Giving Social Robots a Conversational Memory for Motivational Experience Sharing

Saravanan, Avinash; Tsfasman, Maria; Neerincx, Mark A.; Oertel, Catharine

DOI
10.1109/RO-MAN53752.2022.9900677

Publication date
2022

Document Version
Final published version

Published in
RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication

2022 31st IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 29 - September 2, 2022. Naples, Italy.

2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) | 978-1-7281-8859-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/RO-MAN53752.2022.9900677

Giving Social Robots a Conversational Memory for Motivational Experience Sharing

Avinash Saravanan[1], Maria Tsfasman[1], Mark A. Neerincx[1,2], and Catharine Oertel[1]

***Abstract*— In ongoing and consecutive conversations with persons, a social robot has to determine which aspects to remember and how to address them in the conversation. In the health domain, important aspects concern the health-related goals, the experienced progress (expressed sentiment) and the ongoing motivation to pursue them. Despite the progress in speech technology and conversational agents, most social robots lack a memory for such experience sharing. This paper presents the design and evaluation of a conversational memory for personalized behavior change support conversations on healthy nutrition via memory-based motivational rephrasing. The main hypothesis is that referring to previous sessions improves motivation and goal attainment, particularly when references vary. In addition, the paper explores how far motivational rephrasing affects user's perception of the conversational agent (the virtual Furhat). An experiment with 79 participants was conducted via Zoom, consisting of three conversation sessions. The results showed a significant increase in participants' change in motivation when multiple references to previous sessions were provided.**

I. INTRODUCTION

One of the current frontiers of social robotics is to endow social robots with a conversational memory. A conversational memory would allow robots to refer to events and experiences in previous interactions and could potentially improve social awareness, experience sharing, and bonding. The development of such a system is essential for many use-cases in education or health care. One such vital use-case relates to behavior change support systems.

A behavior change support system (BCSS) is designed to help a user master a behavior change [1], for instance, after being diagnosed with obesity or diabetes and adapting both eating and sports habits. A critical aspect of successful behavioral change is the creation of intrinsic motivation. Given that behavior change requires longer-term engagement, keeping up the motivation to commit to the changes is essential. A widely used method of counseling that focuses on encouraging and sustaining such motivation in the management of lifestyle-related diseases is Motivational Interviewing (MI) [2].

Motivational feedback is a central aspect of motivational interviewing therapy. This involves reflection on a patient's goals and behaviors: for example, encouraging a patient for (partially) attaining their goal [3], [4], [5]. Therefore, in an MI-based behavior support system, reminding the

This work was supported by the 4TU Human Technology Research Center and Medical Delta

[1]The Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands

[2]TNO Human-Machine Teaming, Kampweg 55, 3769 DE, Soesterberg, The Netherlands

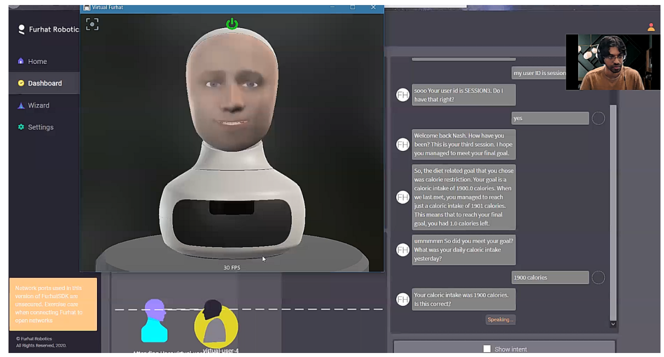

Fig. 1. Illustration of Coaching Dialog between participant and robot coach (the dialog took place in a Zoom meeting, top right shows the test leader for starting and rounding off the session).

user of progress and insights gained in earlier sessions might strengthen the reflection process and positively impact motivation. To provide such reflections over a period of time, a robot would need to remember the user and the progress made, monitor their affective response (sentiment), and provide a variability of appropriate motivational feedback phrases. Yet, this aspect has not been investigated in the context of MI-based behavior support systems [6], [7], [8], [9], [10]. This paper presents the development and evaluation of a sentiment-aware conversational memory model that allows a robot to provide such feedback, positively referring back to a user's previous insights gained and progress made across sessions.

II. BACKGROUND

Being diagnosed with a chronic disease can not only be traumatic but often also requires a change in lifestyle and adherence to specific nutrition regime [11], [12]. Finding the motivation to follow through with the needed changes can pose a serious challenge for many people. Self-management based educational programs have been shown to increase the lifestyle change success [13]. However, such programs often lack adaptability to the individual requirements of the patient. A more person-centered approach to lifestyle change guidance is Motivational Interviewing (MI).

Motivational Interviewing is a recognized counseling method that supports people to talk about and start changing their behavior [14]. With diabetes, this approach has shown promising results in diet, and physical activity management [15], [16]. Although encouraging, MI is particularly resource-consuming since it requires an educated professional for regular individual sessions with the patient. A potential solution to this problem lies in the field of social robotics. A social robot designed to do Motivational Interviewing could regularly provide individualized support for lifestyle management. For the domain of lifestyle-related healthcare, there have been attempts to create conversational agents that could help motivate people to adopt healthy behaviors [17]. A crucial step in the direction of enabling MI-based human-agent dialog was the development of MI-dialog acts and statements [6], [7], [8], [9], [10].

However, an essential component of motivational interviewing has not been taken into account in the context of social robotics. For MI to succeed, it needs to occur regularly and build upon previous experiences. An MI-centered social robot needs to incorporate a memory model or architecture to continue and build on previous dialogue acts in a systematic motivational way.

Within the conversational agents' community, determining how to design an agent's memory is an active area of research. The architectures of particular interest in this context are often presented by the term 'episodic memory'. The term refers to remembering personal experiences. In the context of conversational agents, these are usually experiences shared by the user throughout previous interactions with the agent (also referred to as shared memories). For example, [18] and [19] extract and store the contents of shared memories in knowledge graphs, specifically, when the context triggers a similar pattern in the conversation (actor, place and time triples in [18] and the H5W "who", "what", "where", "when", "why" and "how" structure in [19]). Although showing promising results, the models in these studies don't address the user experience with the affective aspects in the conversation. [20] developed a memory system for a chatbot in the form of an online diary for teenagers to reflect on their experiences. The architecture they developed also stores all the mentioned experiences in knowledge graphs, but on different levels: memory line (an overall timeline of when events occurred), general event (contents of the events in a 4W data-structure - "what", "where", "when" and "who"), lifetime period (events referred to a lifetime period, such as "preschool", "high school" etc.). The memory retrieval is triggered by syntactic cues in the conversation when a similar pattern is found in the memory database. No affective information was taken into account and no motivational or goal attainment data was recorded. [21] introduced a memory system in a tutoring scenario that did take into account all these factors. Their model memorizes conversation with the user in the context of the user's goal attainment and emotional state. First, all the expressed information is stored in the short-term memory module. Second, after the interaction, only memories relevant to the goals (e.g. a photography task completed by the user) or the emotion are kept in the long-term memory module, and the rest is deleted. Memories are then retrieved in the next interactions in the moments of conversation when the goal or emotional state of the user is of importance (e.g. bringing up the emotional state of the user by the end of the last interaction or bringing up the history of performance in a certain task when redoing it). They also conducted an extensive analysis of how the memory system and agent's personalty (supportive vs unsupportive) affects user perception of the agent [22]. This is particularly relevant for our study because they monitored users' motivation and learning success. Although the effect of memory on motivation was not significant, the learning success was greater in the supportive memory-aspired agent than in the unsupportive agent without memory. A potential reason for the insignificant effect of memory on motivation could be that memories were not directly used to adjust motivational statements. Also, the conditions with variable frequency of brought up memories were not explored. The system was not used within the context of a tutoring scenario.

To our knowledge, there was only one other study that used an artificial memory system in the context of MI [23]. This study was designed around aiding children with type 1 diabetes. However, the memories' specific effects on children's affect and motivation were not tested [24].

While different memory models were proposed, they did not address the memory-based references and the variety of these references with their effects on motivation, goal attainment and perception of the robot. Therefore, our experiment focuses on evaluating the design of a conversational memory and the corresponding references on these three variables, focusing on the following hypothesis: Referring to previous sessions improves motivation and goal attainment, particularly when the references vary. In addition, the paper explores how far motivational rephrasing affects the users' perception of the conversational agent (the virtual Furhat).

III. Memory Model

In this section, we elaborate on the components of our memory model and its construction for the human-robot interaction. The interaction design is unique in that it encodes shared experiences to use them for the generation of motivational rephrasing.

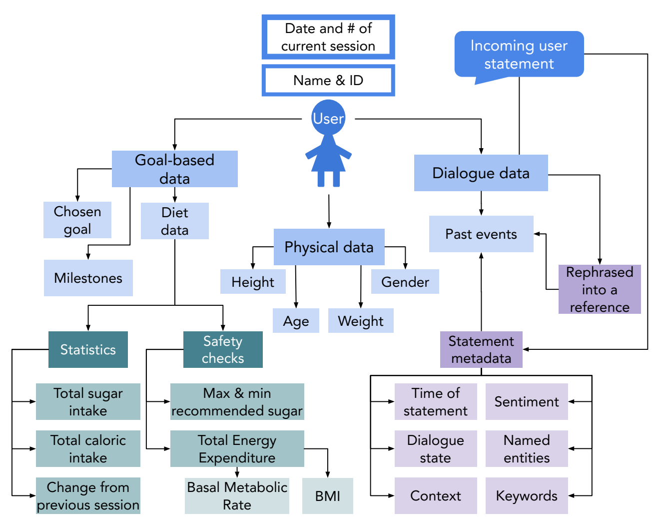

Fig. 2. Illustration of memory components implementation distinguishing goal-based, physical, and dialogue data

A. Memory Components

Our memory architecture collects and updates information about the user in several ways (Figure 2). It stores three types of data: goal-based, physical and dialogue data. The physical data module stores static information such as the user's age (in years), height, weight, and gender. The goal-based data module stores all information related to a participant's chosen goals. It is stored connected to a user's milestones and diet data. The diet data sub-module keeps track of a user's total sugar and caloric intake progress towards his goals and previous report. It also acts as a safety monitor to ensure that the user's sugar intake, Body Mass Index (BMI) and Basal Metabolic Rate (BMR) remain in a medically safe range. On the other hand, the dialogue data module contains memories of experiences shared throughout the interaction. Each user statement goes through two processes. First - rephrasing it into a reference that can be used later on in a conversation. Second, it extracts various metadata from the statement: sentiment, named entities, and keywords. That information is stored along with the time of the statement, dialogue state it was mentioned in, and context of the interaction.

The experience sharing part of our memory model is used for two purposes: to (re)align the user's motivations towards the chosen goal and to praise or criticize the progress (see Figure 3). In this way, we aim to strengthen and clarify sources of intrinsic motivation according to the positive MI-approach. This interaction design is unique in that it is designed to encode shared experiences with the purpose of using them for motivation in the future.

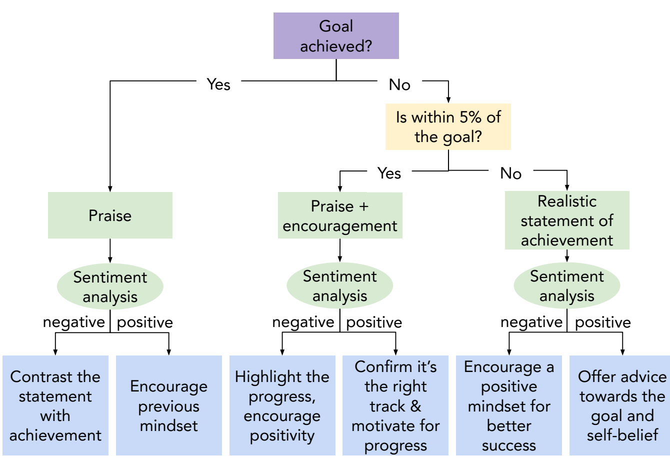

Fig. 3. A generalized decision tree illustrating how progress of the user and sentiment of an answer to a question in the past affect the way in which a motivational phrase is generated.

B. Memory Generation

In our memory generation, we ensure that statements relating to intrinsic motivation are considered as well as answers to questions that ask the user about their reasons for working towards a goal. We also include the users' feelings about a goal by utilizing the user's own statements regarding their mindset from previous sessions. This technique is adapted from the strategies and techniques used in motivational interviewing [25] to encourage goal-oriented mindsets.

For the decoding, the saved data, which are previous statements given in reply to motivational interviewing style questions, needs to be extracted in the right context to provide appropriate feedback to support the behavior-change process. It is beneficial to reflect on past experiences to spot problematic behavior and set goals accordingly to improve goal setting and achieve a more long-lasting behavior change [26]. However, not all past experiences bear the same emotional importance to a user. Some that bear a particularly positive or negative valence might be particularly well suited to be brought up in interaction to trigger behavior change. In our system implementation, the rephrasing of past experiences takes place in the second and third sessions.

To achieve this we employ an array of techniques: reword the original statement into a reference that can be used as motivation or encouragement that can remind the patient of why they are working towards their goal and realign their motivations accordingly. To do so, the following is needed in the decoding: the question, the answer given, and context and sentiment for a motivational phrase.

By using these techniques in combination we hope to capture the user's mindset beyond what is explicitly stated by the user. We relate these phrases to the user's goal progress to estimate the user's confidence in achieving their goal. We tailor the levels of encouragement or criticism accordingly in the generated motivational phrase. Three strategies used to refer to the past are:

1) Simple recitation of an answer
2) Summarizing by keywords/key-phrases
3) Summarizing according to sentiment (positive/negative valence) rather than content

We used the Twitter corpus provided with the NLTK toolkit to train a naive-bayes classifier to distinguish between answers containing either positive or negative sentiment. Using 10-fold cross-validation, we achieved an average prediction accuracy of 89%. We used combined sentiment in a user's answer with progress made and tailored our response accordingly as illustrated in Figure 3

Rephrasing as feedback is intended to motivate the participant and falls under the categories of praise and encouragement according to Schunk and Lilly [27]. We add relevant motivational statements after summarizing the question or answer to improve or maintain specific actions based on the user's rate of progress towards their particular goal. Because of this, the phrases used can have a corrective or confirmatory response based on whether the user has met their milestones or not.

An example of a case where a motivational phrase is constructed and used is presented in Figure 4, showing three components of the rephrasing: Question asked, answer given, and motivational statement. The first two are modified programmatically to allow for referral to a past event from a previous session. The question repeated to the user in a later session is the original question repeated with the pronouns changed. The answer provided by the user to that

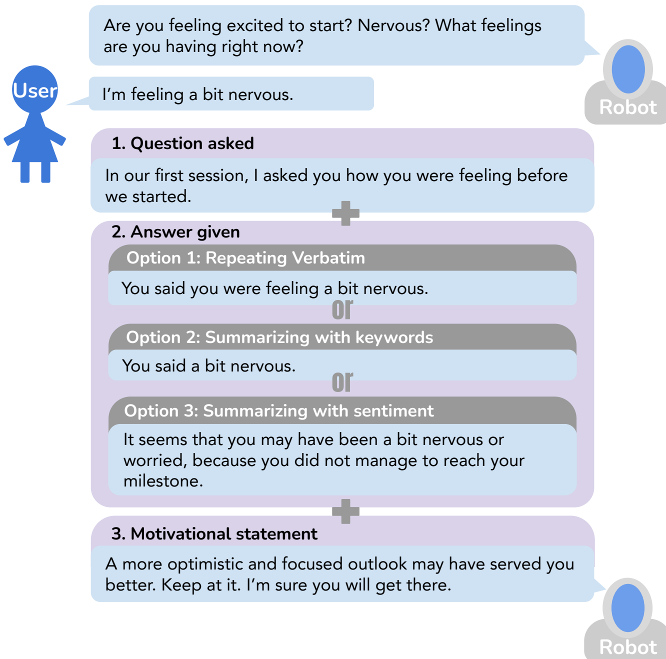

Fig. 4. Examples showing how user's answers are summarized and brought up in the subsequent interactions.

question is repeated by using one of the three summarization techniques which can be seen in Figure 4. The purpose of these two components is to contextualize the motivational statement that follows to emphasize its relevance to the user's personal situation. The third component (i.e., the motivational statement) utilizes a decision tree to choose an appropriate handcrafted statement based on the question asked, the answer given, and the user's progress towards their desired goal. Regarding the answer given, the sentiment found within the answer using sentiment detection acts as one of the conditions on which the tree branches off. A high-level decision tree can be found in Figure 3. We handcrafted a selection of sentences and statements. However, which of these statements is chosen is determined by a rule-based system. Their choice is dependent on the experimental condition and individual progression through the milestones and dialog. This means that while the phrases and subphrases are handcrafted, which one is chosen and how it is assembled is dynamically handled using the rule-based system.

IV. Use Case for Evaluation

The memory system described above was tested within a behavior change support case study. We set the scenario around changing eating and movement behavior to improve a person's health condition. For the *robot*, we used a virtual representation of the back-projected robot-head Furhat ([28], see Figure 1), which is taking on the role of a behavior change support coach. The robot's tasks are twofold: On the one hand, it is supposed to aid the participant to reflect by posing goal-related questions. On the other hand, it is designed to answer questions the participant might have in conjunction with their behavior change.

The interaction was designed to accompany the user through their behavior-change journey across a series of 3 sessions. The sessions have at least 24 hours between each other. Each session follows a general sequence of steps as depicted in Figure 5.

In session 1, the robot learns about the user and helps them set personal goals and milestones for the following sessions.

In sessions 2 and 3, the goals and milestones are reviewed, where the robot refers to the motivational memories from the previous session. The number of references depends on the condition that the user is assigned to as detailed in Section V-C.

The dialog flow is depicted in Figure 5. It is designed to guide the user through sequences of steps related to goal setting and milestones. An autonomous dialogue system was implemented that drives the robot's behavior based on a state-machine implementation. It needs to be noted that the dialog system is dual-initiative. This implies that the participants can also take the turn and ask questions related to nutrition. The question answering is facilitated by querying a database of nutrition information provided by the U.S. Department of Agriculture [29]. To find an acceptable answer to the question asked, we create a word-embedding space using tf-id and rank the appropriateness of responses according to the semantic similarity of food items.

V. Evaluation

This section presents how we evaluated our hypothesis and research question within the use case: Referring to previous sessions improves motivation and goal attainment, particularly when the references vary. Moreover, how far does the motivational rephrasing affect user's perception of the robot?

A. Manipulation

We created three conditions to evaluate the effects of memory references and ran the experiment as a between-subjects design. We equally distributed study participants across those conditions :

1) Condition I: No references to previous events
2) Condition II: References to only one event.
3) Condition III: Multiple references to different events (variety) where no events are ever brought up twice.

The first condition had no references to the past, while the second only repeated references. In the third condition, the agent mentioned two references to different events in every session. This was done, on the one hand, to guarantee variability and no repetitions of referenced information, since the previous research shows that repetitions might lead to decreased perceived intelligence of a conversational agent [30]. On the other hand, to avoid overloading the interaction with past events, which has been shown to be perceived as boring or repetitive [31].

B. Participants

A total of 93 participants took part in the study, of which 79 had valid results (36 female, 42 male, 1 nonbinary). The

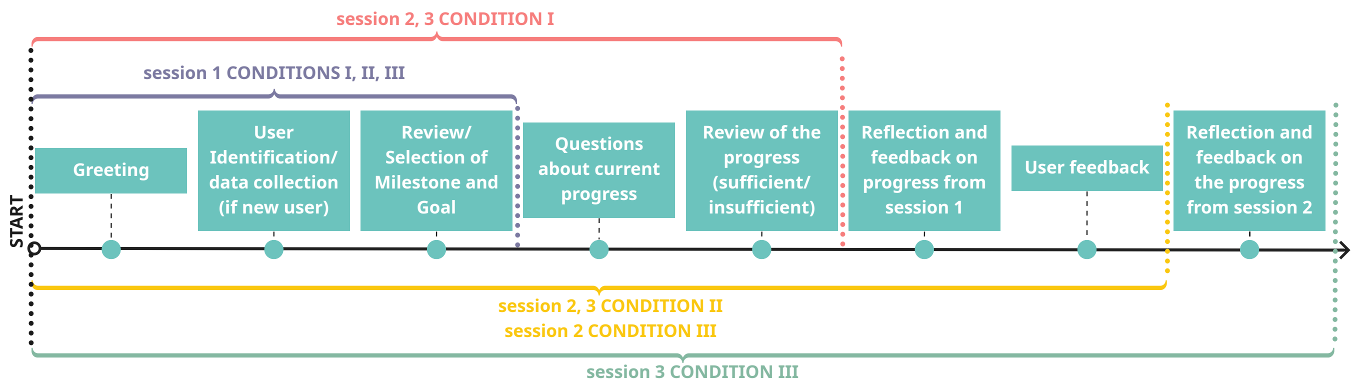

Fig. 5. Initial sequence of Events in the experiment.

reasons why 14 of 93 participants were discarded included not completing all three sessions, technical problems or accent specifics resulting in the robot not understanding participant's speech, and severe connection issues interrupting the interaction. The average age of participants was 31.5, and most participants held a bachelor's level degree. Participants had to have a high degree of English fluency but originated from different countries around the world. Conditions I and II had 26 participants each, and condition 3 had 27 participants.

The majority of participants were recruited through Prolific. For safety reasons, they were expected to not have any metabolic disorders to avoid possible negative consequences. In addition, participants were expected to fulfill the following criteria to participate in the study:

- Above the age of 18
- Not currently following a diet
- Has a working computing device, stable internet connection, and headphones that can be used for video meetings.

Each participant was asked to choose a health-promoting goal of either calorie restriction or sugar reduction to pursue (lifestyle-related).

C. Measures

The following section describes the different methods used to test the main hypothesis and additional research question as stated earlier.

a) Motivation: Values of motivation were self-reported (subjective): from 1 - *"not motivated at all"* to 5 - *"highly motivated"*. The questionnaire contained questions regarding participants' engagement and behavior-change motivation that was perceived prior to the interaction, during the interaction, and after completion of the interaction. Participants were asked to answer these questions on a five-point Likert scale.

b) Goal attainment: To determine the effectiveness of our memory model in helping users achieve their behavior change goals we defined the following matrices: *milestone-adherence,difference,count*, *final-goal- difference, achievement, intention to continue diet* . While the milestone related measures are aimed to measure whether the intermediate goals are met and to what degree the final-goal measures evaluate the behavior-change goals overall as well as the participants' intention to continue the diet.

We used the Godspeed questionnaire from [32] to evaluate the participants' perception of the robot (including full items of animacy, anthropomorphism, likeability, perceived intelligence and perceived safety). Along with the Godspeed questionnaire, we asked questions related to the perceived success of the interaction against obesity and diabetes. The above questions were asked after completing all the sessions of the experiment.

c) Qualitative Interaction Evaluation: In order to evaluate how participants' perceived the number of sessions and capture any additional comments they might have, we included 2 open questions in the questionnaire. The questionnaire was administered at the end of the third session.

D. Experimental Procedure

All participants were recruited through the crowd-sourcing platform Prolific academic and interactions were conducted in English. Each participant was supposed to take part in three sessions spaced in time. Before starting the experiment, participants completed the informed consent form through a qualitrics link. A video was produced to provide experiment-related instructions to ensure that all participants were briefed in the same way. In each interaction, the participant would interact with the virtual embodiment of the Furhat robot through a video-conferencing call.

VI. Results

A. Motivation

Figure 6 shows a box plot of perceived motivation change in different conditions. A Kruskal-Wallis test [33] revealed that there was no significant effect of our memory model on participants' engagement in the interaction between the three conditions $H(2)=0.92517$, $p>0.5$. However, there was a significant effect of our memory model on participants' motivation to engage in behavior change $H(2)=8.7556$, $p<0.05$. A Dunn's post-hoc test revealed that participants assigned to condition 3 showed a significantly higher motivation increase than participants assigned to condition 1 ($p<0.05$) or condition 2 ($p<0.05$).

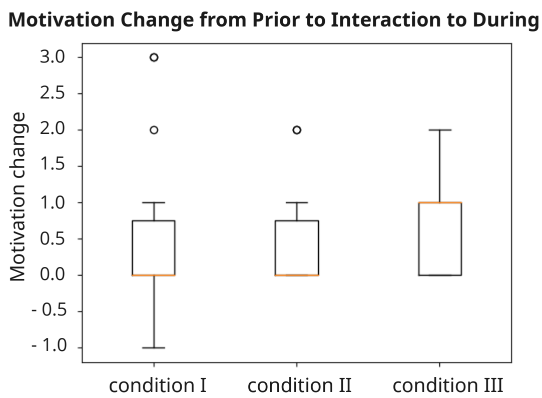

Fig. 6. The figure shows a box plot comparison of the three conditions for change in motivation from prior to interaction with the robot to during interaction.

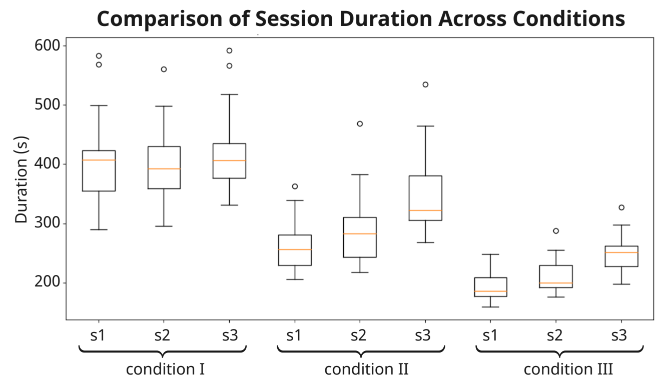

Fig. 7. The figure shows box plots of durations for different sessions across the conditions.

B. Goal attainment

To investigate whether referring to previous sessions improves goal attainment we carried out an ANOVA analysis and a post-hoc Bonferroni test on all milestone and final goal measures. We did not find any significant differences between the experimental groups. We also did not find any significant differences for any measure of the Godspeed questionnaire.

C. Qualitative Findings

In addition to the collected quantitative data, participants were allowed to type additional comments regarding their interaction with the robot and the experiment. Sometimes, they expressed their impressions verbally after conversations. Some of the participants voiced preferences of a robot over a human. The questionnaire did not fully capture these preferences. Three of the participants noted that, unlike a human, the robot is not judgmental or has any prejudices that could result in a negative experience. For a small number of participants, the three day period needed for the experiment was somewhat problematic due to personal circumstances. Some participants may cite a hard day at work which resulted in no time to eat. Similarly, some participants may cite being more active on a particular day, resulting in a higher intake for that specific day.

For many, the impact of simply counting calories and sugar made a significant difference in the perception of their diet. The visual aspects of the robot had a mixed response upon further discussion of the experiment after completion of the three sessions. Some participants did not mind the appearance, while others found the human-like face unnerving. Regarding the voice, this had a mixed response, but most of the participants found the Polly voice to be surprisingly human, with the voice sounding less rigid than other voices. According to many participants, the voice did not sound robotic. The main flaw that the voice had was prosody and pacing, which is a problem that the participants noted is typical in other robotic voices.

a) Does referring to previous sessions improve goal attainment?: Based on the significant results, goal attainment, specifically final goal achievement as well as other performance-related metrics, was not different between the three conditions. We saw 86.1% of participants achieve their goal. When split between conditions, the achievement for condition 1 was 88.5%, for condition 2 it was 88.5% and for condition 3, it was 81.5%.

VII. Discussion

The results show that a variety of reflections on shared experiences provides greater motivation and that motivational memory is capable of motivation when used this way. However, we did not find significant effects between the three experimental groups related to goal achievement, which might have been due to the fact that behavior change usually takes place over longer periods of time. In a future study, it would be interesting to investigate whether differences in goal attainment can be observed by extending the number of sessions.

Whereas other research focuses on creating memory architectures similar to human memory [34], [35], [36], e.g. distinguishing short-term and long-term memory [37], our memory model focuses on supporting users to pursue and attain their behavior change support goals. In this way, it can be viewed more as a goal-driven representation of the experiences that underpins human-robot conversations. Furthermore, memory enables the generation of conversations that follow Motivational Interviewing principles [6], [7], [8], [9], [10]. Interactively, the robot relates to a person's goals and actual behaviors, supporting positivism with appropriate reflection, encouragement, and praise [3], [4], [5]. The generation of motivational sentiment-aware dialogues can be viewed as an extension of other memory models, such as Kasap et. al [37] work on task engagement and social presence and Sanchez' [38] focus on topic related goals.

It should be noted that by using references in an interaction, the duration of that interaction will be increased compared to conditions without references (see Figure 7). Differences in duration between conditions are not noticeable for session 1, but are noticeably larger for the succeeding sessions for conditions 2 and 3 on the order of 30-60 seconds due to the use of shared experiences which naturally add to the duration of the interaction. It should be noted, however, that because of factors such as lag on Zoom and Internet quality, the differences in duration that were measured may not be fully representative of the actual duration of the interaction. Future research is needed to test the effects of dialog duration on motivation.

More than one-third of participants indicated that they would prefer the interaction to be longer while the remainder found the length of the interaction to be appropriate. These findings are encouraging in that behavior change is a longer-term activity. The fact that participants are willing to interact for a longer period of time with our robot is not only encouraging from a system-design perspective but also from a behavior-system support point of view who typically target long term goal achievement over a period of months [39].

We could show that our memory system positively affects participants' motivation to engage in diet-related behavior change. While these results are very encouraging as a first step towards long-term person-adaptive human-robot interaction, there are also some limitations. In its current set-up, the interaction is very much use-case dependent. The system asks a predefined list of questions related to diet goals. It follows a predefined sequence of conversational goes with specific reflection moments in which memories are being referred to. This is quite effective within a period like the one described in this work, but if one were to extend the number of sessions, such methods may become quite repetitive over time.

In our future research, we will study the effects of a physical robot, focusing on more long-term interaction. We will investigate what to remember and what to forget (cf., [40]), and explore when and how often memories should be referred to in a conversation to maximize task and social objectives (cf., [37] and [23]). Further, we aim to interactively evaluate the individual reaction of the user to the specific memory. We believe that the robot support can be applied for type II diabetes (T2D) patients in the use case. Diabetes II is the most common form of diabetes [41] which makes up 90% of all diabetes cases resulting in more than 392 million people being diagnosed with it worldwide. Preventative measures for type II diabetes require management techniques similar to the behaviors supported by the robot, such as caloric reduction and sugar reduction.

VIII. Conclusion

We designed and evaluated a memory model for behavior-change-support dialog in the social robot Furhat. The memory model took both content and sentiment-related information to store and refer to past experiences. We investigated how the frequency of referred to memories effected both participants' interaction evaluation and behavior-change goals. We found that the experimental group exposed to more than one memory had a significantly higher increase in motivation than the participants exposed to none or only one memory. In the future, we will increase the number of sessions and participants to investigate whether this will influence goal achievement.

IX. Acknowledgement

This research has been partly funded by the Dutch-Swiss ePartners4all project (Grant no. TKI-LSH-T2019).

References

[1] H. Oinas-Kukkonen, "A foundation for the study of behavior change support systems," *Personal and Ubiquitous Computing*, vol. 17, pp. 1223–1235, July 2012.

[2] D. Christie and S. Channon, "The potential for motivational interviewing to improve outcomes in the management of diabetes and obesity in paediatric and adult populations: a clinical review," *Diabetes, Obesity and Metabolism*, vol. 16, no. 5, pp. 381–387, 2014.

[3] R. Koestner, M. Zuckerman, and J. Koestner, "Praise, involvement, and intrinsic motivation.," *Journal of personality and social psychology*, vol. 53, no. 2, p. 383, 1987.

[4] D. H. Schunk and M. W. Lilly, "Sex differences in self-efficacy and attributions: Influence of performance feedback," *The Journal of Early Adolescence*, vol. 4, no. 3, pp. 203–213, 1984.

[5] J. R. Tudge, P. A. Winterhoff, and D. M. Hogan, "The cognitive consequences of collaborative problem solving with and without feedback," *Child development*, vol. 67, no. 6, pp. 2892–2909, 1996.

[6] O. A. Blanson Henkemans, P. J. Van Der Boog, J. Lindenberg, C. A. Van Der Mast, M. A. Neerincx, and B. J. Zwetsloot-Schonk, "An online lifestyle diary with a persuasive computer assistant providing feedback on self-management," *Technology and Health Care*, vol. 17, no. 3, pp. 253–267, 2009.

[7] R. Looije, F. Cnossen, and M. A. Neerincx, "Incorporating guidelines for health assistance into a socially intelligent robot," in *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 515–520, IEEE, 2006.

[8] R. Looije, M. A. Neerincx, and F. Cnossen, "Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors," *International Journal of Human-Computer Studies*, vol. 68, no. 6, pp. 386–397, 2010.

[9] S. Park, J. Choi, S. Lee, C. Oh, C. Kim, S. La, J. Lee, and B. Suh, "Designing a chatbot for a brief motivational interview on stress management: Qualitative case study," *Journal of medical Internet research*, vol. 21, no. 4, p. e12231, 2019.

[10] J. G. G. da Silva, D. J. Kavanagh, T. Belpaeme, L. Taylor, K. Beeson, J. Andrade, *et al.*, "Experiences of a motivational interview delivered by a robot: qualitative study," *Journal of medical Internet research*, vol. 20, no. 5, p. e7737, 2018.

[11] H. Kolb and S. Martin, "Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes," *BMC Med*, vol. 15, p. 131, July 2017.

[12] M. Uusitupa, T. A. Khan, E. Viguiliouk, H. Kahleova, A. A. Rivellese, K. Hermansen, A. Pfeiffer, A. Thanopoulou, J. Salas-Salvadó, U. Schwab, and J. L. Sievenpiper, "Prevention of type 2 diabetes by lifestyle changes: A systematic review and Meta-Analysis," *Nutrients*, vol. 11, Nov. 2019.

[13] S. Chatterjee, M. J. Davies, S. Heller, J. Speight, F. J. Snoek, and K. Khunti, "Diabetes structured self-management education programmes: a narrative review and current innovations," *Lancet Diabetes Endocrinol*, vol. 6, pp. 130–142, Sept. 2017.

[14] J. Hettema, J. Steele, and W. R. Miller, "Motivational interviewing," *Annu Rev Clin Psychol*, vol. 1, pp. 91–111, 2005.

[15] G. Ekong and J. Kavookjian, "Motivational interviewing and outcomes in adults with type 2 diabetes: A systematic review," *Patient Educ Couns*, vol. 99, pp. 944–952, Dec. 2015.

[16] P. D. Soderlund, "Effectiveness of motivational interviewing for improving physical activity self-management for adults with type 2 diabetes: A review," *Chronic Illn*, vol. 14, pp. 54–68, Mar. 2017.

[17] A. Fadhil and S. Gabrielli, "Addressing challenges in promoting healthy lifestyles: The al-chatbot approach," in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '17, (New York, NY, USA), p. 261–265, Association for Computing Machinery, 2017.

[18] P. Vossen, S. Baez, L. Bajcetić, and B. Kraaijeveld, “Leolani: A reference machine with a theory of mind for social communication,” in *Text, Speech, and Dialogue* (P. Sojka, A. Horák, I. Kopeček, and K. Pala, eds.), (Cham), pp. 15–25, Springer International Publishing, 2018.
[19] C. Moulin-Frier, T. Fischer, M. Petit, G. Pointeau, J.-Y. Puigbo, U. Pattacini, S. C. Low, D. Camilleri, P. Nguyen, M. Hoffmann, H. J. Chang, M. Zambelli, A.-L. Mealier, A. Damianou, G. Metta, T. J. Prescott, Y. Demiris, P. F. Dominey, and P. F. M. J. Verschure, “Dac-h3: A proactive robot cognitive architecture to acquire and express knowledge about the world and the self,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 4, pp. 1005–1022, 2018.
[20] J. Campos and A. Paiva, “May: My memories are yours,” in *Intelligent Virtual Agents* (J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, eds.), (Berlin, Heidelberg), pp. 406–412, Springer Berlin Heidelberg, 2010.
[21] Z. Kasap and N. Magnenat-Thalmann, “Towards episodic memory-based long-term affective interaction with a human-like robot,” in *19th International Symposium in Robot and Human Interactive Communication*, IEEE, Sept. 2010.
[22] Z. Kasap and N. Magnenat-Thalmann, “Long-term social interaction with an expressive robot.,” in *Proceedings of Computer Graphics International (CGI’11)*, (Ottawa, Canada), 2011.
[23] B. Schreuder Goedheijt, “Recalling shared memories in an embodied conversational agent : personalized robot support for children with diabetes in the pal project,” 2017.
[24] M. A. Neerincx, W. van Vught, O. B. Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, and B. Bierman, “Socio-cognitive engineering of a robotic partner for child's diabetes self-management,” *Frontiers in Robotics and AI*, vol. 6, Nov. 2019.
[25] S. Olafsson, T. O'Leary, and T. Bickmore, “Coerced change-talk with conversational agents promotes confidence in behavior change,” in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, ACM, May 2019.
[26] L. Andre, A. E. M. van Vianen, T. T. D. Peetsma, and F. J. Oort, “Motivational power of future time perspective: Meta-analyses in education, work, and health,” *PLOS ONE*, vol. 13, p. e0190492, Jan. 2018.
[27] D. H. Schunk and M. W. Lilly, “Sex differences in self-efficacy and attributions: Influence of performance feedback,” *The Journal of Early Adolescence*, vol. 4, pp. 203–213, Sept. 1984.
[28] S. A. Moubayed, J. Beskow, G. Skantze, and B. Granström, “Furhat: A back-projected human-like robot head for multiparty human-machine interaction,” in *Cognitive Behavioural Systems*, pp. 114–130, Springer Berlin Heidelberg, 2012.
[29] U. D. of Agriculture, “Database.” https://fdc.nal.usda.gov/, 2021. Accessed: 2021-05-01.
[30] D. Aneja, D. McDuff, and M. Czerwinski, “Conversational error analysis in human-agent interaction,” in *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, IVA ’20, (New York, NY, USA), Association for Computing Machinery, 2020.
[31] J. Campos, J. Kennedy, and J. F. Lehman, “Challenges in exploiting conversational memory in human-agent interaction,” in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1649–1657, 2018.
[32] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots,” *International Journal of Social Robotics*, vol. 1, pp. 71–81, Nov. 2008.
[33] *Kruskal-Wallis Test*, pp. 288–290. New York, NY: Springer New York, 2008.
[34] J. E. Laird, A. Newell, and P. S. Rosenbloom, “SOAR: An architecture for general intelligence,” *Artificial Intelligence*, vol. 33, pp. 1–64, Sept. 1987.
[35] J. E. Laird, C. Lebiere, and P. S. Rosenbloom, “A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics,” *AI Magazine*, vol. 38, pp. 13–26, Dec. 2017.
[36] J. R. Anderson, M. Matessa, and C. Lebiere, “ACT-r: A theory of higher level cognition and its relation to visual attention,” *Human–Computer Interaction*, vol. 12, pp. 439–462, Dec. 1997.
[37] Z. Kasap and N. Magnenat-Thalmann, “Building long-term relationships with virtual and robotic characters: the role of remembering,” *The Visual Computer*, vol. 28, pp. 87–97, Sept. 2011.
[38] M.-L. Sánchez, M. Correa, L. Martínez, and J. Ruiz-del Solar, “An episodic long-term memory for robots: The bender case,” in *RoboCup 2015: Robot World Cup XIX* (L. Almeida, J. Ji, G. Steinbauer, and S. Luke, eds.), (Cham), pp. 264–275, Springer International Publishing, 2015.
[39] J. O. Prochaska, “Decision making in the transtheoretical model of behavior change,” *Medical Decision Making*, vol. 28, pp. 845–849, Nov. 2008.
[40] D. Richards and K. Bransky, “Forgetmenot: What and how users expect intelligent virtual agents to recall and forget personal conversational content,” *International Journal of Human-Computer Studies*, vol. 72, no. 5, pp. 460–476, 2014.
[41] G. Danaei, M. M. Finucane, Y. Lu, G. M. Singh, M. J. Cowan, C. J. Paciorek, J. K. Lin, F. Farzadfar, Y.-H. Khang, G. A. Stevens, M. Rao, M. K. Ali, L. M. Riley, C. A. Robinson, and M. Ezzati, “National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants,” *The Lancet*, vol. 378, pp. 31–40, July 2011.