Health Counseling by Robots: Modalities for Breastfeeding Promotion
Prasanth Murali, Teresa O’ Leary, Ameneh Shamekhi, Timothy Bickmore

Abstract— Conversational humanoid robots are being used increasingly for health education and counseling. Prior research provides mixed indications regarding the best modalities to use for these systems, including user inputs spanning completely constrained multiple choice options vs. unconstrained speech, and embodiments of humanoid robots vs. virtual agents, especially for potentially sensitive health topics such as breastfeeding. We report results from an experiment comparing five different interface modalities, finding that all result in significant increases in user knowledge and intent to adhere to recommendations, with few differences among them. Users are equally satisfied with constrained (multiple choice) touch screen input and unconstrained speech input, but are relatively unsatisfied with constrained speech input. Women find conversational robots are an effective, safe, and non-judgmental medium for obtaining information about breastfeeding.

I. INTRODUCTION

Conversational robots (CRs) and virtual agents have been demonstrated to be effective media for delivering health education and counseling across a range of topics, with robots successfully promoting rehabilitation exercise [1], diet [2], and physical activity [3], and conversational agents successfully used to promote an even wider range of health behaviors [4]. Although studies have demonstrated that robots are more engaging and persuasive than virtual agents at some tasks [5], it is unclear whether they are advantageous for health counseling when physicality does not provide specific instrumental or communicative affordances (e.g., by demonstrating exercises [1]).

In addition, while there is some evidence that CRs may be effective at eliciting self-disclosure of sensitive topics in general [6-9], the effectiveness of CRs for health education and counseling about sensitive topics, such as addiction counseling or domestic violence, has not been evaluated.

In this work, we report on the evaluation of a CR that provides education on and promotion of breastfeeding, a potentially sensitive topic for women. In addition, we systematically explore a range of input and output user interface modalities to determine whether these have any effect on user acceptance, learning, and motivation to follow recommendations.

Breastfeeding is an important health topic: many major US public health and medical organizations including the Centers for Disease Control and Prevention (CDC) have been actively promoting breastfeeding and recommend exclusive breastfeeding for the first six months of life [10]. However, only 16.3% of mothers in the US follow this recommendation [11]. Most mothers receive breastfeeding education or support from a medical professional, but many feel the education and support they receive is cursory and inadequate, and many receive misinformation or encounter practitioners hostile or indifferent to breastfeeding [3].

A. Related Work on Health Counseling Robots and Agents

A number of CRs and virtual conversational agents have now been developed to counsel patients on health problems. Kidd et al conducted a longitudinal evaluation of an in-home diet coach robot that helped users track their eating and exercise behavior, finding significantly greater compliance compared to other standard tracking techniques [2]. Fasola, et al, describe a humanoid robot that motivates older adults to perform physical rehabilitation exercises, demonstrating that users prefer it over a virtual agent [1]. Winkle, et al, evaluated a humanoid robot that persuades individuals to perform exercise (wrist rotations), and demonstrated that when it used certain persuasive techniques it was more effective at gaining compliance [3].

Virtual agents have been used in an even wider range of health interventions, including several large-scale randomized clinical trials. A sampling of applications include preconception care counseling for young women [12], medication adherence counseling for patients with atrial fibrillation [13], and exercise promotion for geriatrics patients [14]. Virtual agents that display relational skills such as empathy, social dialogue, nonverbal immediacy behaviors, and other behaviors have been shown to increase working alliance in automated health behavior change interventions [15]. Zhang, et al., demonstrated that virtual agents can effectively play the role of a decision coach and facilitate shared decision making between patients and clinicians regarding prenatal testing for Down syndrome [16]. Similarly, virtual agents have also been shown to be effective for delivery of healthcare information and decision support [17, 18].

B. Related Work on Technologies for Breastfeeding Promotion

Several hundred non-automated breastfeeding interventions have now been evaluated, with generally positive outcomes [19]. A few automated systems have been developed to promote breastfeeding among mothers. The interventions developed to date have been educational materials deployed on static web pages or multimedia CD-ROMs ([20, 21]). Joshi, et al, describe a bilingual touch-screen tablet-based intervention to promote breastfeeding.
However, the system is not a longitudinal intervention (designed for single contact use), and has not been evaluated in a clinical trial (only formative usability testing results are reported) [22]. Emrick describes the development of “Latch Master”, an iPhone game designed to teach mothers about correct breastfeeding positions and latching, although no evaluation is reported [23]. Shi et al., developed a feminist virtual agent that plays the role of a lactation educator and showed that participants with stronger feminist orientations to prefer the feminist agent [24]. Balaam, et.al. describe the development and evaluation of a mobile app that provides a crowdsourced map of public breastfeeding locations [25]. None of these projects involved the use of a robot, nor systematically explored the impact of various input and output modalities.

C. Related Work on CR Embodiment: Humanoid Robot vs. Virtual Animated Agent

There have been several studies comparing user interaction with robots to virtual agents that indicate there may be differences in how users discuss sensitive health topics with these systems, depending on embodiment. Li conducted a meta-review of robot vs. agent studies conducted up through 2013, finding that physical embodiment generally leads to improved behavioral compliance and attitude change in users [5], however, none of the studies involved discussion of sensitive topics. Powers et al., compared differences in users’ responses in a health interview between a virtual agent, a remote robot projected life sized on a screen, and a co-located robot, finding that users forgot more and disclosed least with the co-located robot, and forgot least and disclosed most with the virtual agent [26]. While users may be more willing to disclose to a virtual agent, a few studies indicate that robots may be more effective than agents at attitude and behavior change. Fasola and Mataric compared virtual to humanoid robot exercise coaches for older adults, demonstrating that elders preferred the robot, even though there were no differences in actual exercise behavior [1].

D. Related Work on CR Input Modality: Constrained vs. Unconstrained

Most health counseling CRs evaluated in clinical trials have used fully-constrained user input (i.e., multiple-choice menu of allowed utterances, updated at each turn of the conversation). This is done primarily so CR counseling can be fully validated (in part, to avoid safety issues [27]), but also because the dialogue engines use formalisms that prescribe a small number of user choices at each state of the agent-initiated counseling interaction. Tomko, et al, conducted one of the few studies directly comparing a constrained input speech system (“Speech Graffiti”) to unconstrained speech, in a limited domain of queries to a move information database. Rather than continuously displaying prompts of what users can say, Speech Graffiti uses a restricted command grammar that users must be trained in the use of. In a within-subjects user study with 23 participants, they found that participants preferred the constrained speech system along multiple dimensions of subjective satisfaction, while not finding any significant differences on task completion rate or task completion time [28].

II. DESIGN OF A CONVERSATIONAL ROBOT FOR BREASTFEEDING PROMOTION

We developed a CR (“Tanya”) to play the role of a virtual lactation consultant. We videotaped counseling sessions with a lactation consultant and developed a 20-minute CR dialogue intended to motivate women to follow the CDC recommendations. The topics covered in this dialogue include: greeting; asking user about her most important breastfeeding topic; review of the CDC recommendations; review of benefits of breastfeeding for the baby; breastfeeding benefits for the mother; breastfeeding “101” (latching, basic nursing positions); review of CDC recommendations; and farewell.

Dialogues are scripted using a custom hierarchical transition network-based scripting language and template-based text generation. The CR speaks using a speech synthesizer, and uses a range of nonverbal conversational
behavior, including facial display of affect, communicative eyebrow movements, directional gazes, and head nods, with most nonverbal behavior automatically generated using BEAT [29].

Five variants of the user interface were developed for comparative evaluation (Figure 1). In all conditions except Unconstrained Speech, a separate touch screen monitor was used to show allowed user utterances, as well as additional educational text and media. In the Unconstrained Speech condition, the second monitor was only used to show the educational materials. The five conditions were:

- **Robot Touch** – Users interact with a Furhat humanoid robot (Figure 1) by selecting their utterances from a touch screen of choices.
- **Robot Constrained Speech** – Users interact with Furhat by speaking one of the allowed utterances displayed on a screen of choices. A research assistant maps their speech to a selection (Wizard of Oz control).
- **Robot Unconstrained Speech** – Users interact with Furhat by speaking, without prompts or constraints. A research assistant maps their speech to one of the allowed choices at each state of the dialogue (Wizard of Oz control).
- **Agent Touch** – Users interact with a virtual animated agent (Figure 1) by selecting their utterances from a touch screen of choices.
- **Agent Constrained Speech** – Users interact with a virtual agent by speaking one of the allowed utterances displayed on a screen of choices. A research assistant maps their speech to a selection (Wizard of Oz control).

Unconstrained speech was added to the Robot condition only, in response to requests for this condition by participants in our pilot studies. Participants in the Robot condition indicated more often and more strongly that they were disappointed by not being able to ask Tanya their own questions and that they felt that the system lacked functionality required to reformulate and articulate information in new ways in response to user questions.

[P3] “I think that actually talking to her would be better than just selecting an option and maybe you would want to ask something not listed. You’d get what you need from the system faster.”

[P12] “It would have been better if it was only an interaction with her where she could understand. Instead of having to tap a button. That would improve the interaction and make it a 1-1 kind of thing.”

III. EVALUATION

We conducted an experiment to evaluate the ability of our CR to increase breastfeeding knowledge and motivation. In order to explore a range of possible interaction modalities, we conducted a five-treatment between-subjects experiment, evaluating satisfaction, attitudes towards the robot, and change in breastfeeding knowledge and motivation as outcomes. The study was approved by our institutional IRB.

A. Participants

Participants were recruited from campus fliers and an online job posting site, and were required to be 18 years or older, female, no children yet but interested in having children in the future, and native speakers of English. Participants were compensated for their time.

B. Measures

**Satisfaction**

We used a 7-item questionnaire to measure user satisfaction on 7-point scales (α=.8).

**WAI:** Attitude towards the CR was assessed using the bond subscale of the Working Alliance Inventory (WAI) [30]. WAI contains 12 items on 7-point Likert scale.

**Intention to adhere** to the US consensus recommendations on breastfeeding was assessed by asking “When you have a baby, how long do you plan to feed him or her nothing other than breast milk?” (response in number of months). Intent has been associated with actual length of breastfeeding [31].

**Breastfeeding knowledge** was assessed using the validated Iowa Infant Feeding Attitude Scale (IIFAS) [32]. IIFAS consists of 17 items and assesses pro-breastfeeding knowledge and attitude specifically in contrast with formula feeding.

C. Procedure

Following administration of informed consent and demographic and baseline measures, participants had a 20-minute interaction with Tanya, followed by post-intervention measures and a semi-structured interview.

During the semi-structured interview, participants discussed their experiences with and attitudes about the Tanya system as well as how the system affected their perceptions of breastfeeding and breastfeeding intent. Interviews consisted of a series of open-ended questions about the users’ overall impressions of both the session and the Tanya system. Users were asked to discuss their session with Tanya and compare it to a similar interaction with a doctor, friend, or family member.

Next, we asked additional questions focused on specific aspects of the session, including what users liked and disliked. Questions elicited user perceptions of the credibility and trustworthiness of Tanya, suggestions to improve the usability of the system, as well as future intentions to use the system. Interviews were audio recorded and transcribed verbatim. Coding and analysis followed an iterative process where new codes were added and refined.

Common patterns that emerged during data analysis were discussed and grouped into themes. A code book of 42 codes was developed and interviews were analyzed using NVivo 12 software.

IV. RESULTS

47 women, aged 18-35, participated in the study.
Table 1: Outcome measures

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>Robot (N=29)</th>
<th>Agent (N=17)</th>
<th>Total (N=46)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Touch (N=10)</td>
<td>Constrained Speech (N=7)</td>
<td>Unconstrained Speech (N=12)</td>
</tr>
<tr>
<td>Breastfeeding Knowledge Score</td>
<td>7.6(3.5)</td>
<td>11.7(3.3)</td>
<td>12.5(7.3)</td>
</tr>
<tr>
<td>Change: post - pre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding Intent Change:</td>
<td>1.6(1.7)</td>
<td>2(1.7)</td>
<td>.58(1.0)</td>
</tr>
<tr>
<td>Month post - pre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (composite scale</td>
<td>4.7(1.0)</td>
<td>4.4(1.0)</td>
<td>5.3(1.1)</td>
</tr>
<tr>
<td>– 7 items)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Alliance (composite</td>
<td>5.1(0.9)</td>
<td>4.3(1.5)</td>
<td>5.3(9.3)</td>
</tr>
<tr>
<td>scale – 12 items)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A. Quantitative Results:

Overall, participants across the 5 conditions were highly receptive of Tanya and considered the CR a useful medium to discuss breastfeeding. Participants rated their comfort discussing breastfeeding with Tanya significantly higher than neutral (Mean(SD)=5.9(1.07),Z=5.7,p<.001). A paired sample t-test also showed a significant pre-post increase in breastfeeding knowledge (t(45)=-14.41,p<.001) and breastfeeding intent (t(45)=-6.39,p<.001) across all conditions (Figure 3).

A one-way ANOVA across all 5 modality variants indicated that there were no significant differences among them on any measure, nor were there any significant differences between robot vs. agent embodiment on any measure. However, a one-way ANOVA across the three user input modalities yielded a significant difference on satisfaction, F(2,43) = 3.57, p<.05, and WAI, F(2,43) = 3.8, p<.05). Post-hoc tests indicated that Constrained Speech resulted in significantly lower scores in comparison to the other two conditions. However, no significant differences were observed between Unconstrained Speech and Touch input modalities (Figure 4).

B. Qualitative Results:

Credible, Trustworthy, and Knowledgeable.

Most participants in all conditions accepted Tanya as a credible and trustworthy resource and were willing to learn about breastfeeding from Tanya. Participants reported overall that the system was informative and knowledgeable. These statements were unprompted and comprised of detailed narratives clearly demonstrating newly acquired knowledge.

[P4] “Because someday when we have kids it’s really important. You know just because I am working if I choose formula feeding and it would lead to childhood obesity. And other disadvantages. And breastfeeding is very good. And I did not know that. I thought it would stress the mother out I

![Figure 3](image-url)  

**Figure 3**: a) Breastfeeding knowledge change across all conditions (post - pre); b) Breastfeeding intent change across all conditions (post - pre)
did not know that the mother would feel good.”

Participants stated that prior to interacting with Tanya they had not thought about breastfeeding. Overall, most participants stated that information gained during their session with Tanya will influence their future healthcare decisions about breastfeeding.

[P35] “I didn’t have much of an idea about it. Like the position. Now I have a good knowledge about breastfeeding. In the future, when I plan to have kids this will definitely come back to me.”

Discussing Sensitive Health Topics with a CR.

Most participants described both the CR and agent as behaving ‘like a person.’ During our analysis we found a pattern that established that this ‘like a person’ status had benefits when discussing sensitive healthcare topics such as breastfeeding.

[P37] “Sometimes when you talk to a doctor, you almost feel like you have to be like, "Yeah, I'll try this," because you're talking to a doctor, but it's like ... then again, it's not a person. It's like there's less of a feeling like you have to take their advice and more just like information. You're not going to hurt her feelings by being like, well, I don't want to do this.”

This sentiment was repeated by most participants who stated that they would feel more comfortable talking to Tanya about breastfeeding than a physician. A few respondents went so far as to say that sometimes they do not ask doctors questions because they are too shy or embarrassed.

[P03] “I think you open up more to someone you know is not going to judge you. A machine would be a better option for you rather than a person. Because you can be yourself and say whatever you want to. And that’s why. Especially health related issues people are always shy about that. They don’t want to open up to someone else. So, I felt like that’s a good option for women.”

The Tanya system was viewed by our participants as ‘human-like’ but not necessarily human. This category allowed participants to quickly understand how to interact with the system, as well as providing them with a comfortable, safe, and judgment-free way to obtain information about a sensitive health topic.

[P27] “Yeah, I really would actually use it again. There are certain sensitive topics that are definitely easier to have with Tanya then a medical professional. You don’t have to deal with the awkwardness of talking to a human.”

[P09] “Interacting with robot is more comfortable. You can uh, go back to the questions you can easily interact with the robot. Than with a human person. Because there would be awkwardness to ask some questions. So, I think robot is appropriate.”

Additionally, many participants mentioned that a healthcare provider is often unwilling or unable to spend significant time during a routine medical care appointment to speak with participants about sensitive healthcare topics. Several participants mentioned that the limited time constraint associated with medical care appointments also impacts a physician’s ability to share information with a patient. As a result, participants described the interaction with the system as direct, informative, and complete when compared to an appointment with a physician.

[P04] “I was perfectly fine with it. It was backed with data. Even if I went to a doctor or such they would say the same thing. And they probably wouldn’t be like look I have charts to prove this. The human doctor might skip something rather than just go over every point. They may not back it up with factual information. A doctor would give me something very brief. So, I get a lot more out of this.”

The comfort and acceptance of receiving information from the Tanya system was echoed by 21 participants during the interviews. Participants endorsed the need for a reliable easy-to-use resource to acquire medically-verified information that they otherwise may not have access to.

V. CONCLUSION AND FUTURE WORK

We found that all modalities resulted in significant increases in user knowledge and intent to adhere to breastfeeding recommendations, however there were few differences among them. Overall, women did report—in both quantitative and qualitative assessments—that they were comfortable obtaining information about breastfeeding from a conversational agent or robot (regardless of embodiment), perhaps more comfortably than obtaining it from a doctor. Women reported that CRs provide an effective, safe, and
non-judgmental medium for obtaining information about breastfeeding.

Our one significant finding on modalities is that users did not like constrained speech input—when they were allowed to speak but could only say one of an allowed set of utterances—whether with a robot or virtual agent. There are several possible explanations for this. The use of speech may raise expectations about the natural language capabilities of the system, which are then violated when users are not allowed to say anything they want (a kind of conversational “uncanny valley” effect [33]). It also could be that this modality took the most time and effort to use, given that users had to read all of the options and then speak their response sequentially at every turn of the dialogue.

Regarding the lack of any significant differences between physical and virtual embodiment, it may be that it doesn’t matter for applications with high cognitive load such as counseling, or for applications that involve sensitive topics, such as breastfeeding, at least when the physicality of a robot is non-essential.

Future work includes replication of the study in other sensitive domains and with more participants to determine the optimal modality for health counseling CRs. We are also interested in exploring the space of conversational agent modality options in AR and VR. Although we attempted to design our virtual agent to look as similar as possible to the humanoid robot, comparisons of embodiments always suffer from many potential confounds of subtle details of the agent design which can only be addressed through replication over many design instances.

REFERENCES


8. Ring, L., Bickmore, T., Pedrelli, P.: An affectively aware virtual therapist for depression counseling. (Year)


