [13]. Consequently, we broaden the definition of social task to be social interaction supported by AI systems for the context of this study. In addition, we have chosen a conversational agent as an artefact for conducting test and evaluation regarding AI support systems. The motivation for this chosen type of technology is based on the familiarity and success that CAs have, making it one of the most used AI supported technologies worldwide. In 2019, it was reported that more than 3 billion digital voice assistants were in use, while forecasting that the amount would surpass 8 billion by the end of 2024 [20]. However, recent research also indicates that the usages of CAs at homes decline over time, highlighting challenges for adoption regarding voice user interface (VUI) at home [10]. Furthermore, a systematic literature review on CAs in business by Bavaresco et al. (2020) showcases that research revolving around CAs is generally towards Q&A as their primary goal, highlighting no research on CAs in group settings [4].

We find this curious, since CAs are a well-establish technology that helps users complete tasks in various settings every single day. Especially, when taken into consideration that support systems which facilitates better group meetings has been a popular topic, with its own field of research that rose in popularity in the 1990's, namely Group Support Systems (GSS) [2, 5, 14, 16, 29, 32]. This further signifies that support systems for group facilitation is yet to be exploited within the HCI community, in relation to the gap of research on AI support systems for social tasks, which we aim to address in this paper.

We investigate the barriers and challenges that may hinder efficient social interaction in group meetings. Through data collection and evaluation using a conversational agent artefact, we aim to achieve insight on the research gap of AI support systems and social interactions in the labour market. This leads to our following research questions:

- *RQ 1: What are the barriers for adopting Human-AI support systems regarding social interaction in group meetings?*
- *RQ 2: What are the needs and expectations of end-users in relation to conversational agents to support social interaction?*

Our findings highlight a high usage of devices during meetings that causes disruption to the communication. Furthermore, it was identified that reducing mundane tasks could have a positive impact on efficiency. Transcribing and taking notes for meetings were found very time-consuming, and could cause loss of information. Features for time management, agenda and the aforementioned were highly sought after, and received positive feedback. The CA received mixed feedback in regards to both its features for reducing miscommunication, and inclusion of participants. Inclusion of participants was to be facilitated by the meeting holder and not the CA itself. It was discovered that augmenting the facilitator would obtain a higher acceptance, without putting pressure on the participants.

In this paper, our main contributions are: (1) *Seven recommendation to HCI practitioners and researchers, to create a better understanding of the challenges, needs and expectations of adopting Human-AI support systems for social interaction in group meetings.* For example, we highlight the importance of engaging participants during meetings, and the impact of individual adaptability of conversational agents for adoptions; (2) *Contribute to the open research gap of AI support systems across the labour market.*

2 RELATED WORK

In this section we provide an introduction to the impact conversational agents may have in business' and the importance social interaction have in the labour market. We present earlier work examining the use of CA in group facilitation and how the adoption of AI can be improved. Subsequently, we highlight some of the earlier work done within the area of CA regarding social interactions, features to improve cooperation and facilitate better group dynamic.

2.1 Conversational Agents in business

Conversational agents are advanced dialogue systems which implements a variety of applications for usage. By default, they are set to recognise the users' speech, text, gesture or other form of interaction, then process and respond to that information by conveying it back to the user, often in a similar fashion. A conversational agent can be designed to take many different roles, all depending on the their respective domain [18]. Some of the most well-known CA are Google Home and Apple's Siri, both sharing the top spot in a 2019 survey by 36%, where Amazon Alexa was used by 25% and Microsoft Cortana by 19% [19]. Globally, more than 3 billion digital voice assistants are in use, while forecasting the amount would almost threefold by the end of 2024 [20].

In our previous study, it was discovered that very little research was conducted on AI support systems and social interaction in the labour market, within the HCI community [15]. Therefore, we set to investigate what research has been conducted on CA in businesses generally, to broaden the scope of our understanding. A systematic literature review by Bavaresco et al. (2020) encompasses literature on CA in business of the last decade through a defined methodology [4]. The aim of their study was to review CAs in the business domain, focusing on machine learning. Here they discovered that no study combines self-learning, personalisation, and generative-based responses. These are rather important aspects in terms of CA and machine learning. Self-learning is needed to improve the CA interactions, personalisation is

a feature for adapting to tasks for the specific user, and response generation is the technique for response dialogue in CA. Furthermore, the majority of research for business domains was conducted for commerce, having their primary goals as (Q&A) and customer support. No research specifically addressed CA in the form of group settings. Out of 58 studies only five were regarding internal support as their primary goal, and only one was about some form of social interaction [4].

The article that had elements of social interaction were a study by Zhao et al. (2018) regarding a social intelligent negotiation dialogue system, namely SOGO. The proposed system can negotiate with people while building social bond. This is achieved by a two-phase computational model; a task phase and social phase. By implementing a dialogue model for deciding task moves, and social language generator for social skills, they achieved a system that behaves in sync with the user. This is mainly due to the system developing human-like qualities, resulting in achieving higher agreement rate, securing beneficial deals, while also building interpersonal rapport. They further draw the distinction that strategies for building close relationship in human-human interaction is also possible to a human-agent. This article showcases that generating task-related social moves is crucial for maintaining conversational strategies, which relates to the adoption of CA [31].

To further investigate CA in business, a study by Ling et al. (2021) conducted a systematic review of what factors that influenced users' adoption and use of CA. Here they identified 23 factors that could be categorised into; usage-related, agent-related, user-related, attitude, evaluation, and other factors. User acceptance is generally tied to usage benefits, which agent and user characteristics contribute to. Here utilitarian benefits was the most reported factor when influencing the users in relation to usage settings. When looking specifically at social interactions, Ling et al. (2021) identified that in terms of usage, CAs that are designed to offer companion-, and relationship building influenced the users for acceptance and adoption. Implementing human-like cues and communication modalities by having roles of appearance, empathetic and social abilities was found to significantly influence elderly users to adopt the CA [21].

These papers help encapsulate some of the important aspects when it comes to the adoption of CA in businesses, regarding acceptance, features and social capabilities.

2.2 Social Interaction in the labour market

The process of sharing knowledge and shared understanding from one person to another is known as communication [23]. The quality of communication is determined by factors in the communication process. The transmitter and receiver are two essential components of every communication transaction. The communication is started by the transmitter. The recipient is the person who receives the communication. The transmitter encodes the concept by composing a message using *words, symbols, or gestures*. The result of encoding, could take the shape of *spoken, nonverbal, or written language*, is the message. The message is conveyed across a medium or channel, which serves as the communication's carrier. A *face-to-face* discussion, a *phone call*, an *e-mail*, or a *written report* can all be used as the medium. The receiver converts the received message into useful data. Anything that disrupts the message is considered noise. Noise includes different views of the message, linguistic obstacles, interruptions, emotions, and attitudes. Finally, feedback emerges when a receiver replies to a message and returns a new message to the sender. The sender receives feedback to evaluate if the message was received and comprehended [23].

When interacting in social contexts, we might encounter different barriers or noises. Lunenberg (2010) lists a few; *Process Barriers, Physical Barriers, Semantic Barriers* and *Psychosocial Barriers*. All these barriers may influence the quality of the communication and therefor hinder the social interaction in some way [23].

Tschan and Semmer (2004), split social interaction into two categories *Task-Related* and *Private* interactions at work [28]. *Task-related* interaction has its focus on completing a task, e.g. a doctor and a nurse interact with each other, discussing how to best help the patient or explaining features of a product to a customer. During these interactions there may occur what they call *Private* interactions, or what we can refer to as small talk. The main interaction is about the task at hand, and there would be no interaction if not for this task. Secondly, they describe *Private* interactions at work, as a more informal way of talking to each other. Discussing "public" events, such as the Russian invasion of Ukraine or how they spent their weekend [28]. Social interaction has a big impact on how we work, and how we interpret the "message" can be prohibited by many aspects. The literature suggest, when researching social interaction in the labour market, that there always is some sort of communication involved.

As described in the introduction, there were only discovered two papers revolving the area of AI support systems and social tasks. Firstly, a paper by Suh et al. (2021) investigating the importance that AI may have in influencing human social dynamics. They observed 30 people (15 pairs) write a musical phrase in pairs, both with and without AI. The findings suggest that AI can influence social dynamics during creativity [27]. Secondly, a paper by Shamekhi et al. (2018) was discovered during our literature review. The paper exams the value of embodiment in CAs, more specifically on group facilitation tasks [25]. These papers help enlighten certain areas in the research of AI in social tasks, and how it might be facilitated in this context.

3 METHOD

We initially interviewed five participants to create and understand which challenges they might encounter in the context of social interactions when conducting meetings in the workplace. We used semi-structured interviews to gain insight into participants thoughts, experience and opinions about group meetings. We applied Reflexive Thematic Analysis as created by Braun et al. (2019) to analyse these interviews and create scenarios based on the themes generated from the analysis, see 3.1.3 [6, 7]. Subsequently, these scenarios were developed into small videos, approximately 1 minute each, and presented to four employees at a larger international windmill manufacturer in Aalborg. After each scenario, a small discussion, guided by a flexible interview protocol and a debriefing with the participants, was conducted. This was to ensure sufficient collection of data for further analysis.

3.1 Thematic Analysis

To analyse the data provided from the interviews, we made use of the method *Reflexive Thematic Analysis*, developed by Braun and Clarke [6]. Originally, it was only called *Thematic Analysis*, however in 2019 they changed the name to *Reflexive Thematic Analysis* [7, 8]. Reflexive Thematic Analysis (hereafter referred to as Reflexive TA) is a widely used qualitative analytic method within psychology studies [7]. The method is mainly applied to sets of text, interviews or transcripts. The concept of Reflexive TA is to reveal patterns or "themes" across qualitative datasets, with the purpose of answering both narrow and broad research questions [7].

The method has become increasingly popular within the research area of HCI and can according to Brulé (2020), help HCI researchers reach a deeper understanding of their end-user. Reflexive TA can provide HCI researchers with an analytical assessment of statements, views, opinions, knowledge, experiences or values, and what impact they may have, beyond the prototype, and not just summarise how users liked or disliked it [8].

Braun and Clarke (2019) describes three schools of TA, *Coding Reliability*, *Reflexive TA*, and *Codebook TA*. *Coding Reliability*, is an approach that, according to Boyatzis (1998), is a translator between qualitative and quantitative researchers [7]. Researchers using this approach share values with those of the quantitative school in the matter of

reliability and replicability of the observations. Secondly, the Reflexive TA approach “[...] is conceptualised as a fully qualitative approach with data collection and analysis techniques underpinned by a qualitative philosophy or paradigm” [7]. Lastly, the Codebook TA approach is mapped between the two other approaches and revolves around the concept of having some if not all the themes predetermined before the full analysis.

We chose the Reflexive TA approach, since we wanted to conceptualise themes derived from our initial interviews, which could be used to create different scenarios. Reflexive TA fitted our analytic approach, since we did not have clear themes or codes to apply to the dataset.

To discover these different patterns there are four approaches to consider before analysing data. The method distinct between an inductive or deductive approach, inductive being an approach where the data determine the themes, and deductive involves approaching the data with some sort of predetermined themes, found in the data. Subsequently, Braun and Clarke (2006) also talk about semantic and latent approaches to the data. Semantic is where the evident is explicit and latent involves reading into the subtext [6]. In our initial interviews, we utilised the inductive approach in combination with the semantic approach, due to our lack of knowledge on barriers and interruptions that might occur during group meetings. Consequently, we wanted the data to develop our themes based on the information retrieved from the interviews. In our analysis of the interviews we followed the six steps provided by Braun et al. (2019); *familiarisation with the data, generating codes, constructing themes, revising and defining the themes, and producing the report of the analysis* [7].

3.1.1 Familiarisation. After the interviews were conducted we cycled through notes and audio recordings from the interviews, to become more familiar with the data. We began taking notes of potential points of interest for the analysis, and these points were then timestamped.

3.1.2 Codes. After getting familiar with the data, we systematically code each interview, separately. We each went through the interviews searching for representations of talk, that had meaning to the topic of social interaction, barriers, interruptions, facilitation and digital assistants. The interviews were not fully transcribed, but instead noted in timestamps where we found points of interest, along with a transcription of that particular point and related code. These were put in a chronological order in a shared Google Document. Afterwards, we compared notes and discussed the initial codes and iterated upon these. Codes with high similarity were condensed into single codes. This process generated 43 single codes from the initial interviews, which is going to be applied in the next part of our analysis.

3.1.3 Generating Themes. Derived from the initial interviews these 43 codes generated eight themes, see section 4. The process is to organise these codes in a way which reflect the similarities in what our participants expressed during the interviews. We made use of the approach given by Braun et al. (2019), and started to construct themes, or what they identify as *candidate themes* [7]. They describe, that themes are not fully emerged from the data, but a process of constructing them is required [7]. We started to connect similar codes, along with their respective data into *clusters of meaning*, which is defined as *codes as building blocks* [7]. We discussed these derived themes, and their representation of the data, some were discarded, others merged into new themes. Finally, we arrived at three overall themes; *Barriers*, *Common Courtesy* and *Facilitation*. These three themes serves as the foundation for the scenario-based evaluation, in the next part of our analysis.

3.2 Scenario-Based Evaluation

As a part of the process of presenting the HCI community with recommendation in the area of adopting AI into social tasks and interaction in the work space, we created a scenario-based evaluation. We wanted to uncover more areas of group meetings. Based on our initial findings we created, as mentioned earlier, three short videos to represent the three main themes, as seen here: <https://youtu.be/aNNpBPE0Khc>. Looker et al. (2008) describes scenarios, as a narrative of an interaction between user and system [22]. Since a functioning prototype is not developed, we wanted to create a setting in which we would be able to obtain data. The videos present the viewer with three challenges; *barriers*, *common courtesy* and *facilitation*. The three videos were created in *Adobe Premiere Rush*, and acting was done by the authors, with help from colleagues. Before each scenario was presented, a small briefing on the topic was conducted, in order to set the context.

These scenario-based evaluations served two purposes. The first purpose was to highlight some of the barriers and challenges the initial participants had, by facilitating some resemblance they might have from their own experiences, when conducting group meetings. Could the participants presented with the scenarios recognise the same challenges as derived from our initial interviews? Secondly, what was their perception of our concept, what needs, and expectation would they have for this type of technology? To answer this, we again applied *Reflexive Thematic Analysis* to ensure an explicit understanding of the data collected [7, 8]. However, the second time, we used a mix of the semantic and latent approach while analysing the data. This was to identify underlying features, which may have been hidden "between the lines". Otherwise, the same procedure was followed in terms of Reflexive TA. Below is a short description of the three scenarios.

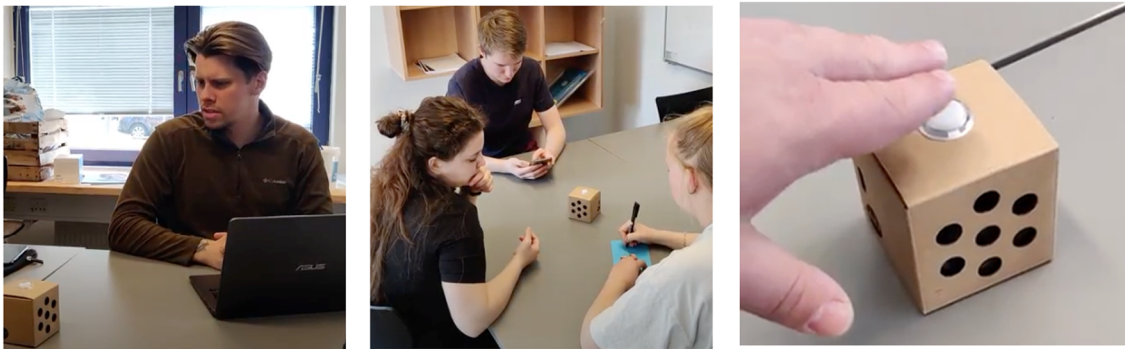


Fig. 1. Snippets from the scenarios created from the initial interviews

3.2.1 Scenario One: Barriers. In the first part, the narrator starts by introducing Jens, who is a fictive person in this scenario. Jens is a UX designer who has arranged a meeting with the engineering team to receive feedback on some designs he made. When presenting his designs, Jens uses frequent terminology within his field, since it is what he works with on a daily basis however, many of the engineers might not be accustomed to them. The CA registers this and gives a signal that a lot of terminologies are used, therefore notifying Jens and the others. Jens then uses the CA to give a brief explanation of the terms used to ensure a mutual understanding of the subject.

In the last part, we are continuing in the same meeting. Here another fictive character is introduced, Tim. Tim is very engaged and can be incredible dominant during the meetings. The meeting discussion is primarily between Tim

and another participant, where Tim is loud and indirectly controls the conversational flow. In the background is Jens, who is quiet and does not participate in the discussion. The CA can through voice recognition identify, that Jens has been passive for a while and intervenes by stating “*Jens, I have noticed that you have been quiet for a while. Do you have any inputs?*”. This startle Jens but gives him incentive to participate, which results in a lot of valuable inputs. He just found it difficult to participate since Tim often could be very dominant during meetings.

3.2.2 Scenario Two: Common Courtesy. The scenario takes place during a meeting with a thriving discussion and energy. The flow of the meeting gets disrupted when a participant start to receive notification on his phone, resulting in him losing attention and starts to interact with his phone instead. The participant is a project leader, who has been in the company for many years. Therefore it can be difficult to confront him about this disruptive manner. The CA therefore interrupts the meeting by reminding people that: “*Just as we keep our phones idle when attending the cinema, remember to show the same courtesy for your colleagues during the meetings. Answer only in important situations*”.

3.2.3 Scenario Three: Facilitation. The scenario takes place at the start of a meeting, were the CA is asked to highlight the agenda, and give a recap from the last meeting. During the meeting the CA manages the defined time to ensure every important point of the agenda is accounted for. After reaching one of the points of the agenda, the CA initiate a conversation by asking “*In relation to the given agenda, we have 30min presentation, 5min break, 30 min discussion. Afterwards, the meeting is concluded. We have reached the first 30min. Have you concluded the presentation?*”. The facilitator tells the CA that they are not done yet, in which the CA replies “*Okay, that is notified. Please give me an update when you are done, since we have many interesting points to accommodate together!*”. After the meeting is concluded, the facilitator tells the CA to give a recap of the meeting, ensuring that the agenda was reached and the meeting successful. Afterwards, the CA sends the recap to every participant to prevent loss of information.

4 FINDINGS

In this section we will present our findings from our five initial interviews (P1-P5), and our four scenario-based evaluations (P6-P9), see table 1. During our analysis, it was discovered that every participant had previous knowledge regarding digital assistants, but not conversational agents, and only two participants used it regularly. In the first part of this section, we present the findings from a our initial interviews with five master students from interaction design and their views, experiences and values on group meetings and what challenges that may create. The second part of our findings address the discoveries that were uncovered during the presentation of three scenarios with four participants from a large, international, windmill manufacturer.

4.1 Initial Interviews

The semi-structured interviews with fellow interaction design master students yielded many interesting thoughts, experiences, and opinions in terms of group meetings. The majority of answers were homogeneous and related. This correspond well with the use of Reflexive TA, as Bruan et al. (2019) state that the method is suited for smaller studies with five to six participants where the answers are relatively homogeneous, and the data is rich [7]. Systematically going through the data, we generated a sample of codes, that would be converted into eight candidate themes. Through iteration, we further convert these themes into three overall themes: *Barriers*, *Common Courtesy*, and *Facilitation*. These three themes covers all the sub-themes and their codes, and served as structure for creating scenarios to be evaluated upon.

Nr	Gender	Age	Occupation
P1	M	25	Master Thesis Student
P2	M	25	Master Thesis Student
P3	F	26	Master Thesis Student - Full time UX designer
P4	M	27	Master Thesis Student - Part time UX designer
P5	M	27	Master Thesis Student
P6	M	34	Process Engineer
P7	F	60	Engineering Change Management
P8	F	27	Team Leader
P9	F	34	Project Leader

Table 1. Participant Sample

4.1.1 *Barriers*. Barriers were the only theme that did not consist of any sub-themes. However, it was the theme with most coding, therefore making it rather significant for converging into an overall theme. Some of the occurring codes that formed this theme were; *language, miscommunication, interpersonal, engagement, and meeting size*.

Part of the coding revolving language were explicitly explained by P1 who stated that working in a non-native setting could be difficult; *"We just had a short period with a Danish and foreign supervisor, and it is just two very different ways to facilitate meetings"*. Something similar had occurred for P5 as well. Even though both participants are fluent in English, it still creates barriers for conducting meetings in English, as further amplified by P5; *"I have also experienced in terms of speaking English, that we often can be talking past each other"*. Here P2 also states that; *"Language and misunderstanding are the biggest barriers I encounter"*.

Miscommunication is a reoccurring topic that manifested a lot of barriers. P1 explained that in his professional work, he encounters it to be problematic, by stating; *"I gave some suggestions based on common knowledge in relation to the project, but they did not really understand any of it"*. P4 when asked about the most dominant problem in meetings stated; *"I would say it is understanding of what I am saying. Is it received correctly? It also happens that I make use of terms that some of my other colleagues do not understand"*.

Interpersonal and meeting size were discovered as barriers, and can create even more in correlation. As explained by P3 for meeting size; *"In larger meetings, I can often feel a bit intimidated, so I often withhold opinions"*, which emphasised that larger groups of people can contribute to more insecurities. Furthermore, P3 stated in relation to interpersonal that; *"The more people or dominate they may be, the more difficult it usually becomes to state one's opinion"*, which further cements the notion of interpersonal and size as a combined factor.

Regarding the interpersonal experiences, P5 stated something similar; *"It comes down to the interpersonal level. If you feel like the opinions you have might be controversial, it might get discarded. The better you feel about people, the more confident you are"*. It gives a clear picture that without good interpersonal relationships, the chances for encountering barriers when conducting group meetings increases.

Lastly, engagement was discovered, as explained by P2; *"I have experienced several times that people either says nothing or speak very little in order to not get noticed"*. Engagement and interpersonal can also create more barriers when the two of them are combined, as quoted by P4; *"My colleague can be really dominating, I knew her before I started working here, so I have had small interactions with her before, and therefore I knew that she sometimes can sound a bit harsh, but she doesn't mean it"*.

4.1.2 *Common Courtesy*. Throughout our analysis we identified the theme of common courtesy, regarding the usage of phone and laptops during meetings, respect, and policies. The theme was categorised with three sub-themes: *interruptions, distractions, and respect*.

For interruptions, codes such as device interruption, and environmental interruptions were used. P5 stated; *"Generally mails and notification can interfere with the meeting. I therefore put my devices on idle for that very reason"*.

Challenges were identified in relation to people using their phones during meetings, and not showing respect towards one's colleagues. This was identified in a statement from P3 *"I really hate when people are on their phones, either formally or informally, because you have no idea what they are doing. It's a lack of respect"*. P2 commented on this issue as well; *"Phones, can be disrupting for me, because I lose concentration if a phone starts to ring"*.

However, this is very different from participant to participant, depending on the context of the meeting and situation. P4 who work at a production company, as a UX designer had this to say on the topic of mobile phones during his meetings: *"During meetings, people are usually very understanding, and don't have their phones on them, however if they get a call from the production, they need to answer it. So there is an acceptance of this, because it is work related, and has a higher priority"*. This indicates that phones can be necessary during meetings, even though they might create disruptive elements.

In addition, during the interviews we asked about the use of policies in meetings. All but one participant stated that they made no use of such policies, but that it was more of an unspoken rule, that they did not make use of their phones unless it was very important. P2 stated; *"At my mentor meetings, it is an unspoken rule, that we do not have our phone on us during the meeting"*, when asked about it. Only P3 said they enforced a rule during meetings *"When software (people) has stuff with them, we have made a rule that nobody has anything in front of them, since they need to pay attention to the presentation at hand"*.

For the environmental interruptions, the most common were either people walking in during meetings, or sound. P3 commented on this and said; *"I hate when people dump in 5 minutes late. We all have stuff to do, so it is frustrating to waste time on such matters"*. When asked about the biggest problem in meeting context P2 said; *"It is properly sound, as from notification from phones and laptops, and people walking in during a meeting, and interrupting the flow"*.

During the interviews, it was made clear that technology could have a disruptive effect on meetings, if not controlled properly. All participants seemed aware of this and described how they took necessary precaution. P3 stated; *"I have a smartwatch and laptop I turn off at meetings, since they easily can distract me"*. But even when saying this, both P1 and P5 also commented that they have made use of their phones during meetings. P1 contradicted himself and even justified it, though he earlier stated that he hated it when others made use of their devices; *"I have a philosophy that if I don't gain anything from the meeting, I can do something else. It could be on the phone or something productive"*, this indicated a lack of respect towards others and their time, and P5 reflected upon past experiences during the Covid-19 lockdown; *"In the past (corona period) I often thought I could multitask online meetings. I later found out that I couldn't"*. This indicates a loss of information, since peoples engagement could be low due to distractions from devices

4.1.3 *Facilitation*. Several encounters regarding the facilitation of meetings and how CA could function as a facilitator during meetings was discovered during our analysis. The theme *Facilitation*, was given two sub-themes, *meeting structure*, and *conversational agent*. When systematically going through the interviews, it was identified how different meetings were structured and facilitated. This topic was briefly presented in the previous section. Furthermore, various ideas of how CA could help and take roles in the context of meetings were presented by the participants.

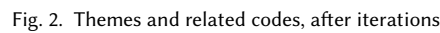
Facilitation of time was discovered as common issue throughout the interviews and P5 commented on their meeting structure; *"We got to have an agenda, otherwise we often stray from the scheduled time, which is bad"*, and further elaborated; *"We sometimes discuss how to manage our time better. It often subsides however"*, indicating a need for time management. P4 described how they at his company usually managed time poorly; *"[...] usually meetings are booked for 30 minutes, but they always take double or even triple that time"*. During the interviews the participants were asked if they could see CA having a role during their meetings, in which P3 responded; *"It could be very helpful, if it could control the agenda in terms of time management. Also, in terms of people who are not saying much, by taking them into consideration"*.

Subsequently, an area of facilitation was identified, namely the inclusion of participants during a meeting. All participants had either experienced not being included or absence, or witnessed others. We identified a reason for this being personalities, meaning that some participants during a meeting can be very dominate. P1 had experienced this before; *"X can be very dominant. If there is some silence, then X can quickly take control. Therefore, I often fall back and take a more subtle role"*. P5 have had similar experiences, and felt that it sometimes could be hard to get back into the conversation, and expressed how CA could have an impact on this interaction; *"It could be 'woke' and listen to the amount of time people spent talking. Then it could help include everyone"*. P4 expressed how even in small meetings it could be hard to include or engage people. But also reflected on the use of CA in such a context; *"I don't know if it would be acceptable if it took the lead. That would required for it to know how was present in the room, and would that be more acceptable than listing to a human?"*

Throughout the analysis, facilitation of meetings were heavily discussed, and how a CA could help improve the meeting proceedings. Besides time management, facilitation of the agenda was also commented on by both P5 and P3, along with recaps of meetings. There was a general excitement revolving the possibility of having a system listing in on the meeting both to facilitate the inclusion of participants, but especially also for recording and transcribing meetings. P4 stated; *"It would be really nice if it could recap, what happened at the meeting"*, and further commented on this; *"I only work Monday and Tuesday, so when i come back the next week I might have forgotten something, so a system of sorts to help me recap on what happened last week, to help ease the information load, would be great"*. In addition, P3 expressed; *"If it could take notes and thereby remove the focus you otherwise had on writing down, then you had more focus at hand"*, further advocating for a feature that would remove the focus from computers and more on the human-human interaction. As a side feature, P3 would also like to have it identify who said what during the meeting.

4.2 Findings From Scenarios

Based on the findings in the initial interviews, we created three scenarios, as mentioned in 3.2, with the purpose of testing the concept and how end-user would react to this technology, in specific situations. After the evaluation, we applied *Reflexive TA*, to the data, and systematically examined it. This generated four themes, *facilitation, challenges of meetings, interpersonal* and CA, with four sub-themes, see fig. 2.



One of the first comments from P6 explained that; “We don’t have meetings with people where people’s opinions does not matter. In larger, formal meetings, maybe. In our meetings, we need to hear everyone out” but later showed frustration on how some meetings were facilitated by stating that; “I often feel we get divided into groups, two-on-two, during meetings. Especially during brainstorming, where I feel like everybody should be sparring”. This contradicts or at least conflicts with P6 perceptions on how the meetings are conducted, based on the two statements. This is further signified from a statement P6 made; “You might have an agenda, but you can easily get off course. Then you might have difficulties getting back on track”, hence suggesting that the structure of meetings are somewhat fleeting at times. Different perceptions and experiences were discovered during the analysis of the scenarios. Here P7 and P8 contradict each other regarding the usage of phones. P7 stated; “You can say a lot about the meeting culture here, but I have never encountered a problem with people being on their phones or being distracted. I don’t know regarding Teams meetings, but not physically”, which directly contradicts with what P8 expressed during our evaluation; “I wouldn’t say it is a problem, but everyone sits with their phones during meetings. Then you can do something parallel, listening with one ear. Especially under corona and online”.

Throughout the evaluation, we identified mismatch in how the perception of facilitation was handled, and how it could both go unnoticed and be frustrating. This could indicate an insufficient facilitation, in terms of a misconception regarding the meeting.

Agenda was a topic that all our participants shared an opinion on, and based on scenario three, we received a lot of positive comments on this topic. P6 reflected upon how the CA could be a part of facilitating specific parts of the meeting, and stated; *"In terms of brainstorming ideas, I believe that one should present their idea, and then we can talk about that. But the way we do it right now, we kinda spilt up in groups. So a way for it to facilitate the agenda of brainstorming would be really nice"*. P9 also saw the potential of helping with facilitating breaks; *"Could be nice with reminders for breaks or agenda management. Sometimes we often forget breaks or time runs out, when there are still important aspects to discuss"*, and P7 came with another idea for how it could be helpful in facilitating the agenda by stating; *"Provide the CA with the agenda, it transcribes into the different topics, based on user input from the meeting facilitator"*. Management was identified as the primary factor, when addressing facilitation, and something that the participants were very aware of.

4.2.2 Efficiency. Efficiency was identified as a theme for this, and involves areas, such as *time management, augmentation, loss of information and problems of being a facilitator*. During the course of our Reflexive TA, we started to notice how efficiency was a reoccurring topic for our participants, and how they wanted to be more efficient at meetings. They quickly started to see how CA could fill this role. Especially, time was heavily discussed, and P9 stated; *"It could help facilitate time so we manage to get all out of the meeting, and so that we remember to wrap up"*. P8 further explained the problem, of time management; *"We have a huge problem with wrapping up meetings, so we can get to the next meeting"*. P8 further elaborates on why time is important, and how the CA could provide a solution for this; *"Time often runs out for us. A notification from the CA could be beneficial to inform of the time. This place (the company) is huge, so you have to leave in advance, otherwise you might miss another meeting because you have to walk. And that is super interrupting when people arrive late!"*. P7 expressed ideas on how the CA should interact with us for time management and said; *"Light or a bar visible to the facilitator indicating the % of the agenda missing and time left"*, underlying the improvement of the facilitator.

In scenario three, we presented the participants with the possibility of recapping meetings, or creating resumes of a meeting. This idea was not to our surprise, based on the initial interviews and own experiences, highly sort after. All participants saw the idea of the agent taking notes, as an extremely helpful tool, P6 explained why; *"When being a facilitator, typically we take notes as the meeting happens, and there will be a loss of information here"*. Through our evaluation, it was discovered that the facilitators was under a lot of pressure, hence they had to present the meeting, facilitate it, make sure to include different participants to ensure an adequate meeting and take notes for the recap. The participants did not express discomfort towards this process of having a meeting directly, but was aware that by allocating resources to taking notes, they were in danger of losing engagement and information. In addition, P6 further stated that the idea of receiving important data and information in terms of related assignments, would be very beneficial, but also expressed scepticism in that regard; *"If it (CA) has to tell something technical in terms of our products, it needs to be updated and specific. Things change all the time here, so it is necessary to be up-to-date with that"*.

4.2.3 Challenges of Meetings. We sought to acquire knowledge on the type of challenges, the participants came across in their meetings. This will help us identify new challenges, and compare with the challenges from our initial interviews. As discovered in our initial interviews, the usage of phones and laptops seems to be a reoccurring topic in these evaluations. P9 expressed her feelings about the issue; *"It is very distracting when we look around and see people with their laptops, while someone is presenting"*. Both P6 and P8 had used their phones in previous meetings, confirming the

distractions of technology and the challenges of keeping people engaged in meetings. To this P6 commented when presented with scenario two, see 3.2.2; *"I think it can be very useful to set the "tone" of the meeting, so all know how it is done"*, this was also addressed by P9; *"I think it would be really nice, that if we just in general, were presented with the guidelines at the start of a meeting"*. This indicates, that even though we identified that policies was not used in general, but more as unspoken rules, a need for clearer policies was in demand.

Three other topics of challenges in meetings were discovered through our analysis; *being late, disruptive talking and miscommunication*. All participants had experienced colleagues being late to meetings, and all but one also admitted to being late at least once, and P7 stated; *"Sometimes people are late, and that is annoying"*. Even though they try to be on time, it will occasionally happen that one is being late. The reason it seemed frustrating to the participants, was the amplifications it would have on the meetings, as explained by P8; *"People coming late to meetings can be really annoying because we have to recap"*. P8 further explained that, a reason for these challenges of being on time, came from the management side of things, and that they sometimes had a meeting from 9-10, and at 10 they would have another meeting. However, the 10 o'clock meeting was located on the other side of the factory (a 15 min walk). So a better facilitation of meeting time was also discussed as a potential role of the agent.

The second, disruptive talk, was more of mixed challenge, since not all participant had experienced this. When asked about the biggest interruptions during meetings, P9 answered; *"The most annoying thing is, if two people in the corner start to have their own conversation"*, and P7 had a corresponding answer; *"chit-chat during a meeting, prolonging the meeting is an issue for me"*.

Lastly, we identified miscommunication as a challenge, since several participants expressed positive attitudes towards the first scenario, see 3.2.1, in certain contexts of meetings. P9 could see the potential of the agent, and the help with terminology at larger meetings where all do not possess the same knowledge. But in smaller meetings, she had difficulties seeing its usefulness; *"In terms of meetings with similar occupations, I don't see a need for CA to describe terminologies. In more diverse meetings I definitely can"*. P6 when asked about this, answered similarly; *"The use of terminology understanding, is dependable of who is attending the meeting. If it is a smaller meeting, usually all that are involved typically know what is discussed and are there for a reason. On bigger meetings, like department meetings, there might be a use for it"*. P7 was very intrigued by the idea and stated; *"I can really relate to the first scenario. In my daily work, I have contact with so many people, and they all use different terms"*, she further explained, that due to this miscommunication, she often has to redo a task, which she feels is waste of time. However, P8 found it hard to see the usefulness of the agent in terms of terminology; *"Typically the people at the meetings know about the information, and I do not think it is a problem"*.

4.2.4 Interpersonal. Throughout our systematic analysis of the evaluation, we identified a theme we named interpersonal. As the word implies, this theme encircle the relationship and communication between people. Codes such as *engagement, dominance and courtesy* were used. After the presentation of the first scenario P6 commented; *"There are always someone who is very dominant in group meetings, and it affects how the meetings are being conducted"* and added to how it affected the meetings; *"[...]people sometimes, fall back and don't join the conversation, due to dominate participants during a meeting"*. This further cements, that personality has a role in the challenges of group meetings. P8 also reflected on this, and told that she could recognise this in her own work and meetings.

Engagement and lack of it, was something we came across in our discoveries as well. During meetings P9 had experienced such insufficiency; *"In my previous job, we often had very large meetings for design reviews. Everyone with*

a leader function was there all day. When something was not relevant for them, even though people were presenting for feedback, they would be on their computer and phones. That was very annoying and disruptive”.

We discovered, that people tend to work on their own tasks, if they are not involved in the meetings, and even our participants had done the same. P8 commented on this and said; *“I often get something done while (sitting on the phone) at the meeting, since it is not always totally relevant, even though I know I might miss some information from the meeting”*. It was surprising, that even though she knew, that she would lose information, she still felt it was justified. When asked about the ramification of this, she said that she did not feel it was problematic, since everyone is doing it at some point, during a meeting. This last point of view was very different from what P7 experienced. When presented with the scenario, she could not recognise the dilemma, and felt that all were engaged in the meetings. P6 also reflected on this and stated; *“the issue of engagement, and that when the meeting is not revolving around you, you pick up your phone to keep yourself activated, instead of being present at the meeting”*.

During the evaluation, it was identified that interpersonal played a role in how we perceive meetings. Different personalities play a role in the facilitation of the meeting, and again some sort of courtesy is missing, in terms of being engaged in other people.

4.2.5 Conversational Agent. The agent played a main role in the scenarios and derived a theme of its own. We were able to identify needs, potentials, expectations, impressions, and reflections of use, function, accessibility, and capabilities. *Conversational Agent*, where given a sub-theme named *Human perception*.

To touch on the previous topic of engagement and the usage of phone during meetings, the participants were presented, in scenario two, with a solution to how an agent could help facilitate such a situation, see 3.2.2. P6 really liked the idea around this, and felt it could be beneficial, and the fact, that it was delivered it in a funny way, made it easier to relate to. He further commented; *“I could see it as useful to engage people more”*. In addition to needs and potential, P7 presented us with another idea; *“The agent could notify the facilitator about certain people, that have not said anything. Not including the other participants in the meeting, and therefore enhancing the facilitator”*. We asked her about the problem of confronting people in the open, for not engaging, to which she answered; *“I feel it could be very transgressing to be called out like that, I do not think it would have the desired effect”*. P9 expressed a similar opinion on the topic; *“It can be very interrupting if a system interrupts the ongoing conversation”*. This again depends on the personality of the participants at a meeting, as P6 elaborates on his view of the topic; *“It can be difficult for introvert people to join in, even if they are forced ”*.

After the evaluation the idea of only notifying the meeting facilitator, was taken into consideration, since the purpose of the agent was to help the interaction between humans and supporting this feature. P9 also talks about how the concept of CA could be able to enhance the human, and talks about some of the same features as P7; *“It can help the facilitator with providing information, and then let the facilitator make a decision based on the data provided”*.

It was identified that the participants from the evaluation are obligated to take notes as the facilitator of the meeting, these notes are then formed in to what they call *minutes of meeting*, or a resume of the meeting. Our participants all shared an excitement for the agent being able to recap meetings and therefor lessen their workload, in the matter of taking notes. P8 commented; *“Analysis, recap, and sending out minutes of meetings would make our lives so much easier”*. The idea of the transcribing feature is to divert the focus from laptops and phone and into the conversation. P7 stated that it would help her with the information load, for her to focus on asking the right questions, and ensure she did not forget anything, because of the duality when both facilitating and taking notes. This was also confirmed by P9 who

stated; *“It is hard to handle both recap of meeting and agenda”*. A feature of the agent could be to help minimise the information load, and provide a solution to ensure engagement in the conversation, without losing information.

Lastly, P8 recommended how the CA could establish contact or inform of time; *“It could tell, by sound or light, 5 mins left of the meeting remember to finish up”*, P7 also recommended that; *“If it could visualise and show how much time there was left, then it could be very helpful”*.

4.2.6 Human Perception. The topic of human perception, was included as a sub-theme under *Conversation Agent*, to highlight the way in which something is regarded, understood, or interpreted of our participants. We discovered, both positive and negative perceptions and opinions towards the concept of having AI help facilitate the social interaction of group meetings. We encountered topics such as; *augmentation, automation, involvement, fear of tech, training of human abilities, losing our human side, and mistrust*.

P6 had a positive view on the agent and its involvement; *“It has the possibility to control the meeting facilitation, however, it will not delude the content of the meeting”*. But is also aware that his view might not be shared and further comments on this; *“I think it will divide the waters so to speak. You have to be adaptable for this kind of technology”*. This is mutual for P8 who had this to say about the topic; *“We are typically open about any tech, but I think it would take some getting used to”*, P8 also expressed; *“It is exciting, but it is something we would have to adjust to. I see the potential it could have”*. On the other hand P7 was more sceptical, and showed more concern towards AI having a large of an impact and stated; *“I fear we get too technologically dependent as humans. We need to learn about empathy for other human beings. Otherwise, we will lose essential human traits to technology”*, in addition she stated; *“There is always a pill for anything, and there is always a technology to fix anything. Instead of focusing on ourselves as humans and trying to improve”*. This is a common phenomenon, mistrust in IT or fear of it. This is an area of which we have to be aware and take into consideration when designing such technologies. P6 also stressed the importance of the experience with the technology; *“When we were in lockdown (corona), we had to use Teams (Microsoft Teams) instead of physically meeting, everyone hated it. Now we don’t want to part with it. The easier it is to use, the faster you adopt it”*.

However, P7 was not completely against the idea and stated that in the context of CAs trying to improve the meeting facilitator, she could see potential. We asked her to elaborate on this; *“It could function as a tool to learn how to include people in meetings, and facilitate the good conversation/discussion”*.

Generally, the participants was intrigued by the concept of the agent, and could see it used in a way to improve how to facilitate meetings, presentations, and management. It could help move the focus from the screens to the human-human interaction, and lessen the workload on menial tasks such as taking notes for resumes.

Consequently, our evaluation identified 14 barriers and challenges regarding facilitation of meetings and adaptation of agents, showcased in table 2. Overall, we identified a positive perception of the concept, with a few things to consider when moving forward with recommendations to the HCI community.

Challenges of meetings	Challenges of adoption
Phone and laptop usage	Mistrust to conversation agents
People being late, and or interrupting the meeting	Advancement of the technology
Time management	Acceptance
Miscommunication	Context and settings
Engagement	The presence of the conversational agent
Interpersonal	Perception
Loss of information	
Unspoken rules	

Table 2. Challenges and barriers identified

4.3 Design Expo

As part of our master thesis, Aalborg University arranged the Design Expo, which is an exhibition for students to showcase their thesis projects to companies that may be interested in our field and expertise. Various companies were presented with the data from our pre-thesis, research, methods, findings, and limitations from our master thesis project. Concretely, we had a dialogue with eight out of fourteen companies regarding our project. All the company representatives that attended our stand could share similar experience or thoughts when presented with our research questions and findings from our evaluations. Specifically, the majority of them gave valuable inputs from their own meeting experience, procedure, and format. They explained their own context, and the potential of the CA when it comes to facilitating better human-human interaction during meetings. In addition, they also saw great leverage and motivation for automating note-taking and transcription, since it would further support their human-human interaction. A representative from an IT department that works with logistical systems explained that; *"I have never thought of it that way. We spent so much time facilitating meetings on how we can optimise our systems, but we encounter many of the same issues that you have in your findings"*. Another representative from a software team, working at a bank, exhibited fascination for the efficiency a CA could have regarding meetings; *"It is quite funny. We have meetings to ultimately make things better, because meetings are essential for every businesses. However, we seldom look at how we can facilitate better meetings, which I see now we might benefit more from"*. Lastly, a remark for the usage of CA technology was from two senior software developer working at an IT consultant firm, where one stated that; *"This type of technology has great potential for the solutions you try to accommodate. Though it is gonna take a very long time to reach it, as certain aspect is not up-to-date with the expectations, yet."*

Generally, we received positive feedback from the representatives. The majority of our findings correlated with challenges they had experienced when facilitating group meetings, and expressed an interest in the possibilities of what a CA could provide.

5 DISCUSSION

We aim to identify the challenges and barriers there might be when adopting AI support systems for social interactions in the labour market. Furthermore, we wanted to identify the needs and expectations of the end-users, in order to explore the possibilities of implementing CA in the fore mentioned context.

Through our systematic approach of analysing both interviews and evaluations, we identified eight challenges regarding the facilitation of meetings, and six challenges when trying to adopt CAs in social interaction, as listed in table 2.

We identified a positive perception of conversational agent in group meeting facilitation, with a high focus on helping with management and easing the participants menial tasks, to prevent loss of information and increase substance of meetings. In addition, we discovered an unexpected high usage of devices during meetings, which had a negative impact. Next, we present our seven recommendation to the HCI community, based on the findings from both the initial interviews and scenario-based evaluations. Lastly, we discuss the implications of collecting qualitative data during this study.

5.1 Recommendations

The recommendations is based on the analysis of our findings, in correlation with the investigated research conducted throughout this study. They serve as initiatives to accommodate the challenges that were discovered and presented. Specifically, we address the recommendations to designers and practitioners that wishes to use CAs to overcome barriers and challenges that exist in group meetings.

5.1.1 Establish inclusion of all relevant people present during meetings. Throughout our analysis we were presented with the challenge of engaging participants. We identified two main reasons for lack of engagement, *interpersonal* and *device distractions*. The area of interpersonal addresses participants and the abilities of the facilitator. A common issue was regarding the difficulties of including participants at meetings, or ensuring their opinions being heard. Both introvert and dominant personalities have to be accounted for in meetings. For this we recommend; *a solution, in which the conversational agent augment the facilitator, providing data of participation of the attendees*. This approach was discovered during the evaluation, and was positively received, by the participants. This is supported by Cutler et al. (2021) that argues that member inclusion is a factor for better collaboration [12].

5.1.2 Reduce the usage of devices during meetings. According to Lunenburg (2010), phones is a physical barrier that hinders efficient communication [23]. Therefore we hypothesis, in correlation with our discoveries, that phones and laptop usage during meetings create disruptions for communicating. We tried to accommodate this problem with a solution, that briefly when detecting a phone being used, commented on this. By designing the feature as a fun, recognisable speech, we tried to incorporate comedy for something mundane, and hereby increasing the messaged received. Based on our discoveries we advocate: *Focus on reducing device disruption through the conversational agent to enhance the physical interaction and awareness during meetings*.

5.1.3 Ensure a feeling of cooperation with the agent. Human-AI collaboration is vital for the adoption of the agent. To establish this we recommend an approach were we support the feeling of cooperation. In a paper on incorporating AI in group facilitation, Suh et al. (2021) found that with the presence of an AI, most participants felt it supported them progressing faster [27]. This indicates a higher rate of acceptance if the participants can see and feel the effects of AI in their work tasks.

During the analysis we discovered mixed feelings towards different aspects of the system. Features such a transcription were highly sought after, along with time management. However, a feature such as intrusion for inclusion, even though identified earlier as important, was less sought after. We hypothesise, that the less liked features could be due to the association of cooperation with the agent. This is also documented by Laban and Araujo (2020), in which that perceiving

an agents as cooperative is associated with perceiving the agent as more anthropomorphic, and hereby perceiving the performance as being better [17]. This is supported by the claims of Ling et al. (2021) which stated that user acceptance is linked to usage benefits [21].

5.1.4 Reduce menial tasks to enable end-users to focus on their main goal. When designing CA for social interaction, it might be beneficial to look into which menial tasks it can help facilitate. We want to take into account how we can design better facilitation to overcome barriers and challenges for the users. For this we looked into the benefits of removing note taking, and automating recaps of meetings, hereby facilitating a more productive interaction between the participants. We wanted to direct the focus away from computers and over to the conversation. The role of a “secretary”, taking notes from the meetings was met with very positive feedback. The participants believed that this would enhance the meeting experience, since they no longer had to focus on taking notes, therefore reducing the loss of information and potentially enhancing the engagement in meetings.

Consequently, by letting the meeting participants focus more on the main goal, by reducing menial tasks, we hypothesise an increase in interaction, which could lead to a more productive meeting. Based on this we recommend; *identifying and reducing these menial tasks, with the purpose of generating a better meeting environment, with focus on the social interaction.*

5.1.5 Create individual adaptability to ensure better adoption. In our related work, we highlighted a systematic literature review by Bavaresco et al. (2020), which suggested future directions for CAs in business related fields. One of the directions were *personalisation*, a feature in which the CA adapts dialogue or tasks to the specific user [4]. From our findings, we identified that the CA must be able to adapt to various users and meeting contexts. Larger and formal meetings have another set of formalities and requirements, in relation to a small, informal meeting, where the participants have another set of structure and interpersonal relationship. Participants that are attending the meeting may have important roles, such as facilitators, designers, or users regarding technical, practical, or economical aspects, while some may be simple observers. The CA has to be adaptable for each of these factors, therefore we advocate for *individual adaptability as an important ML recommendation to ensure better adoption in group meetings.*

5.1.6 Incorporate self-learning to enhance better interaction with conversational agents. Another direction that is tied to *personalisation* is *self-learning*. *Self-learning* relates to the use of ML to facilitate better interaction with the CA [4]. This was also discovered during our evaluations, as internal data and structures were ever-changing. For example, requesting up-to-date information when addressed during meetings, in which the CA should participate. It must also be able to learn from the context of conversations and each individuals role, engagement, and persona. In addition, it must also learn the difference between idle chat and work related conversations to improve the efficiency of meetings. Though, it still has to take into account that friendly chat and jokes can also be important for building rapport. This is supported by Ling et al. (2021) as an important design feature, as companionship and relationship building contribute to better acceptance and adoption [21]. We therefore advocate and highlight; *the importance of incorporating the ML techniques of self-learning to enhance the interaction of the CA.*

5.1.7 Provide intervention without being intrusive. Intervention is a recommendation we identified during the analysis of our findings. This was identified as an important feature, as the participants expressed great fondness but also several concerns for intervention. If a discussion unfolds, it is important not to interrupt the conversation. In relation, the conversation must not be unrelated or surpass the specific agenda, hence resulting in loss of information that could require rescheduling another meeting. It correlates with the important aspects that *self-learning* and *individual*

adaptability provides. Furthermore, it is important the CA does not expose or enforce someone into an uncomfortable position however, the inclusion of people must not be unattended due to engagement and dominance. Consequently, it must learn some level of commitment for when to intervene, since there is a fine line between efficiency and disruption.

To design non-intrusive functionalities we recommend; *usage of ambient notifications* since it notifies in a more subtle manner, and *direct notifications to the group facilitator* to augment their roles and enhance human-human interaction.

5.2 Qualitative Data Collection

Research show that expectancy effects can alter the interviewee's behaviour, which might have taken place as the scenarios demonstrated how the CA could accommodate some of the challenges and barriers that occurred during meetings [1]. Conflict between statement and opinions is clearly showcased in our findings, see 4.2. Here P6 tells that everyone's opinions is taken into consideration, but yet feels frustrated being divided into small groups. Even though they all work together and often facilitate the same meetings, they each have very different perspective on how things unfold. This is also seen by P7 who never experienced people on their phones or not fully participating, whereas both P6 and P8 states they often write e-mails, are on their phone or something they felt was productive for them. This may be related to biases or fading retrospective memories. Observation and diary studies could provide more objective data in that regard.

Different aspects of tasks and friendship are a factor that applies when facilitating meetings at the university, where professionalism and rules applies to the context of the business industry. In addition, it further showcases the differential factors between university and business, when interviewing the second pool of participants. Here sub-themes such as *meeting context* and *interpersonal* played a major role in the industry context, since there were different meeting sizes, contexts, profiles, types, and structures. This generated a lot of code for the theme *interpersonal*, since engagement, courtesy, and personality varied a lot based on the meeting context.

However, the themes and scenarios that were created from our initial interviews reconciled with the scenario-based evaluation. Three out of four participants from the scenario-based evaluation shared very similar experiences and opinions, highlighting familiarity between the two groups of participants. Furthermore, the findings from our analysis were also identified by the eight companies that participated during the Design Expo, signifying a common thread regarding barriers and challenges that occurs during group meetings in most settings.

6 LIMITATION AND FUTURE WORK

For our first interviews, we chose to interview fellow interaction design students. This was based on the assumption of seeing them as experts, since they have spend almost five consecutive years working in groups, which is a mandatory part of the Aalborg University model. In addition, they all have had jobs beside the university, where they also have experienced group meetings in some form or another. However, one of the complications that may occur from this, is that group meetings at the university and in the labour market can differ a lot, see 4.1.

Given more research and time, it could have been optimal to conduct some observations in meetings from different companies, to see how each respective meeting was conducted. Furthermore, by conducting a diary study over a longer period of time, where we equip the participants with journals to write experiences down as they take place, we could achieve more precise remembrance when interviewing and analysing it afterwards. These findings could have provided with different perspectives and illustrations on other aspects to meetings, that we might not have identified during our study. Therefore, we would like to conduct observations and diary studies in the future, that potentially could lead to new discoveries and recommendations as well.

For analysing the gathered data, we made use of Reflexive TA by Braun et al. (2019) to generate code and themes to systematically review the data. Here Braun et al. (2019) state that in terms of conducting interviews, the minimum amount should be at least five to six interviews for a small project, if the data is rich, the answers homogeneous and research questions are well-defined [7]. This was the case for our initial interviews, as they generated very similar and rich data in relation to our research questions. However, the scenario-based evaluation were only conducted with four participants, which is below the minimum suggestion by Braun et al. (2019). The answers were relatively homogeneous, with some deviations, and as mentioned earlier had similarities to the initial interviews as well. This is seen as a limitation, due to the methodology when analysing and generating themes. In the future, we would have to acquire a larger participant sample, to ensure more validity in terms of what the conducted methodology proposes.

As highlighted in our related work, the research emphasises the importance of social interaction, provided by the CA, to ensure better acceptance and adoption, see 2. This was under-explored during our scenario-based evaluation, since the participants had no real interaction with the CA. The CA we used were limited to few actions, hence why the scenario-based evaluation was a more approachable method to conduct an evaluation. We argue that this may be the largest limitation since no real testing of a CA has taken place.

For future work we would like to implement better social skills, as emphasised in our related work, into the CA. In addition, we would like to test and evaluate these social skills, roles, and relationships the CA should uphold, to clarify what factors can contribute to better acceptance, adoption, and social interaction in group meetings.

7 CONCLUSION

In our pre-thesis, we discovered a gap in the research through a methodological review of 80 relevant studies. Based on this, we set out to investigate the issues with adopting AI in social context. We wanted to highlight the needs and expectations of the users in the labour market for AI support systems. We provide a systematic Reflexive Thematic Analysis, identifying fourteen challenge and barriers when trying to adopt AI support systems for social interaction and facilitating meetings.

Our study is based upon previous work within the HCI community, interviews with fellow master students and scenario-based evaluation with employees from the labour market. Through our analysis, it was discovered that the reduction of menial tasks, with the help of a conversational agent, would have an effect on productivity. In addition, the inclusion of participants during a meeting was found to be a challenge. This both concerned the interpersonal aspects of the participants as well as the facilitator. With the help of a conversational agent it could be possible to augment the facilitator to enhance the meeting quality and experience. Our findings indicated, that the usage of devices during meetings was found to be more disrupting than beneficial, however in some cases it was a necessity. Lastly, it was identified that the adaptability of the conversation agent was crucial for the adoption in social interaction, along its ability to self-learn.

While we are seeing an increase in the use and development of conversational agents, research in the area of AI and social interaction on the labour market is limited. By providing seven recommendation to practitioners and researchers of the HCI community, and potentially other related fields, we aim to highlight some of the challenges when adopting AI support systems for social interaction, in the labour market. We hope our study can contribute to future research, and accommodate the research gap that is highlighted for developing AI support systems for social interactions, across the labour market.

REFERENCES

- [1] Nicole M Adams-Quackenbush, Robert Horselenberg, Josephine Hubert, Aldert Vrij, and Peter van Koppen. 2019. Interview expectancies: awareness of potential biases influences behaviour in interviewees. *Psychiatry, Psychology and Law* 26, 1 (2019), 150–166.
- [2] David Arnott and Graham Pervan. 2015. A critical analysis of decision support systems research. In *Formulating research methods for information systems*. Springer, 127–168.
- [3] Suket Arora, Kamaljeet Batra, and Sarabjit Singh. 2013. Dialogue system: A brief review. *arXiv preprint arXiv:1306.4134* (2013).
- [4] Rodrigo Bavaresco, Diógenes Silveira, Eduardo Reis, Jorge Barbosa, Rodrigo Righi, Cristiano Costa, Rodolfo Antunes, Marcio Gomes, Clauter Gatti, Mariangela Vanzin, et al. 2020. Conversational agents in business: A systematic literature review and future research directions. *Computer Science Review* 36 (2020), 100239.
- [5] Robert P Bostrom, Robert Anson, and Vikki K Clawson. 1993. Group facilitation and group support systems. *Group support systems: New perspectives* 8 (1993), 146–168.
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [7] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.
- [8] Auteur Emeline Brulé. 2020. Thematic analysis in HCI. <https://sociodesign.hypotheses.org/555>
- [9] Jessy Ceha, Ken Jen Lee, Elizabeth Nilsen, Joslin Goh, and Edith Law. 2021. Can a Humorous Conversational Agent Enhance Learning Experience and Outcomes?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [10] Minji Cho, Sang-su Lee, and Kun-Pyo Lee. 2019. Once a Kind Friend is Now a Thing: Understanding How Conversational Agents at Home Are Forgotten. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) (DIS '19). Association for Computing Machinery, New York, NY, USA, 1557–1569. <https://doi.org/10.1145/3322276.3322332>
- [11] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, et al. 2019. What makes a good conversation? Challenges in designing truly conversational agents. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [12] Ross Cutler, Yasaman Hosseinkashi, Jamie Pool, Senja Filipi, Robert Aichner, Yuan Tu, and Johannes Gehrke. 2021. Meeting Effectiveness and Inclusiveness in Remote Collaboration. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 173 (apr 2021), 29 pages. <https://doi.org/10.1145/3449247>
- [13] Enrique Fernández-Macías and Martina Bisello. 2021. A comprehensive taxonomy of tasks for assessing the impact of new technologies on work. *Social Indicators Research* 159, 2 (2021), 821–841. <https://doi.org/10.1007/s11205-021-02768-7>
- [14] Jerry Fjermestad and Starr Roxanne Hiltz. 1998. An assessment of group support systems experimental research: Methodology and results. *Journal of Management Information Systems* 15, 3 (1998), 7–149.
- [15] Thomas Ildstrøm and Alexander Rath Hakonsen. 2021. Artificial Intelligence in Work: What do HCI Researchers Study?. In *Interaction Design Pre-thesis*. 1–16.
- [16] Starr Roxanne Hiltz Jerry Fjermestad. 2000. Group support systems: A descriptive evaluation of case and field studies. *Journal of Management Information Systems* 17, 3 (2000), 115–159.
- [17] Guy Laban and Theo Araujo. 2019. Working together with conversational agents: the relationship of perceived cooperation with service performance evaluations. In *International Workshop on Chatbot Research and Design*. Springer, 215–228.
- [18] Raina Langevin, Ross J Lordon, Thi Avrahami, Benjamin R Cowan, Tad Hirsch, and Gary Hsieh. 2021. Heuristic evaluation of conversational agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [19] Federica Laricchia. 2022. Digital assistants usage by vendor worldwide 2019. <https://www.statista.com/statistics/1134020/digital-assistants-usage-worldwide/>
- [20] Federica Laricchia. 2022. Number of voice assistants in use worldwide 2019-2024. <https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/>
- [21] Erin Chao Ling, Iis Tussyadiah, Aarni Tuomi, Jason Stienmetz, and Athina Ioannou. 2021. Factors influencing users' adoption and use of conversational agents: A systematic review. *Psychology & Marketing* 38, 7 (2021), 1031–1051.
- [22] Nicholas Looker, David Webster, Duncan Russell, and Jiudong xu. 2008. Scenario Based Evaluation. *Proceedings - 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing, ISORC 2008*, 148–154. <https://doi.org/10.1109/ISORC.2008.56>
- [23] Fred C Lunenburg. 2010. Communication: The process, barriers, and improving effectiveness. *Schooling* 1, 1 (2010), 1–10.
- [24] Elayne Ruane, Abeba Bihane, and Anthony Ventresque. 2019. Conversational AI: Social and Ethical Considerations.. In *AICS*. 104–115.
- [25] Ameneh Shamekhi, Q. Vera Liao, Dakuo Wang, Rachel K. E. Bellamy, and Thomas Erickson. 2018. Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173965>
- [26] Maxim Stepanov. 2020. Some stats about voice assistants. <https://uxdesign.cc/some-stats-about-voice-assistants-1c292476584>
- [27] Minhyang (Mia) Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 582, 11 pages. <https://doi.org/10.1145/3411764.3445219>

- [28] Franziska Tschan, Norbert Semmer, and Laurent Inversin. 2004. Work Related and “Private” Social Interactions at Work. *Social Indicators Research* 67 (06 2004), 145–182. <https://doi.org/10.1023/B:SOCI.00000007338.60393.bf>
- [29] Murray Turoff, Starr Roxanne Hiltz, Ahmed NF Bahgat, and Ajaz R Rana. 1993. Distributed group support systems. *MIS quarterly* (1993), 399–417.
- [30] Jinping Wang, Hyun Yang, Ruosi Shao, Saeed Abdullah, and S Shyam Sundar. 2020. Alexa as coach: Leveraging smart speakers to build social agents that reduce public speaking anxiety. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [31] Ran Zhao, Oscar J Romero, and Alex Rudnicky. 2018. SOGO: a social intelligent negotiation dialogue system. In *Proceedings of the 18th International Conference on intelligent virtual agents*. 239–246.
- [32] Ilze Zigurs and Bonnie K Buckland. 1998. A theory of task/technology fit and group support systems effectiveness. *MIS quarterly* (1998), 313–334.