Designing Automated Facilitation for Design Thinking:
A Chatbot for Supporting Teams in the Empathy Map Method

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

Omid Shoury
University of Hamburg
omidshoury@hotmail.de

Abstract

The Empathy Map Method (EMM) in the Design Thinking approach is a powerful tool for user centered design but relies on the methodological skills and experience of rare facilitation experts to guide the team. In a collaboration engineering effort, we aim to make this expertise available to teams without constant access to a professional facilitator by packaging facilitation knowledge into structured process support and state-of-the art technology. Based on requirements from scientific and practitioners' literature, we introduce the concept of a conversational agent in the form of a chatbot to take over the role of the facilitator of the EMM. We present an initial wizard of oz evaluation to derive insights and implications for improvements and the software implementation towards the ambitious goal of automated, non-human facilitation of EMM.

1. Introduction

Human-Centered Design enables a deeper understanding of people's needs and empathizes with their problems, for which a service or product is designed [30]. Problem identification and solution occur in close collaboration with customers [12]. Hence it is crucial to support the targeted user groups with innovative solutions by identifying and solving real problems [30]. Due to the shift towards service-based business models, companies need to improve their innovation processes and customer focus [55]. Design Thinking (DT) is a popular and powerful approach for this challenge [16]. Curedale defines DT as "[...] a people centered way of solving difficult problems. It follows a collaborative, team based cross disciplinary process. It uses a toolkit of methods and can be applied by anyone from the most seasoned corporate designers and executives to school children" [16]. One method in DT is the Empathy Map Method (EMM), a complex creativity technique to develop empathy for potential customers and gain new insights into their needs [41]. A

facilitator guides the collaboration process, who needs amongst others domain specific methodological knowledge [14] in DT and EMM. Identifying and hiring a facilitator with these specific skills is costly. In addition, EMM makes high demands on the facilitator's social and cognitive abilities. A promising approach for making complex collaboration techniques widely available to non-method-experienced practitioners is to document facilitation knowledge by means of Collaboration Engineering (CE) approaches in structured process designs [35, 51]. Method knowledge can then be implemented in pre-configured IT systems in order to perform the process in a semi-automated way [10]. As part of this research, we aim to make EMM success independent of expert human facilitation by digitalizing EMM in the form of a facilitation chatbot (electronical EMM; eEMM). The paper is structured as follows: Section 2 presents related work on chatbots and the EMM and requirements are derived from this literature base to ground our design. Section 3 outlines design and evaluation methodologies. Section 4 presents the chatbot concept. The concept has been validated in an initial wizard of oz instantiation, which is described in section 5. In section 6, we discuss implications for the chatbot design and contributions to theory and practice before summing up the work and giving an outlook on further research.

2. Related Work

Table 1. Related research topics

	Research Topics				
Article	Agent	DT	EMM	CE	Facilitation
(Harding & Swarnkar 2013)	+	-	-	-	+
(Strohmann et al. 2017)	+	+	-	-	+
(Graesser et al. 2001)	+	-	-	-	+
(Dyke et al. 2013)	+	-	-	+	+
(Kumar & Rosé 2011)	+	-	-	+	+

HICSS

Table 1 shows a selection of related work from a preceeding literature analysis, which include two or more of the following topics: agent, DT, EMM, CE and/or the facilitation concept. To the best of our knowledge, no scienctific literature adressed the digitalization of the EMM yet. The next sub sections go into detail on related work and how it impacts our design choices.

2.1 Chatbots

Various terms can be found for chatbots like chatterbot, conversational agents or machine conversation system. Lieberman describes an agent as a computer program that can be considered as an assistant for users [37]. Schlicht defines chatbots as "service[s], powered by rules and sometimes artificial intelligence, that you interact with via a chat interface" [47]. He differentiates chatbots by their functionality, i.e. whether they act rule-based or based on AI. In this context, proactive (activity is initiated by chatbots acting in advance of future situations) and reactive actions can be distinguished [39]. A distinction between a macro and micro level architecture to instantiate proactive and reactive reactions is often used in chatbot design, e.g. [18, 34], which we adopt for our purpose. Chatbots simulate natural language [8]. The architecture of a chatbot integrates a language model with calculation algorithms. A chatbot's power to act autonomously as a conversational agent for tasks that are usually performed by humans is determined by its AI. As part of this work a largely rule-based concept of a chatbot will be presented as a controllable, baseline approach for the structured method EMM, which could potentially be extended in the future, if certain AI functionalities are beneficial and reliable. Like this, we are able to simulate the core process support functionalities without potential distractions from unexpected AI behavior. Toward the same end, we chose a simple text-based bot over other forms of speech based or embodied agents that might have different effects on the interaction. An advantage of this choice is that participants and researchers can easily access the complete conversation documentation. Recommendations on how to design chatbots for different aplications are predominantly found in practitioners' literature to date, as organizations have started to explore their potentials but research is still in an early stage. Table 2 gives an overview of recommendations from practitioners' literature we used to derive requirements for the eEMM.

Table 2. Chatbot recommendations

Recommendation	Sources
Give adequate introduction for how to start and for individual tasks	[5, 15, 22]
Provide a clear and easy to understand menu	[5]

Communicate, why a chatbot is used and strengthen the user's expectations	[5, 15, 48, 52]
Communicate all available functionalities	[5, 15]
Provide efficient dialog structures. Chatbot should help users in as few steps as possible	[15, 46, 48]
Create awareness of process and progress	[15, 46, 48]
Chatbot should be able to terminate interaction on user request	[17]
Present commands by clickable buttons instead of free text commands	[5, 15, 46, 48]
Chatbot should recognize synonyms	[23]
Chatbot should have multiple answers per text pattern	[22, 23]
Chatbot should use multimedia, e.g. images or audio.	[48]
User-friendly design of input/output fields	[23]
Chatbot should be personalized	[5, 15, 22, 23, 48, 50, 52]
Avoid deceptions (chatbot awareness)	[46, 48]
Communicate expected answer type	[15, 17]
Appropriate handling of intents that are not recognized	[15, 17, 23, 50]
Provide continuous support option	[15]
Check user inputs for correctness	[15, 17]
Chatbot should have general knowledge	[8, 13, 23]
Chatbot should catch improper posts	[13, 22]
Ability to recognize dialogue history	[50]

2.2 Empathy Map Method (EMM)

EMM is a creativity method that can be applied in DT during DT phases 'Understand' and 'Perception' [1, 38] for a) synthesizing observations from all team members, b) identification of customer needs and c) gaining new customer insights [12, 28, 31]. The EMM leads the team towards a consistent view of observations, customer needs and insights derived from them [41]. The identified customer needs and insights are fundamental for developing innovative ideas. Three to four participants are recommended for productive EMM sessions [38]. Figure 1 shows the six topics addressed during EMM and their procedural order (numbers one to six) [42, 49, 20, 21, 19].
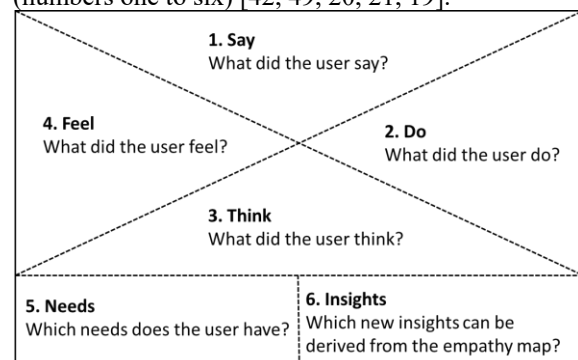
Figure 1. Empathy Map Method [19]

The EMM leads to customer insights [38]. An *Insight* provides information about what a customer needs to fulfill their desires and needs. It is an unexpected event, leading to a surprising effect and views the

design challenge from a new perspective. [29, 38]. To gain *Insights*, either the contents from two different quadrants or the contents in a quadrant are set in relation to each other [29, 30]. We chose the EMM as workpiece for digitalization, as it continues recent efforts to introduce chatbots as facilitators for creativity tasks [45], but for a more complex, procedural method that is both challenging - due to its creative, usually visual approach - and promising, as many teams could benefit from it without a professional facilitator.

3. Methodology

At first, a literature review has been conducted to derive initial requirements of the eEMM. The scientific databases ABI/INFORM, EBSCO Host, Springer Link and ACM Digital Library were searched for design-relevant publications in the fields of DT, CE and chatbot technology. The identified requirements are shown in Table 3. Based on them, the chatbot's conceptual knowledge base was derived and represented via the MindMap technique to reflect the architectural structure of different intents (e.g. macro and micro level). As proof of concept, a conceptual prototype has been developed, which instantiates the communication between the chatbot and other users. For an initial evaluation of the concept, it was prototypically implemented with the use of an instant messaging service within a group support system (MeetingSphere). In MeetingSphere, a session has been configured, which implements the process view of the developed concept. MeetingSphere functionalities were configured to reflect the chronological process structure of the EMM. This instantiation allows to execute an electronic EMM session, in which a human acts as the chatbot facilitator to evaluate and learn how to improve the concept before actual software development. In this so-called Wizard of Oz (WoO) study, the uninformed participant is told to communicate with an autonomous system, which is in fact controlled by a real person. The wizard is a trained human, who simulates a chatbot [9, 32]. Wizards are provided with strict rules and a conclusive knowledge base to approximate the behavior of a chatbot [11, 16]. WoO studies are characterized by a high degree of user participation [9]. Thus, they have proven useful for the evaluation of interactive conversational systems before their actual implementation [32]. With the help of WoO, the usability and feasibility of planned functionalities can be tested before implementing the system [7]. In line with [40], three wizards are provided to reduce the complexity of tasks for each wizard. Wizard 1 is responsible for the macro-level (see section 4) and is therefore equipped with the appropriate section from the eEMM concept to make Ava's contributions from

the macro level, either time or event driven. Wizard 2 and 3 are both responsible for the micro-level, 2 for phase-dependent and 3 for phase-independent dialogues. The user test was conducted with three subjects from a university. Each of them got a specific role within a fictious scenario. The EMM session lasted two hours. Utterances were coded for their intent and issuer. Evaluation results are used to identify problems and derive suggestions for improvement.

4. Requirements and Concept for Chatbot

Overall, 39 functional requirements were identified. 7 are crucial, 16 conditionally important and 16 less important, but helpful in terms of usability. Due to space limitations, Table 3 shows the 23 eEMM requirements that fall into the crucial (**bold**) and conditionally important categories. Essentially, they are referring to the micro-level, macro-level, cross-level behavior of the chatbot and usability. Communication with the chatbot takes place via natural language. This requires a chat service for the interaction within the group and between chatbot and the group. Users must be able to save posts for performing the activities scheduled in an EMM session. In addition, a chatbot needs both proactive and reactive ability to control a session and be able to respond to user questions. Further, multiple users must be accepted by the system as EMM is conducted in teamwork. For a frictionless session, a chatbot must be able to adequately enforce an ideal approach of the EMM with all intermediate and final results. Thus, the chatbot must have the required domain knowledge. All other requirements contribute to the system usability and are only partially necessary. For the chatbot concept, the disembodied character "Ava" was defined. The underlying knowledge base was developed in line with the requirements and Ava's set of utterances was framed to reflect the aspired personality. Ava has been designed retrieval and rule based with finite domain knowledge. The knowledge base consists of a macro and a micro level. Macro level knowledge executes proactive, micro level knowledge steers reactive behavior of Ava. On the macro level, the blueprint process of EMM was modeled and complemented with utterances for facilitating the EMM. In a script based chronological order, the macro level comprises all introductions and instructions Ava needs to provide to guide users through the tasks. Ava's contributions on the macro level are issued proactively after preset time intervalls. In contrast, Ava's reactive behavior is located on the micro level and activated by the recognition of predefined patterns in user utterances. It contains domain knowledge that allows users to ask for assistance, if something is unclear.

Table 3. Chatbot requirements

Title	Specification of requirements	Source	Ty-pe
Instant Messaging as communication channel	Chatbot and users interact with each other through an IM service.	[44, 47]	Chatbot behaviour across levels
Natural Language	A chatbot interacts with users using natural language.	[8, 13, 23, 43]	
Closed domain	It is already known at the beginning what knowledge a chatbot must have for the implementation of the tasks required for the facilitation of EM sessions. Accordingly, it is already known, which domain knowledge is needed in the knowledge base of the chatbot. Thus, a chatbot with fixed domain knowledge is implemented whose knowledge base is static and thus limited.	[2, 35]	
Retrieval-based	For facilitating EMM, a chatbot must be able to answer user requests considering the content of the domain. Potential user requests are already known in the development process and are restricted by the considered domain. Accordingly, a chatbot needs to be able to select and provide statically predefined answers from the knowledge base to answer user requests.	[2, 3]	
Rule-based	The chatbot must be equipped with a rule-based mechanism for processing user queries.	[2, 3]	
Reactive and proactive ability to act	For a script-based control of the group session and the moderation of the tasks, a chatbot must be able to proactively contribute. Therefore, mechanisms are required that determine the intervals at which a chatbot contributes and that coordinate the workflow of the EMM. A chatbot needs to be able to answer situational questions from users about the content of the corresponding domain and how the method works in order to facilitate an EMM session. In order to answer user questions during an EMM session, a chatbot must be able to respond to user requests. It must have both reactive and proactive capacity to manage the EMM .	[35, 39, 53]	
Questioning results	A chatbot needs to encourage members to look at their thoughts and work from different perspectives and establish a broad frame of reference in terms of DT principles.	[14, 38]	
Adequate handling of unrecognized user requests	If a chatbot can not understand a user request because no pattern was detected in the request, the chatbot should inform users and allow users to continue a regular conversation.	[15, 17, 23, 50]	Micro level
Verification of user input	The chatbot must intervene, if the user responds with an incorrect answer type to a question that requires a certain type of response. In this case, the chatbot must alert the user by naming the expected data type. If a user repeatedly phrases a request in slight variation that is not recognized, it can be assumed that the user wants to exercise a specific command that either does not exist or is called differently. The chatbot must recognize patterns and either propose obvious commands or inform the user that such commands are not available.	[14, 15, 17]	
Required domain knowledge	The chatbot acts as a facilitator for the EMM and must therefore have knowledge about DT. In particular questions to the procedure of the EMM must be answered by the chatbot.	[14, 35]	
Adequate introduction	A chatbot has to explain to the members initially, how the eEMM can be started, which functionalities it offers and how they can be invoked. The chatbot will guide the group during the EMM session and familiarize them with the activities in the process. It must thus explain at the start of the activity what the group has to do to solve a task and how the required results can be realized.	[5, 14, 15, 22, 35]	Macro level
Chatbot usage and user expectations	At first the users must be informed about the purpose of the chatbot. The chatbot has to make the user aware of possible interactions. The interaction limits of the chatbot are shown in order to adjust the expectations of the users. It has to inform the group members about the importance of the results of the collaboration process and point out the importance of each participant's contributions to strengthen the personal responsibility of the group members and promote group responsibility. It must also be shown how the results of the collaboration process can be used effectively in the project.	[14, 22, 23]	
Design Challenge	The team must keep an eye on the design dhallenge during an EMM meeting as an essential part of the group's problem. The group also needs to focus on the results of the EMM. The chatbot needs to show the team members the influence of the results on the project and how important the design challenge is. This ensures the effectiveness of the EMM session.	[14, 29, 31, 42]	
Rules of brainstorming	The EMM is a creativity method. Thus, the group members must consider and internalize some rules for the purpose of a productive implementation of the EMM. A chatbot must introduce the rules to the team members at the beginning of the EMM session.	[14, 29, 30, 38, 41]	

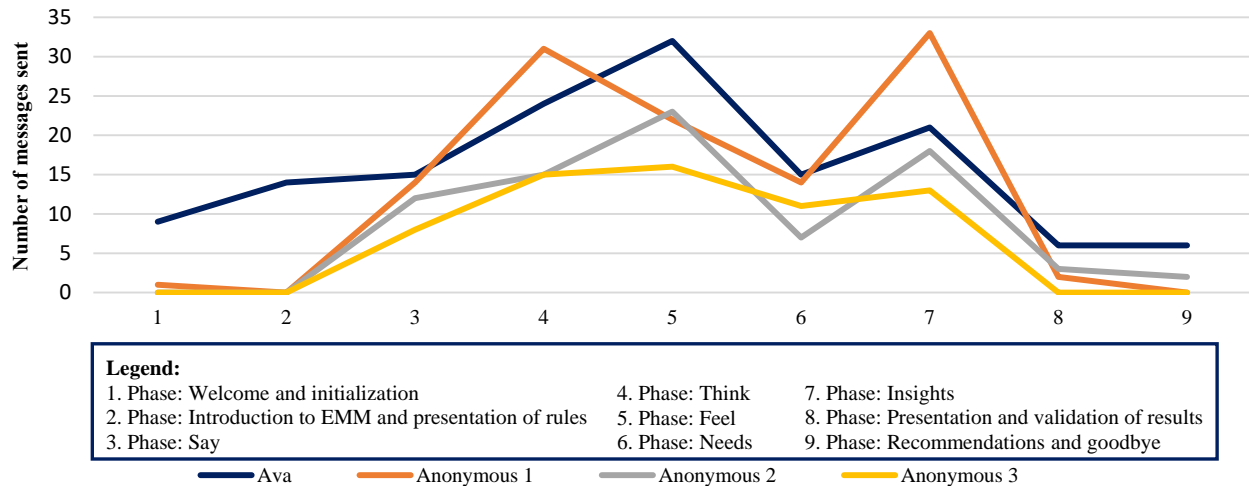| Title	Specification of requirements	Source	Type
Procedure and chronological sequence of the EMM	eEMM must digitally map the ideal procedure of an EMM. A chatbot guides the group through the process and through the defined steps. Mechanisms must be implemented in the eEMM that ensure the scheduling and coordination of the meeting program and group activities. A chatbot is responsible for a secure flow of information and communication within the group. Thus, relevant intermediate and final results must be provided at the right time.	[14, 30, 35, 38]	Macro level
Common group goal	Certain group results must be generated during the EMM meeting. Accordingly, all group members share a common group goal. The team's goal within an EMM session is to capture new insights of potential customers. These findings will be used primarily to specify the design challenge and the target group and to identify user needs as a basis for generating innovative ideas. The goal must be presented to the participants by the chatbot at the beginning.	[14, 29, 30, 38, 41]	
Validation of results	Certain results must be available after the session. These results represent the group products of the collaboration process. The group products must lead to the achievement of the group goal. The chatbot must visually juxtapose the group products and the group goal to assess whether the group goal has been reached.	[33, 35, 36]	
User-friendly input and output fields	Clearly identifiable controls must be available. Users must be able to track the text input over several lines. Likewise, a submit button must be available to submit an entry. The output elements must be presented to the user in a comprehensible manner. The font and font size must be selected according to the group characteristics.	[23]	Usability
Memory function	Due to the nature of an EMM, an eEMM must be able to save user contributions and transmit and provide them in subsequent phases.	[43, 53]	
User support	A chatbot should be able to continually provide guidance on individual activities, procedures, or general topics in the domain.	[14, 15]	
Iterations	Users must be able to either stop interacting with the chatbot at all times or jump back to prior EMM phases. This gives the group the possibility to proceed iteratively.	[17, 38]	
Anonymity of participants	Since the teams in DT are interdisciplinary and the members of the team can belong to different hierarchical levels, the traceability of contributions to individuals must be prevented. In this way, the team members can act unrestrictedly and deal creatively with the tasks.	[4, 24, 26, 35]	
Multiple user	A number of three to four participants is recommended for conducting productive EMM Sessions. For this reason, eEMM must enable the participation of multiple users. A chatbot must be able to process the requests of all participants.	[16, 19, 41]	
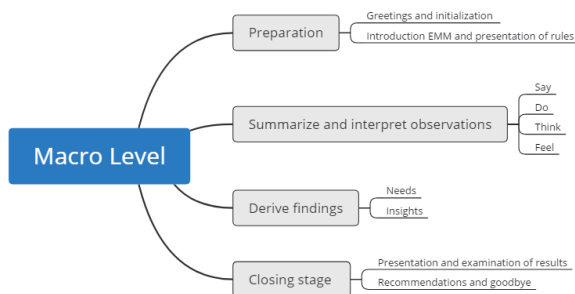
Figure 2. Macro Level

The micro level consists of phase dependent and phase independent dialogues. Phase dependent dialogues cover questions and their answers that refer to one specific phase of the EMM. Phase independent dialogues refer to questions that Ava can answer irrespective of the current phase of the EMM. Thus, Ava e.g. has a repertoire of reactions to small talk.
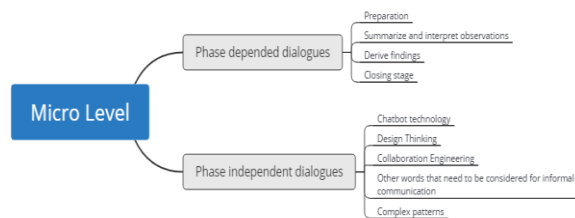
Figure 3. Micro Level

5. Evaluation Results

An initial evaluation of the chatbot concept was executed in a laboratory WoO user test with a group of three users with only minor knowledge in design thinking methodology. The aim of the evaluation was to explore, whether the eEMM with Ava has the potential to substitute a human facilitator. The qualitative analysis of the dialogue protocols and the self-assessment by the participants provides indication that the *macro level* is suitable for the intended use. The chat analysis showed that all team members participated actively in the session and contributed creative ideas to the group task. 402 utterances were issued in total with 91.9% of them being task-oriented rather than social messages and 31.9% being new thematic contributions. This result points to a high motivation and interest of the participants to collaborate creatively. Furthermore, they engaged intensively with the evaluation, elaboration and correction of ideas of their peers. This behavior shows that the participants understood and accepted the rules and suggestions for collaboration, that Ava introduced at the beginning of the session. Self-assessment of the users after the session supports this observation, e.g. *"Ava gave much and helpful information on Design Thinking and the EMM. Ava provided very useful hints and examples for the*

Figure 4. Speech portion in each phase

Legend:
1. Phase: Welcome and initialization
2. Phase: Introduction to EMM and presentation of rules
3. Phase: Say
4. Phase: Think
5. Phase: Feel
6. Phase: Needs
7. Phase: Insights
8. Phase: Presentation and validation of results
9. Phase: Recommendations and goodbye

Ava — Anonymous 1 — Anonymous 2 — Anonymous 3

procedure upon request." (interview excerpt, user 1). The group members interacted intensively with each other and with Ava throughout the collaboration session (see figure 4). Noteably, only 28 requests like "Ava, example for 'need'?" were directed towards Ava by users, which might be an indication that either her proactive instructions left little room for questions or that users for some reason do not make full use of her assistance. 17 out of the 28 requests for Ava could be adressed and answered by her successfully. Eleven requests could not be adressed with the current knowledge base, e.g.: *User 1: "Ava state Design Challenge" Ava: "Sorry, I didn't understand that. Are you sure that you have spelled everything correctly?"*

These requests point toward necessary extensions of Ava's knowledge base with additional patterns and knowledge elements. Apart from these limitations, which will be addressed in the next section, Ava's reactive abilities are sufficient for facilitating the eEMM according to the users' assessment and enable even groups that are new to the EMM to conduct a successful session. As figure 4 depicts, team members were most active in the process steps that require their contribution (phases 3 to 7) and attentitively picked up Ava's instructions in the initial phase and during wrap-up. This user behavior indicates that the instructions provided on the *macro level* concerning the tasks and the process structure have been perceived and interpreted correctly by the participants. Thus, users focused on the activities requested by Ava to a large extent and hardly got distracted. In such, Ava's utterances were expedient, have attracted the participants' interest for the task and made the importance of the task and of active participation clear. Qualitative user feedback with respect to the macro level content was in line with these findings and

largely positive. Minor critcism refered to the instructions was too lengthy and detailed: "Overall, the tasks were very clear. Occasionally, Ava's instructions at the beginning of a new subtask were a little to extensive so that one could loose oversight at first" (user 3).

Furthermore, Ava's answers to user requests have been evaluated as clear and helpful by the users, which also suggests a successful design of the *micro level* knowledge base: *"Ava's answers were very detailed and precise. They helped to take the right perspective, to process the test person interviews and to gain a deeper understanding for these persons"* (user 1). However, participants wished for a more active moderator that contributes more flexibly, and context depended to the task solution. *"In general, Ava was sufficient for the moderation. A more active role would have been nice, e.g. Ava could have provided active feedback, if e.g. an entry is still missing"* (user 2). Users integrated Ava actively in their discussions, when questions on the methodology arose or they faced obstacles, e.g.: *User 1: "Ava, example [for] feeling?" Ava: "Of course I give you an example ☺. This phase is about bringing your observations to a more abstract level, the level of feelings. Thus, show empathy for the customer and imagine, how you would feel, if you were in their situation. Let's assume the customer has the following thought: [...]. In my opinion, the customer might feel stressed in this situation, because [...]"*

The *usability* of the eEMM has been evaluated positively by the users. However, further improvement should strive for implementing the identified requirements for the user interface and layout of the eEMM, which have been out of scope of this initial WoO instantiation due to the use of standard software.

Usability suggestions from the interviews are in particular: *"To increase usability, it would be helpful to display the prior results and the necessary information for the current task in short form apart from the chat window."* (user 3). *"A visualization with the four quadrants would have been advantageous."* (user 2)

Overall, the evaluation results provide indication that the script-based facilitation of the eEMM session was accepted well by the users.

6. Discussion

These initial findings need to be taken with caution, as different team constellations may lead to different group dynamics and our early stage observations are based on a very small user test with only three users. However, the user feedback helped us to identify further user requirements to improve the concept. These requirements should be iteratively integrated into the prototype [30]. The user interviews revealed that Ava's examples for specific concepts and the possibility to ask Ava for such examples has turned out to be especially useful for the participants. Thus, this core functionality should receive the necessary attention in the software development efforts and should even be extended. The introductions that Ava provides for each phase have proven to be helpful. One important functionality that was noted by the participants was to be able to ask Ava for results of prior phases: *"It was possible to ask for the previous results. I liked that."* (user 1) The wrap-up of the results that have so far been collected at the beginning of each new phase (e.g. Ava: *"Very good! You saved four potential feelings in total, one for each thought. The first feeling is ..."*) was evaluated positively by users. These comments lead to the assumption that users appreciate a continuous progress and process overview as well as they need continuous access to the intermediary results, as follow-up activities build on these.

The participants were well aware and made use of the macro level elements for supporting the progress and process overview. The knowledge that Ava provided on DT and the EMM has shown to be sufficient for executing the EMM session. Likewise, Ava's micro level replies to user questions in case of problems have turned out detailed and helpful. However, we found the description of Ava's scope of interaction do require improvement, as users occasionally faced difficulties deciding what they could ask Ava and what not: *"With complex questions, we hesitated to address Ava."* (user 1)

Overall, users asked for a moderator, who could assess contributions in relation to their context and detect relations between utterances. However, this wish seems predominantly related to the system's usability, as users expressed that they would not have assumed the results to be better in general, if they had had a human moderator. One user expressed a differentiated opinion: *"I think that with a good [human] facilitator, one might be able to reach even better results, but with a bad facilitator also much worse."* (user 1) This comment points to the potential of achieving standardized process quality through automated facilitation, especially, if no large number of really good human facilitators are accessible.

They only assumed that with a human facilitator, they might potentially have been able to reach these results in a shorter time. Ingle [31] and Crandall [19] both estimate a duration of thirty minutes for an EMM session with four to ten participants. In our WoO study, the prototypical EMM session took two hours. However, the necessary duration should not be generalized, as it strongly depends on the scope and type of material that is assessed in the EMM session.

Another critical note from participants was that Ava's intelligence is not advanced enough to detect and mediate conflicts. This comment is kind of surprising, as only 1.5% of all user utterances have been identified as critical towards contributions by other participants in the chat analysis. Thus, there was

Table 4. Problems and Suggested Adaptions

Problems	Suggested Adaption
Too detailed instructions in the macro level (high reading effort)	- reduce the reading effort for users - shorten instructions
Too little multimedia content	- use multimedia content in the macro level, especially for the description of tasks
Interaction scope of Ava is not clear enough	- describe interaction scope more clearly at the macro level
Complicate storage mechanism	- provide a button on the user interface with which selected posts can be saved
Too much back-scrolling to see relevant items	- show results statically outside chat window
No visual representation of the results	- visualize the eEMM using the 4 quadrants and the 2 derived
No visual representation of Ava	- give Ava an avatar that matches the defined character of Ava
Lack of coordination between macro and micro level	- define rules for coordinating Ava's proactive and reactive contributions
Posts starting with 'Ava' have limited pattern matching	- the pattern recognition process must be performed for every contribution made
Uncertainty about dealing with multiple patterns	- complement the pattern recognition process with rules for the prioritization of identified patterns

no need for Ava to intervene in a conflict. However, the interview comment points to a need for the facilitator's ability to mediate in case of conflicts.

With respect to the information storage mechanism, participants reported difficulties and described it as too cumbersome. After they agreed on the framing of e.g. a feeling in the discussion, they had to repeat it with the command *"Ava, save feeling: The user is afraid of spoiled foodstuffs, the related inconveniences and the loss of control."* Ava would reply with: *"I have saved the feeling. It has the number 4."* Thus, this functionality will need to be revised. Chat record revealed that the participants had to agree several times on who would save certain contributions. Thus, either a role concept with one group member being responsible for documentation or a more advanced automated storage mechanism would be suggested for further improvement of the process flow.

Concerning the sighting of results and the design challenge, participants noted high effort. Users had to scroll a lot between results, which makes it hard to keep an overview and might lead to relationships being overlooked. For that reason, the presentation format of relevant elements should be developed further towards better usability. Thus, we propose to statically pin relevant results and the design challenge outside the chat window to constantly access them throughout the session. For improving the visibility of relationships between results, we recommend a visual presentation of the results in line with the EMM phases in six fields. Although struggling with some deficiengies of the chat interface, participants acknowledged the chat form of the conversation in general due to its documentation effect, as earlier discussions could be traced and followed upon in lather phases, which would not be easily available in spoken conversation: *"Ava had strengths, as we could always look up what had been written so far. With a human facilitator, one would always have to ask again for the information and could misunderstand the spoken answer."* (user 2)

During the WoO study, the wizards could identify three problems that need to be adressed. First, up to now, there is no coordination between the macro and the micro level. Thus, no rules have been defined for Ava to prioritize proactive or reactive behaviors, if both should be performed at the same time. Special Rules are necessary for the coordination between micro and macro levels. We suggest the extending rule that Ava prioritizes utterances on the macro level over those on the micro level to maintain the flow of speech. Otherwise, users would be in danger of losing track of the conversation. After macro level activities, user request should be adressed on a first come first served basis [54]. To achieve this, user requests need to be collected in a queue and processed periodically. A second problem identified by the wizards concerns the defined mechanism for pattern detection. So far, pattern comparison of user utterances with Ava's knowlege base were only initiated, if the utterances were addressed towards Ava. This mechanism should be extended towards certain text patterns that are directed towards other users, such as insults, sexist or racist utterances. Pattern recognition must cover all parts of the conversation to detect deviant behavior reliably. Third, wizards recognized that there were requests that could not be adressed by the current chatbot unambiguously. It might happen that several patterns may be detected in a single request. In the WoO study, one participant asked: 'Ava, what is an insight?' The pattern comparison revealed two patterns in the knowledge base for this request. On the one hand, the pattern „{what} is an {insight}" is recognized, in which Ava explains, what insight means. On the other hand, the request matches the pattern „{what} do we have to do", in which Ava, dependent on the current phase the users are in, explains what the participants should do. To solve this problem, an extension to the pattern recognition approach is necessary. We suggest adding the rule that patterns with a higher specificity are prioritized over patterns with lower specificity. This rule will choose the pattern that contains the highest number of recognized words. For the specific example, that implies that Ava would chose to answer to „{what} is an {insight}" as two words from this pattern are recognized. All problems and related suggestions for improvement that have been identified in the evaluation are summarized in Table 4. Some of the suggested improvements can not implemented with the current conceptional method and tool choice. Thus, the revised concept needs to be instantiated with other, more advanced forms of technical prototypes. Furthermore, a next step will be to execute a full size experimental study with several groups in order to detect commonalities.

7. Conclusion

We were able to show that the eEMM is effective, although further improvements need to be done. Overall, apart from the identified needs for improvement, Ava has performed successfully in her task to facilitate the eEMM session. Participants were satisfied with the results that have been produced in WoO session and evaluated it as successful. Many chatbots that are currently developed have only a reactive ability to act. It can be assumed that in the future more approaches will be pursued in which chatbots can also act proactively. Only through

proactive abilities chatbots can perform the tasks and activities required to facilitate group work in the context of DT. Further research should extend on that by implementing the presented concept, extending its level of intelligence and by exploring more applications of automated facilitation.

8. References

[1]   www.cio.de/a/design-thinking-im-it-service-management,3260384, accessed 5-30-2018.

[2]   www.info.contactsolutions.com/digital-engagement-blog/a-chatbot-framework, accessed 5-8-2018.

[3]   www.arxiv.org/pdf/1705.01214v2.pdf, accessed 5-29-2018.

[4] Al Shishany, A. and J. Adams, "Anonymity Enhancing Technologies (AETs) in GDSS Supported Meetings", The 3rd International Conference on Information Technology & E-Services, ICETeS' 2013, Sousse, Tunisia, 2013.

[5] www.alistapart.com/article/all-talk-and-no-buttons-the-conversational-ui, accessed 5-29-2018.

[7] Benzmüller, C., A. Fiedler, M. Gabsdil, H. Horacek, I. Kruijff-Korbayova, M. Pinkal, J. Siekmann, D. Tsovaltzi, B. Quoc Vo, and M. Wolska, "A Wizard of Oz Experiment for Tutorial Dialogues in Mathematics", Proceedings of AI in Education (AIED 2003) Workshop on Advanced Technologies for Mathematics Education. 2003.

[8] Boden, C., J. Fischer, K. Herbig, J. Liebe, H. Sinning, and U. Spierling, "Chatbots als Instrument der Planungskommunikation - Chancen, Anforderungen und Perspektiven", in GI Jahrestagung, 2006.

[9] Böttcher, B. and M. Nüttgens, "Überprüfung der Gebrauchstauglichkeit von Anwendungssoftware", HMD Praxis der Wirtschaftsinformatik, 50(6), 2013, pp. 16–25.

[10] Briggs, R.O., G.L. Kolfschoten, G.-J. de Vreede, S. Lukosch, and C.C. Albrecht, "Facilitator-in-a-Box: Process Support Applications to Help Practitioners Realize the Potential of Collaboration Technology", Journal of Management Information Systems, 29(4), 2013, pp. 159–194.

[12] Brown, T., "Design Thinking", Harvard Business Review, 86(6), 2008, pp. 84–92.

[13] Christensen, A., "Virtuelle Auskunft mit Mehrwert: Chatbots in Bibliotheken", Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft (222), 2008.

[14] Clawson, V.K., R.P. Bostrom, and R. Anson, "The Role of the Facilitator in Computer-Supported Meetings", Small Group Research, 24(4), 1993, pp. 547–565.

[15] www.smashingmagazine.com/2016/12/conversational-design-essentials-tips-for-building-a-chatbot/, accessed 5-29-2018.

[16] Curedale, R., Design thinking: Process and methods manual, 1st edn., Design Community College, Topanga, CA, 2013.

[17]   www.alistapart.com/article/designing-the-conversational-ui, accessed 5-29-2018.

[18] Dyke, G., D. Adamson, I. Howley, and C.P. Rose, "Enhancing Scientific Reasoning and Discussion with Conversational Agents", IEEE Transactions on Learning Technologies, 6(3), 2013, pp. 240–247.

[19]   www.dschool-old.stanford.edu/groups/k12/wiki/3d994/Empathy_Map.html, accessed 5-29-2018.

[20] Endrejat, P.C. and S. Kauffeld, "Wie könnten wir Organisationsentwicklungen partizipativ gestalten?", Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO), 48(2), 2017, pp. 143–154.

[21] Ferreira, B.M., S.D.J. Barbosa, and T. Conte, "PATHY: Using Empathy with Personas to Design Applications that Meet the Users' Needs", 18th international conference on Human-computer interaction, HCI International, Toronto, ON, Canada, 2016.

[22] Geeb, F., "Chatbots in der praktischen Fachlexikographie/Terminologie", LDV Forum, 22(1), 2007, pp. 51–70.

[23] Gennermann, H. and S. Hack, "Qualitätsstandards für Chatbots in der bibliothekarischen Auskunft in Deutschland", Projektberichte aus dem berufsbegleitenden Masterstudiengang Bibliotheks- und Informationswissenschaft der Fachhochschule Köln: MALIS-Praxisprojekte 2011.

[24] Gerlicher, A., "Computer-Supported Cooperative Work (CSCW) — kollaborative Systeme und Anwendungen", in Kompendium Medieninformatik: Medienpraxis, R. Schmitz, Editor. 2007. Springer: Berlin.

[25] Graesser, A.C., N. Person, D. Harter, "Teaching tactics and dialog in AutoTutor", International Journal of Artificial Intelligence in Education, 12(3), 2001, pp. 257–279.

[26] Gräslund, K. and H. Krcmar, "Anonymität", in CSCW-Kompendium: Lehr- und Handbuch zum computerunterstützten kooperativen Arbeiten, G. Schwabe, N. Streitz, and R. Unland, Editors. 2001. Springer Berlin Heidelberg; Imprint; Springer: Berlin, Heidelberg.

[27] Harding, J.A. and R. Swarnkar, "Implementing collaboration moderator service to support various phases of virtual organisations", International Journal of Production Research, 51(23-24), 2013, pp. 7372–7387.

[28] Henderson-Sellers, B., J. Ralyté, P.J. Ågerfalk, and M. Rossi, "Examples of Constructed Processes", in Situational Method Engineering, B. Henderson-Sellers, J. Ralyté, P.J. Ågerfalk, and M. Rossi, Editors. 2014. Springer Berlin Heidelberg: Berlin, Heidelberg.

[29] IDEO, Human-Centered Design toolkit, 2nd edn., Authorhouse, Palo Alto, 2011.

[30] IDEO, The field guide to human-centered design: Design kit, 1st edn., IDEO, San Francisco, 2015.

[31] Ingle, B.R., Design thinking for entrepreneurs and small businesses: Putting the power of design to work, Apress, New York, NY, 2013.

[32] Kerlyl, A., P. Hall, and S. Bull, "Bringing Chatbots into education: Towards Natural Language Negotiation of Open Learner Models", Proceedings of AI-2006, the Twenty-fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, 2007.

[33] Kolfschoten, G.L., R.O. Briggs, G.-J. de Vreede, P.H.M. Jacobs, and J.H. Appelman, "A conceptual foundation of the thinkLet concept for Collaboration Engineering", International Journal of Human-Computer Studies, 64(7), 2006, pp. 611–621.

[34] Kumar, R. and C.P. Rosé, "Architecture for Building Conversational Agents that Support Collaborative Learning", IEEE Transactions on Learning Technologies, 4(1), 2011, pp. 21–34.

[35] Leimeister, J.M., Collaboration Engineering: IT-gestützte Zusammenarbeitsprozesse systematisch entwickeln und durchführen, Springer Gabler, Berlin, 2014.

[36] Lewe, H., "Gruppenproduktivität", in CSCW-Kompendium: Lehr- und Handbuch zum computerunterstützten kooperativen Arbeiten, G. Schwabe, N. Streitz, and R. Unland, Editors. 2001. Springer Berlin Heidelberg; Imprint; Springer: Berlin, Heidelberg.

[37] Lieberman, H., "Autonomous interface agents", Proceedings of the SIGCHI conference on Human factors in computing systems, Atlanta, Georgia, United States, 1997.

[38] www.static1.squarespace.com/static/57c6b79629687fde090a0fdd/t/58890239db29d6cc6c3338f7/1485374014340/METHODCARDS-v3-slim.pdf, accessed 5-29-2018.

[39] Repka, V., V. Grebenyuk, and K. Kliushnyk, "Pro-Active Multi-Agent System in Virtual Education", Robust Consensus of Multi-agent Systems with Bounded Disturbances, L. Wang and Z. Liu, Editors. 2011. INTECH Open Access Publisher.

[40] Salber, D. and J. Coutaz, "Applying the Wizard of Oz technique to the study of multimodal systems", Third International Conference on Human-Computer Interaction. 1993.

[41] Schallmo, D.R.A., Design Thinking erfolgreich anwenden: So entwickeln Sie in 7 Phasen kundenorientierte Produkte und Dienstleistungen, Springer Fachmedien Wiesbaden, Wiesbaden, 2017.

[42] Schallmo, D.R.A. and L. Brecht, Prozessinnovation erfolgreich anwenden, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[43] Shawar, B. and E. Atwell, "Chatbots: Are they Really Useful?", LDV Forum, 22(1), 2007, pp. 29–49.

[44] Shawar, B.A. and E.S. Atwell, "Using corpora in machine-learning chatbot systems", International Journal of Corpus Linguistics, 10(4), 2005, pp. 489–516.

[45] Strohmann, T., D. Siemon, and S. Robra-Bissantz, "brAInstorm: Intelligent Assistance in Group Idea Generation", 12th International Conference, DESRIST 2017, Karlsruhe, Germany, 2017.

[46] www.ecommerce-chatbots.com/7-laws-chatbotics-human-chatbot/, accessed 5-30-2018.

[47] www.chatbotsmagazine.com/the-complete-beginner-s-guide-to-chatbots-8280b7b906ca, accessed 5-8-2018.

[48] www.medium.muz.li/the-ultimate-guide-to-chatbots-why-theyre-disrupting-ux-and-best-practices-for-building-345e2150b682, accessed 5-30-2018.

[49] Tochetto, J., C. Guimarães, A.L. Maranho, and A.L. Tartari, "Design with Me: I Have Special Needs! The Case for Cerebral Palsy", 10th international conference on universal access in human-computer interaction,, UAHCI, 2016..

[50] Vetter, M., ",,Ich habe Ihre Eingabe leider nicht verstanden" — Qualitätskriterien für Online-Tests von Bots", in Avatare, C. Lindner, Editor. 2003. Springer Berlin Heidelberg: Berlin, Heidelberg.

[51] Vreede, G.-J. de and R.O. Briggs, "Collaboration Engineering: Designing Repeatable Processes for High-Value Collaborative Tasks", Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005.

[52] www.medium.com/topbots/when-bots-go-bad-common-ux-mistakes-in-chatbot-design-c60b252a6abf, accessed 5-30-2018.

[53] Witten, H. and M. Hornung, "Chatbots: Teil 1: Einführung in eine Unterrichtsreihe zu ,,Informatik im Kontext" (IniK)", LOG IN(154/155), 2008, pp. 51–60.

[54] Xoxa, N., M. Zotaj, I. Tafa, and J. Fejzaj, "Simulation of First Come First Served (FCFS) and Shortest Job First (SJF) Algorithms", International journal of computer science and network, 3(6), 2014, pp. 444–449.

[55] Zolnowski, A. and T. Böhmann, "Grundlagen service-orientierter Geschäftsmodelle", in Service-orientierte Geschäftsmodelle, T. Böhmann, M. Warg, and P. Weiß, Editors. 2013. Springer Berlin Heidelberg: Berlin, Heidelberg.