A Robotic Couples Counselor for Promoting Positive Communication

Dina Utami, Timothy W. Bickmore, and Louis J. Kruger, *Non-Member, IEEE*

***Abstract*— Intimate relationships are crucially important in all human societies, yet many relationships are in some degree of distress. Couple psychotherapy has been demonstrated to be effective at reducing relationship distress, yet most couples do not seek help from professionals. Automated couples counselors could provide help to many couples who avoid professional help due to cost, logistics, or discomfort disclosing personal problems. We explore reactions to and acceptance of a humanoid robot that takes the role of a couples counselor in promoting positive communication skills among asymptomatic intimate couples. Couples were comfortable with the robot in this role; displaying intimate behavior during the counseling session. They followed the directions of the robot in practicing interpersonal communication skills, were largely satisfied with the experience, and described several advantages to working with a robot compared to human counselors or self-help materials.**

I. Introduction

Intimate dyadic relationships are the hubs of the social network of our species. On average, throughout the world, more than 85% of people marry by age 50, and in Western countries the majority of those who do not marry enter cohabiting relationships. However, despite high initial hopes and relationship satisfaction, relationship distress is very common, with very high rates of divorce in developed countries (e.g., 43% of all first marriages in the US end in divorce). Rates of relationship distress and dissolution are even higher for cohabiting couples [1].

Relationship distress and divorce can take a staggering toll on our society. Relationship distress and depression are significantly associated with anxiety disorders, drug and alcohol abuse, poor work performance, and overall health and well-being [2]. Distressed couples also have poorer interactions with their children, with more prevalent aggression and hostility and less parental involvement and support, impacting the high rate of depression and suicide among children and young adults [3]. Marital dysfunction has been estimated to cost the US alone billions of dollars annually [4].

In response to these problems, many schools of couple psychotherapy have emerged over the last few decades, and the number of marriage and family therapists have exploded, with one estimate that there are over 1.8 million people in the US currently undergoing relationship counseling [5]. Many studies and meta-analyses have demonstrated that marital therapy is efficacious, resulting in increased relationship satisfaction [6]. However, most people who suffer from relationship distress do not seek help. Some of the most frequently cited barriers include treatment cost, logistical challenges such as lack of time, and a preference to solve problems on one's own. Several relationship researchers have called for the development of nontraditional marital interventions that are more accessible to a greater number of couples [7].

In our work, we are developing an automated, humanoid robotic couples counselor. Although many health counseling agents have been developed and tested in clinical trials using virtual screen-based agent embodiments, we feel that an anthropomorphic robot provides unique affordances in couples counseling. As other research in multi-party interaction has demonstrated, compared to a virtual agent, a robot can more effectively manage turn-taking in a conversation with multiple humans by using its head orientation and gaze as deictic cues to select the recipient of a specific statement or request. Virtual agents suffer from the "Mona Lisa" effect, which limits the communicativeness of their gaze [8]. In addition, robots can be more persuasive than virtual agents, based merely on their physical presence. Bainbridge, et al., compared a video-displayed robot with a physically co-located robot. Their study showed that participants in the co-located condition were more likely to fulfill an unusual instruction from the robot and gave the robot greater personal space, compared to those in the video condition [9]. A physically co-located robot thus may be more appropriate for a task requiring trust and compliance such as ours. Robots may also have some advantages in comparison to human counselors. For example, people are often more comfortable disclosing difficult issues to automated agents or robots compared to other people due to social desirability bias [10].

In this paper we report the results of an initial pilot study to determine whether intimate couples would accept a robot in the role of a couples counselor, whether they feel comfortable engaging in intimate behavior in front of the robot, and whether they follow the directions of the robot when asked to practice positive relationship skills. If successful, this type of robot could be a new kind of "relational agent" that serves to improve the quality of the human-human relationships of its users, and thereby have a positive impact on society.

II. Communication in Intimate Relationships

The ability to communicate effectively is essential for maintaining a happy and satisfying relationship. Research

Dina Utami is with the College of Computer and Information Science, Northeastern University, Boston, MA, USA (e-mail: dinau@ccs.neu.edu).

Timothy W. Bickmore is with the College of Computer and Information Science, Northeastern University, Boston, MA, USA (e-mail: bickmore@ccs.neu.edu).

Louis J. Krugger is with the Department of Applied Psychology, Bouvé College of Health Sciences, Northeastern University, Boston, MA, USA (e-mail: kruger@neu.edu).

has shown that "*communication deficits precede the development of marital distress*" [11] and negative interaction patterns such as criticisms, defensiveness, contempt, and listener withdrawal are predictive of marital instability and divorce [12]. Many individuals turn to self-help books for relationship advice. In 2008, 13.5 million relationship self-help books were sold in the US [13]. However, many self-help books are ineffective because much of the advice they offer is neither grounded in empirical research [14], nor tailored to the needs of each individual couple.

There are several theoretical frameworks for couples counseling. Frameworks such as Emotional Couples Therapy and Behavioral Couple Therapy are intended to remediate specific relationship problems in distressed couples [15]. In contrast, Couples Relationship Education is designed to strengthen non-distressed relationships and be more preventive in nature. This framework uses structured education to teach couples about relationship knowledge, attitudes, and skills. These programs are typically 12-15 hours long and include modeling, rehearsal, and feedback of skills, as well as activities that promote beliefs and attitudes associated with healthy relationships [16].

III. Related Work

A. Health Counseling Agents

Several virtual agents have been developed to deliver health counseling in various areas, such as exercise promotion [17, 18] and medication adherence [19], with generally positive results. In a study evaluating a virtual agent whose role is to explain a hospital discharge instruction, users reported higher satisfaction with the agent compared to a human nurse because "*felt they could take as much time as they needed, and did not feel embarrassed asking the computer agent to repeat itself*" [20]. Several robots have also been successfully used in healthcare: to promote healthy eating habits [21], to help people maintain a healthy weight [22], to increase social behaviors in children with autism [23], and to improve the quality of life of older adults with dementia [24]. All of these agents and robots were designed to interact with a single user at a time.

B. Multiparty Agents

Several researchers have investigated humanoid agents and robots that are able to interact with multiple humans at one time. Their studies have focused on more complex turn-taking strategies [25-27], engaging multiple users [28-30], and speaker and addressee detection [31-34]. Other studies have investigated non-verbal behaviors in human group interactions to estimate group cohesion [35], dominance [36, 37], and agreement/disagreement [38, 39]. In behavior generation, Mutlu, et al. [40] have shown that gaze cues from a humanoid robot can shape participants' conversational roles. Matsuyama et al. [40] developed a framework for facilitation robots that can increase the feelings of group cohesion in users. Finally, Vroon, et al. [41] studied the social positioning patterns of a robot interacting with a human group. Although these studies have demonstrated that agents and robots are capable of appropriately responding to multiple individuals in a single setting, this capability has not to our knowledge been investigated in a health counseling context.

IV. Design

Our robotic couples counselor is designed to interact with a human dyad in a private room in the same way a human couples counselor does (Fig. 1). Users are seated in chairs in front of the counselor so that they can interact both with the counselor and each other. The counseling session is designed to be approximately 30 minutes long, and incorporates standard communication skills training techniques from Couples Relationship Education. We decided to focus on positive communication skills because they are important for all couples, can be simulated without detailed knowledge of a particular couple's problems, and allowed us to minimize risk by working only with asymptomatic couples. We explicitly excluded couples that had any history of domestic violence.

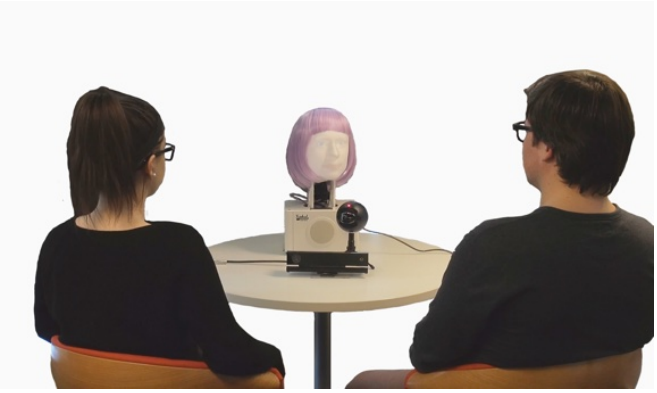

Figure 1. Couple Interacting with Robot Counselor

The robot is a humanoid head developed by Furhat robotics (Fig. 2). It has an animated face, back-projected on a translucent mask, mounted on a two degree-of-freedom mechanical neck that allows it to direct its attention using eye gaze and head pose [42]. The robot's speech is generated using the Windows speech synthesizer and lip movement is synchronized using viseme callbacks from the text to speech engine. Eyebrow raises (for emphasis) and gaze toward/away from users (for turn taking) are generated using BEAT [43]. Dialogue is modeled in hierarchical transition networks using a state chart-based XML formalism. The system was implemented in the open source dialogue system framework IrisTK [44]. In order to assess initial acceptance of the robot counselor without implementing full natural language understanding, user utterances were interpreted by a research assistant in a Wizard-of-Oz framework [45].

The robot maintains the initiative throughout the session, congruent with its role as a high-status expert counselor. Turn-taking is regulated by selecting the next speaker using both robot gaze and speech-based reference to users' given names (programmed into the dialogue system prior to the start of the session). During user speech, the research assistant controls the robot's listening behavior, directing gaze at the speaker and displaying "backchannel" behaviors including head nods and verbal acknowledgments.

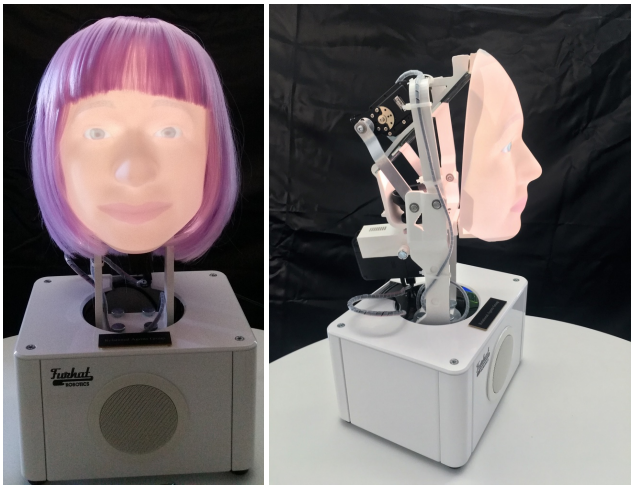

Figure 2. Counselor Robot

Each 30-minute session begins with the robot introducing itself then asking the couple about the history of their relationship. Next, two communication skills are introduced: Active Listening and Effective Speaking. Active Listening is a method to understand a speaker's message accurately and to communicate that understanding to the speaker to demonstrate empathy [46]. The goal of Effective Speaking is to communicate feelings and needs in a clear, non-judgmental manner. It includes several strategies such as using "I statements" (e.g., "I value a clean home" instead of "You are a slob"), and avoiding words that communicate inflexibility such as "always" or "never" [46] . For each skill, the robot first describes the skill, then models the skill, asks the couple to practice the skill with each other, and finally provides feedback to the couple on how well they used the skill. The robot ends the session with a summary and reflection.

V. PILOT EVALUATION STUDY

To assess acceptance of the robotic couples counselor, we conducted a quasi-experimental study with counseling sessions held in a usability laboratory, with measures collected prior to and immediately following each session, as well as 30 days after the session via a web form. Each counseling session and a subsequent semi-structured interview were videotaped. The study was approved by the University IRB and participants were compensated for their time.

A. Study Protocol

Participants were recruited from an online job-posting site and our University web portal, and were required to be at least 18 years old, able to speak and read English, and have been romantically involved with their current partner for at least a year.

Couples were excluded if they indicated any history of domestic violence, using the Conflict Tactic Scale [47]. In addition, the IRB required us to develop a safety protocol to address situations in which couples become violent or abusive during the counseling session.

B. Behavioral Measures

Interactions were videotaped and participant behavior was coded to determine the extent to which participants followed the robot's instructions when requested to practice communication skills. Separate fidelity measures were defined for each of the two skills (Active Listening and Effective Speaking, Table I) to assess how well participants performed the communication skills. Each instance of participant behavior was scored on a 0-10 scale by two judges separately with disagreements resolved by consensus. Adequate inter-rater reliability was demonstrated, with interclass-correlation of 0.9 for Active Listening and 0.8 for Effective Speaking. In addition, a semi-structured interview was conducted with the couple at T1 to probe their reactions to the robotic counselor.

TABLE I. COMMUNICATION SKILL FIDELITY MEASURES

Skills	**Behavior of Interest**	**Operational Definitions**
Active Listening	Attentiveness	Steady gaze, direct body orientation, leaning forward.
	Backchannels	Head nods, vocal assenting behaviors ("*uh-huhs*")
	Indicators of understanding	Any of the following: 1) Paraphrasing (the listener repeated back what the speaker said). 2) Clarifying (the listener asked the speaker questions to ensure that a correct message has been received) 3) Direct expression of understanding (e.g.: "*I agree*!" "*I'm sorry to hear that!*") 4) The listener responds with a statement or a question relevant to the speaker's message.
Effective Speaking	Gentle introduction	The speaker brings up the issue with a greeting or other kinds of opening statement.
	Using "I statement"	The statement focuses on the speaker's feelings or beliefs ("*I feel abandoned and worried when you consistently come home late without calling*") rather than the listener's beliefs or feelings attributed by the speaker ("*Why are you never home on time?*")
	Validation	The speaker communicated sincere understanding and acceptance of the partner's feeling or opinion
	Respect	Gaze/head/body oriented towards the listener, warm facial expression, and gentle tone.

B. Self-Report Measures

In addition to sociodemographic measures, the following measures were collected immediately prior to the session (T0), immediately after the session (T1), or 30 days after the session via web form (T2).

- Attitudes Toward Robotic Facilitator, collected at T1 (Table II). The scale ranges from 1 (not at all) to 7 (very much)
- Enjoyment of the interaction, collected at T1, to assess the enjoyment of the interaction, with a four item composite scale, adapted from [48]. The scale ranges from 1 (not at all) to 7 (a great deal).
- Positive and Negative Affect Scale (PANAS) [49], collected at T0 and T1 to assess changes in emotional state during the session. The scale ranges from 1 (not at all) to 5 (extremely) .
- Inclusion of Other in the Self (IOS) [50], collected at T0 and T1, to assess changes in couples' perceived interpersonal closeness during the session. It is a seven-point graphic (visual analogue) scale.
- Perceived partner's responsiveness to the self, collected at T1, to assess partner's responsiveness during interaction with a four item composite scale ("*My partner seemed to really listen to me.", "My partner seemed interested in what I am thinking and feeling.", "My partner was on 'the same wavelength with me'.", "My partner was responsive to my questions/answers.",* on a scale of 1:Not true at all to 7:Very true).
- Active Emphatic Listening (AELS), measured at T1, to assess perceptions of partner empathy during interaction (adapted from [51]). The scale ranges from 1 (not at all) to 7 (a great deal).
- Relationship Quality (Perceived Relationship Quality Component Inventory, PRQC [52]), Relational Trust (Trust in Close Relationship Scale, TCR [53]), and Communication Quality (Communication Pattern Questionnaire, CPQ-SF [54]) were all measured at T0 and T2 to assess longitudinal effects of the counseling intervention.

TABLE II. ROBOTIC FACILITATOR RATINGS

Ratings of Robotic Facilitator (Anchors 1:Not at all – 7: Very much)	**Mean (SD)**	**p***
How satisfied are you with the facilitator?	5.4 (1.3)	< 0.01
How effective was the facilitator at leading the session?	6.1 (0.8)	< 0.01
How satisfied are you with the interaction experience?	5.7 (1.3)	< 0.01
How much do you like the facilitator?	5.5 (1.6)	< 0.01
How much do you trust the facilitator?	5.1 (2.0)	< 0.01
How natural was the facilitator's behavior?	4.6 (1.7)	< 0.05
How helpful was the facilitator in getting you involved in the interaction?	5.8 (1.1)	< 0.01
How much would you like to continue working with the facilitator?	4.9 (1.0)	< 0.05
How natural was the interaction?	4.9 (1.6)	< 0.01
How interesting was the facilitator	4.8 (1.8)	< 0.01
Would you rather have a live human facilitator lead the interaction?	4.7 (2.0)	< 0.1

*One-sample Wilcoxon signed-rank test demonstrating difference from a neutral score of "4".

VI. RESULTS

We recruited 36 volunteers (18 couples) for the study, of which 2 couples were screened out because of a history of physical abuse. The average age of the 32 participants was 25.1 (range 19 to 76). All had some college education, and 63% were students. All couples were heterosexual and had never participated in a couples counseling session. Of the 16 couples, 7 were seriously dating (do not date other people), 7 were in cohabiting relationship, and 2 were married and were living together. The length of the couples' relationships ranged from 1 to 46 years.

A. Behavioral Results: Intimate Behavior

We observed several examples of intimate behavior during the counseling sessions, indicating that couples were comfortable with the robot and experimental set-up. Non-verbal intimate behaviors included touching, hand-holding, and close proximity during interactions (3). Verbal intimate/affectionate behaviors included intimate self-disclosures (e.g., "*I'm excited to see my mom because I miss her.*"), caring statements (e.g., "*I love you*"), compliments (e.g., "*Hi beautiful*!"), supports statements (e.g., "*Well if you study for the next one you'll bounce back!*"), empathetic statements (e.g., "*I'm sorry to hear that you're stressed …*") and idiomatic terms of endearment (e.g., *sweetheart*, *babe*, *bud*).

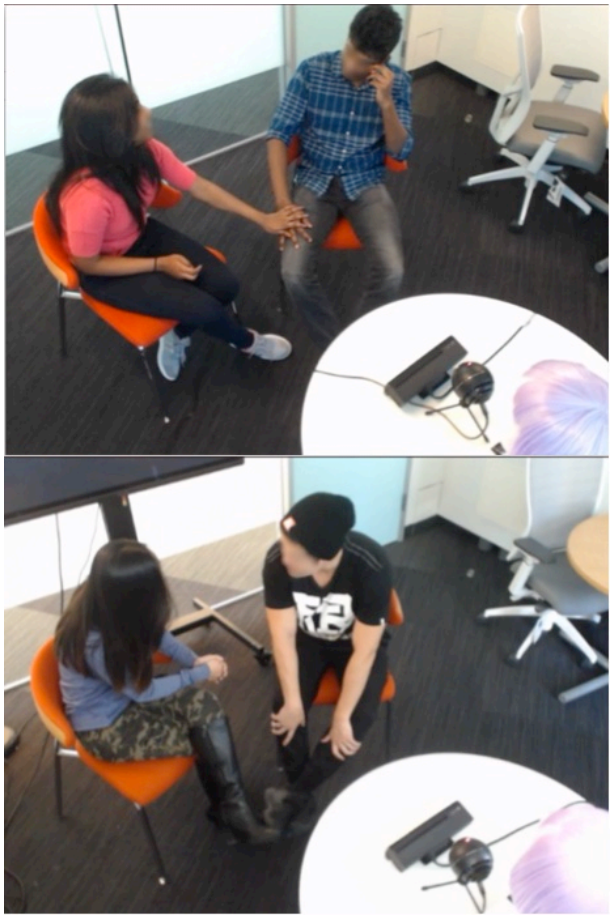

Figure 3. Example Couple Intimate Behavior During Sessions

B. Behavioral Results: Fidelity

Participants actively engaged in practicing communication skills when requested by the robot. There were no instances in which participants refused to comply with the robot's request to engage in practice or role-playing. In addition, participants exhibited good fidelity to the robot's recommendations. For Active Listening, the average fidelity rating (0-10, with 10 perfect) from the two coders was 8.9 (SD = 0.9), and for Effective Speaking the average fidelity rating was 8.8 (SD = 1.0).

C. Self-Report Results

1) *Attitudes Towards Robot:* The ratings for the robot were generally positive across all participants (Table II). No gender effects were found on the ratings of the robot. Participants rated the robot high on satisfaction, effectiveness in leading the discussion, and getting participants involved in the discussion, and they were generally satisfied with the interaction experience. Participants found the robot likeable and trustworthy, and expressed a desire to continue working with the robot. One-sample Wilcoxon signed rank tests demonstrated that most ratings were significantly better than a 'neutral' score of 4.

2) *Enjoyment of the Interaction:* Couples generally enjoyed the interaction with the robot and with each other (M=5.5, SD=1.1, Cronbach's α=0.84), and the response was significantly better than a 'neutral' score of 4 (one-sample t-test, $p < 0.05$).

3) *Positive and Negative Affect:* There was a significant decrease in negative affect after the session (M=12.1, SD=2.9) compared to before (M=14.8, SD = 5.8), paired $t(31)=3.5$, $p<0.05$. There was also trending increase in positive affect after the session (M=32.7, SD=8.6) compared to before (M=31.2, SD=8.9), paired $t(31)=-2$, $p=.056$.

4) Interpersonal Closeness (IOS): We did not find significant differences before (M=5.5, SD=1.2) and after session (M=5.6, SD=1.2) on self-reported interpersonal closeness. A non-parametric two-way analysis of variance yielded a main effect for the participant's gender, $F(1, 4.1)=53.6$, $p<0.05$, with females reporting feeling significantly closer to their partners than males.

5) *Responsiveness of Other:* Participants rated their partner's responsiveness high (M=6.1, SD=1.2, Cronbach's $\alpha = 0.95$), and this was significantly greater than a 'neutral' score of 4 (one-sample t-test, $p <0 .05$). There were no gender effects.

6) *Active-Emphatic Listening:* Couples also rated their partners' active and empathic listening skills generally high (M=5.6, SD=1.1), and this was significantly greater than a 'neutral' score of 4 (one-sample t-test, $p < 0.05$). There were no gender effects.

7) 30-day Follow Up: 21 participants (66%, 14 from the same couple) completed the 30-day (T2) follow up questionnaire on the web. Participants' perception of their partners' relational predictability (TCR subscale) was significantly higher at T2 compared to T0 (36.0 vs. 30.8, paired Wilcox, $p<0.001$). No other significant differences were found relative to baseline (T2-T0) on Relationship Quality, Relational Trust, or Communication Quality.

D. Qualitative Results: Impressions of the Robot Counselor

At the end of each session, we conducted semi-structured interviews to elicit couples experience of the interaction and suggestions for improvement. Our interviews were transcribed and coded using thematic analysis techniques. We started with open coding, followed by clustering of relevant codes into common themes using the affinity diagramming method. Four main themes emerged related to self-disclosure, group set-up, interactivity, and need for tailoring.

1) *Feelings of no-judgment:* Consistent with previous research on increased self-disclosure with virtual humans [Gale,Gratch, 2014], participants expressed their comfort in sharing their thoughts and feelings with the robot, because "*robots don't judge*" [C7,P1]. For example, "*I think you can get people to potentially be more honest with a robot rather than a human because* […] *naturally people felt like humans are judging them especially if you're talking about personal things so they might hold back on what they might say* […]" [C4,P2]. Many participants expressed similar feelings as described in this statement: "*It is sort of embarrassing to talk about these private things with a person*" [C5,P2], in contrast, interacting with robots "*feels more anonymous*" [C15, P1] and private because there were just the couple and the robot in the room. Consequently, this made them feel more "*comfortable to speak up*" [C8,P1] and free to "*being completely our self*" [C13,P2]. Potentially, couple counseling with robots can be a good alternative for those "*who fear going to counselors because they don't want to share something with someone*" [C18,P1].

2) *Benefits of conjoint sessions:* Our results highlight the benefits of counseling sessions held with both members of a couple compared to individual sessions with each partner. As one participant said "*with a relationship, it's not really something that you can work with one person*" [C13,P1], both parties have to work together. One participant gave an example of a situation where it would be helpful to sit together in a counseling session: "*let's say it was a couple who constantly did have some conflict and did have issues, I think that being together and touching on these issues would be helpful*" [C12,P1]. The group set-up enabled couples to exchange non-verbal message instantaneously, for example: "*I liked the fact that we were together because sometimes the robot would say something and then we would both look at each other and we understood what she was saying, we lived a certain situation like that*" [C6,P1]. It also allowed couples to role-play the skills with their partners, instead of just with the robot, which is a more "*meaningful experience*" [C14,P2]. By being there together, they were able to "*build off each other*" [C14,P2]. The group set-up was also thought to be more "*fun*" [C7,P1] and helped users to get comfortable with the robot: "*I think I would have been uncomfortable if it was just me alone with the robot in the room so it was nice to do it in a group*"[C11,P2].

3) *Learning communication skills through practice*: When we asked the participants whether they would prefer

reading self-help books or watching videos to having a couple education session led by a robot for learning communication skills, most participants said they would prefer interacting with the robot. One reason mentioned was because "*it was a two-way communication*" [C7,P2]. They also liked that "*the robot was actually leading the discussion*" and then guiding them through the role play: "*It does force you to do the role playing thing. Like I can imagine reading something and it suggesting 'oh at this point you should role play this with your partner' and me being like yeah okay not really going to do*." [C15,P1]. Interactivity was also cited as being a key characteristic of the interaction. When reading a book and watching a video "*there's the potential to be distracted*" [C14,P1], in contrast, the robot "*guides you into stopping and starting*" [C14,P2] so there's less chance to be distracted. In line with constructivism as a theory of active learning [refs], participants appreciated that they "*had to actually practice*" [C6,P1] during the session. Another participant said, "*I learn more when I do things*" [C6,P2].

4) *Opportunity for tailoring:* While participants like the robot's general teaching strategy and facilitation, the scenarios that we scripted for role-playing practice were "*a little hard to imagine*" [C9,P1]. Participants suggested that we use "*a scenario that happened recently*" [C15,P2]. For example, Couple 15 suggested "*if in the beginning she* [the robot] *said something like what is something you disagree about? ... and then use that in the examples*" [C9,P2]". A couple also suggested that we tailored the content of the counseling according to each couple's need: "*I think our communications are pretty good based on the stuff we learned today so maybe what would be more helpful to us would be a different skill like how to motivate each other*" [C12,P1].

5) *Preference for human facilitators*: Most (53%) participants indicated they would rather have couples counseling with a human facilitator than a robot (last item in Table II), and several comments in the interview supported this. One basis is users' non-acceptance of robots in this role. A 67 years old participant said that she just "*could not take it seriously*" [C10,P1]. Another reason is the feeling that a human counselor would be more empathetic: with a human "*you would sort of bond with them*" [C13,P2]; "*a human would be more sensitive*" [C14,P1]. One participant also said that she/he "*would just trust the human expertise better*"[C15,P1]. Human facilitators would also be able to intervene "*if some confrontation had happened*" [C15,P2]" or "*if a fight erupts*" [C6,P2]. 19% of participants indicated either a human or robot would be acceptable as a couples counselor, depending on the goal of the counseling: "*As education maybe the robot, but as mediation or counseling because of a problem, maybe a human*"[C3,P1].

6) *General impression of the robot:* Several words were used to describe participants' first impression of the robot including: "funny", "excited" and "cool": *"*[...] *I thought it was really cool"* [C13,P2]*;* but also "weird", "surprised", and "anxious": *"I had no idea what to expect when I walked in the room"* [C15,P2]. Over time, however, participants were able to accept it: *"at first it was like: Woah! and then you get used to it*"[C15,P1]. The most frequently mentioned physical aspects of the robot were its hair: *"I liked her pink hair!"* [C6,P1]*; "I would have liked the wig to be a little more normal"* [C9,P1]; and its body: *"We felt there was some sort of presence in the room with us but maybe some body parts would make it feel even more* [...]"[C6,P2]; *"I don't think I would want like a full body* [C11,P1] *I think that might be more intimidating* [C11,P2]". Participants thought the robot did a really good job at facilitating the session and was personalized: *"She was very personalized, she kept saying our names and it was like having a conversation with a person so that was a good thing* [C8,P2].

VII. Discussion

Our study set out to examine the acceptability of a robot designed to facilitate communication skills building sessions for intimate couples. Study participants were very accepting of the robot in this role, expressing high degrees of satisfaction with, liking of, and trust in the robot. They perceived the robotic counselor as being adept at leading the discussion and getting participants involved in the discussion. Participants also found the interaction enjoyable, and their change in PANAS scores demonstrated that the interaction improved their mood. These findings demonstrated that couples reacted positively to the robotic couples counselor.

Perhaps most importantly, participants followed the direction of the robot to practice and role-play positive communication skills with the robot and each other. Based on the post-session interview data, participants expressed that role-playing helped them apply the skills and that they learned more by practicing skills than merely watching videos or reading about them. This finding is consistent with the types of interventions that have produced positive outcomes for couples with human facilitators [55]. Objective ratings from two coders demonstrated that the participants did follow the guidelines of good communications when practicing, and subjective ratings demonstrated that participants perceived their partners to be responsive, and listening actively and empathetically. "*Learning by doing*" is the key idea of experiential learning, a powerful and proven pedagogical approach developed in early 1970s by Kolb [56].

Review of the videotaped sessions and interviews also revealed that participants felt comfortable talking about their relationships and showing affection with each other in front of the robot. Indeed, several of our participants said during their post-session interviews that they would feel more comfortable discussing personal concerns with a robot than with a person. These feelings seem to come from a sense of anonymity; that no person was watching or judging. In the literature, anonymity has been linked to greater disclosure [56]. This result is encouraging given that many couples with problems do not seek therapy because they feel their marital problems are too private to be shared with an outsider [57], implying that for many, a robotic counselor

may actually be preferable to a human in this role.

A. Future Work

Our future work encompasses the development of a fully-automated robotic couples counselor, based on our experience with and data collected from this pilot study. In addition to the challenges of managing such interactions with imperfect speech recognition and natural language understanding, we see turn-taking and automated assessment of communication skill fidelity as particularly interesting research challenges. Extending therapy to multiple sessions in longitudinal treatment is also an important direction of future research. Importantly, adequate strategies for de-escalating conflict during a couples counseling session must also be developed before such systems can be used to intervene on a wider range of relationship problems.

B. Limitations

Our study has many limitations, beyond the small convenience sample of couples used. Our quasi-experimental study was an initial step in testing acceptance and feasibility: true assessment of efficacy must be performed in a randomized, controlled trial. We also note that the majority of participants did state that they would have preferred to work with a human couples counselor rather than the robot. As with most automated counseling systems, we do not see our robot as a replacement for human counselors, but a support tool for those couples who are unable or unwilling to see a human therapist, or even as an adjunct to conventional therapy. We also note that our measure of fidelity may be inflated due to the demands of the research setting. Finally, we recognize that we have only implemented a tiny fraction of what human couples therapists do, especially with symptomatic couples that have specific, ongoing problems. Many problems in natural language understanding, discourse modeling, and the automation of therapeutic protocols must be solved before our robot could be seen as a true therapist, and referring to it at this point as a true couples "counselor" is perhaps a stretch.

ACKNOWLEDGMENT

The authors thank Elise Mason, William Bond, and Arsalan ul Haq for their assistance in conducting the study.

REFERENCES

[1] W. Halford and D. Snyder, "Introduction to Special Series on Universal Processes and Common Factors in Couple Therapy and Relationship Education," *Behavior Therapy,* vol. 43, pp. 1-12, 2012.

[2] M. Whisman, "Marital distress and DSM-IV psychiatric disorders in a population-based national survey," *Journal of Abnormal Psychology,* vol. 116, pp. 638–643, 2007.

[3] A. Krishnakumar and C. Buehler, " Interparental conflict and parenting behaviors: A meta-analytic review," *Family Relations,* vol. 48, pp. 25-44, 2000.

[4] M. Forthoffer, H. Markman, M. Cox, S. Stanley, and R. Kessler, "Associations Between Marital Distress and Work Loss in a National Sample," *J. Marriage and Family,* vol. 58, pp. 597-605, 1996.

[5] American Assoc. for Marriage and Family Therapy. *About Marriage and Family Therapists*. Available: https://www.aamft.org/imis15/AAMFT/Content/About_AAMFT/Qualifications.aspx

[6] W. Shadish and S. Baldwin, "Meta-analysis of MFT interventions," *Journal of Marital and Family Therapy,* vol. 29, pp. 547-570, 2003.

[7] M. Morrill, "The Marriage Checkup: Increasing Access to Marital Health Care," *Family Process,* vol. 50, pp. 471-485, 2011.

[8] S. A. Moubayed, J. Edlund, and J. Beskow, "Taming Mona Lisa: Communicating gaze faithfully in 2D and 3D facial projections," *ACM Trans. Interact. Intell. Syst.,* vol. 1, pp. 1-25, 2012.

[9] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, 2008, pp. 701-706.

[10] G. Lucas, J. Gratch, A. King, and L. Morency, "It's only a computer: Virtual humans increase willingness to disclose," *Computers in Human Behavior,* vol. 37, pp. 94-100, 2014.

[11] H. J. Markman, "Prediction of marital distress: A 5-year follow-up," *Journal of Consulting and Clinical Psychology,* vol. 49, pp. 760-762, 1981.

[12] J. M. Gottman, J. Coan, S. Carrere, and C. Swanson, "Predicting Marital Happiness and Stability from Newlywed Interactions," *Journal of Marriage and Family,* vol. 60, pp. 5-22, 1998.

[13] K. Blakeley. (2006) Self-Help Books: Why Women Can't Stop Reading Them... *Forbes*.

[14] N. Dunbar and G. Abra, "Self-Help Books on Relational Communication: Who's Writing Them and What Advice Are They Giving?," presented at the the annual meeting of the International Communication Association, Dresden, Germany, 2006.

[15] A. S. Gurman and N. S. Jacobson, *Clinical handbook of couple therapy*: Guilford Press, 2002.

[16] W. K. Halford, H. J. Markman, and S. Stanley, "Strengthening couples' relationships with education: Social policy and public health perspectives," *Journal of Family Psychology,* vol. 22, pp. 497-505, 2008.

[17] T. Bickmore, R. Silliman, K. Nelson, D. Cheng, M. Winter, L. Henaulat*, et al.*, "A Randomized Controlled Trial of an Automated Exercise Coach for Older Adults," *Journal of the American Geriatrics Society,* vol. 61, pp. 1676–1683, 2013.

[18] I. de Kok, J. Hough, et. al, "A multimodal system for real-time action instruction in motor skill learning," presented at the ACM on International Conference on Multimodal Interaction (ICMI), 2015.

[19] T. Bickmore, K. Puskar, E. Schlenk, L. Pfeifer, and S. Sereika, "Maintaining Reality: Relational Agents for Antipsychotic Medication Adherence," *Interacting with Computers,* vol. 22, pp. 276-288, 2010.

[20] T. Bickmore, L. Pfeifer, and B. W. Jack, "Taking the Time to Care: Empowering Low Health Literacy Hospital Patients with Virtual Nurse Agents " presented at the Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Boston, MA, 2009.

[21] R. Ros, E. Oleari, C. Pozzi, F. Sacchitelli, D. Baranzini, A. Bagherzadhalimi*, et al.*, "A Motivational Approach to Support Healthy Habits in Long-term Child–Robot Interaction," *International Journal of Social Robotics,* vol. 8, pp. 599-617, 2016.

[22] C. D. Kidd, "Designing Long-Term Human-Robot Interaction and Application to Weight Loss," PhD, Media Arts & Sciences, MIT, Cambridge, MA, 2008.

[23] B. Robins, K. Dautenhahn, R. T. Boekhorst, and A. Billard, "Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills?," *Universal Access in the Information Society,* vol. 4, pp. 105-120, 2005.

[24] W. Moyle*, et al.*, "Exploring the Effect of Companion Robots on Emotional Expression in Older Adults with Dementia: A Pilot Randomized Controlled Trial," *Journal of Gerontological Nursing,* vol. 39, pp. 46-53, 2013.

[25] D. Bohus and E. Horvitz, "Multiparty turn taking in situated dialog: study, lessons, and directions," presented at the Proceedings of the SIGDIAL 2011 Conference, Portland, Oregon, 2011.

[26] D. Bohus and E. Horvitz, "Decisions about turns in multiparty conversation: from perception to action," presented at the Proceedings of the 13th international conference on multimodal interfaces, Alicante, Spain, 2011.

[27] K. Laskowski, J. Edlund, and M. Heldner, "A single-port non-parametric model of turn-taking in multi-party conversation," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 5600-5603.

[28] D. Bohus and E. Horvitz, "Models for multiparty engagement in open-world dialog," presented at the Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue, London, United Kingdom, 2009.

[29] C. Oertel and G. Salvi, "A gaze-based method for relating group involvement to individual engagement in multimodal multiparty dialogue," presented at the Proceedings of the 15th ACM on International conference on multimodal interaction, Sydney, Australia, 2013.

[30] Q. Xu, L. Li, and G. Wang, "Designing engagement-aware agents for multiparty conversations," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 2013.

[31] K. Laskowski, M. Ostendorf, and T. Schultz, "Modeling vocal interaction for text-independent participant characterization in multi-party conversation," presented at the Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, Columbus, Ohio, 2008.

[32] E. Shriberg, A. Stolcke, D. Z. Hakkani-Tür, and L. P. Heck., "Learning When to Listen: Detecting System-Addressed Speech in Human-Human-Computer Dialog," presented at the In INTERSPEECH, Oregon, Portland, USA, 2012.

[33] Y. Takemae, K. Otsuka, and N. Mukawa, "An analysis of speakers' gaze behavior for automatic addressee identification in multiparty conversation and its application to video editing," in *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on*, 2004, pp. 581-586.

[34] M. Katzenmaier, R. Stiefelhagen, and T. Schultz, "Identifying the addressee in human-human-robot interactions based on head pose and speech," presented at the Proceedings of the 6th international conference on Multimodal interfaces, State College, PA, USA, 2004.

[35] H. Hung and D. Gatica-Perez, "Estimating Cohesion in Small Groups Using Audio-Visual Nonverbal Behavior," *IEEE Trans. Multimedia,* vol. 12, pp. 563-575, 2010.

[36] D. B. Jayagopi, H. Hung, C. Yeo, and D. Gatica-Perez, "Modeling Dominance in Group Conversations Using Nonverbal Activity Cues," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 17, pp. 501-513, 2009.

[37] R. Rienks and D. Heylen, "Dominance Detection in Meetings Using Easily Obtainable Features," in *Machine Learning for Multimodal Interaction: Second International Workshop, MLMI 2005, Edinburgh, UK, July 11-13, 2005, Revised Selected Papers*, S. Renals and S. Bengio, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 76-86.

[38] K. Bousmalis, M. Mehu, and M. Pantic, "Towards the automatic detection of spontaneous agreement and disagreement based on nonverbal behaviour: A survey of related cues, databases, and tools," *Image and Vision Computing,* vol. 31, pp. 203-221, 2// 2013.

[39] S. Germesin and T. Wilson, "Agreement detection in multiparty conversation," presented at the Proceedings of the 2009 international conference on Multimodal interfaces, Cambridge, Massachusetts, USA, 2009.

[40] B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita, "Footing in human-robot conversations: how robots might shape participant roles using gaze cues," presented at the Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, La Jolla, California, USA, 2009.

[41] J. Vroon, M. Joosse, M. Lohse, J. Kolkmeier, J. Kim, K. Truong*, et al.*, "Dynamics of social positioning patterns in group-robot interactions," in *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*, 2015, pp. 394-399.

[42] S. Al 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: COST 2102 International Training School, Dresden, Germany, February 21-26, 2011, Revised Selected Papers*, A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann, and V. C. Müller, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 114-130.

[43] J. Cassell, H. H. Vilhjálmsson, and T. Bickmore, "BEAT: the Behavior Expression Animation Toolkit," in *Life-Like Characters: Tools, Affective Functions, and Applications*, H. Prendinger and M. Ishizuka, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 163-185.

[44] G. Skantze and S. A. Moubayed, "IrisTK: a statechart-based toolkit for multi-party face-to-face interaction," presented at the Proceedings of the 14th ACM international conference on Multimodal interaction, Santa Monica, California, USA, 2012.

[45] N. Dahlback, A. Jonsson, and L. Ahrenberg, "Wizard of Oz Studies -- Why and How," in *Readings in Intelligent User Interfaces*, M. T. Maybury and W. Wahlster, Eds., ed San Francisco, CA: Morgan Kaufmann Publishers, Inc., 1998, pp. 610-619.

[46] R. Miller and D. Perlman, *Intimate relationships*: McGraw-Hill Higher Education, 2009.

[47] M. A. STRAUS, S. L. HAMBY, S. BONEY-McCOY, and D. B. SUGARMAN, "The Revised Conflict Tactics Scales (CTS2)," *Journal of Family Issues,* vol. 17, pp. 283-316, 1996.

[48] S. Sprecher, S. Treger, and J. D. Wondra, "Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions," *Journal of Social and Personal Relationships,* vol. 30, pp. 497-514, 2013.

[49] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *Journal of Personality and Social Psychology,* vol. 54, pp. 1063-1070, 1988.

[50] A. Aron, E. N. Aron, and D. Smollan, "Inclusion of Other in the Self Scale and the structure of interpersonal closeness," *Journal of Personality and Social Psychology,* vol. 63, pp. 596-612, 1992.

[51] G. D. Bodie, "The Active-Empathic Listening Scale (AELS): Conceptualization and Evidence of Validity Within the Interpersonal Domain," *Communication Quarterly,* vol. 59, pp. 277-295, 2011/07/01 2011.

[52] G. J. O. Fletcher, J. A. Simpson, and G. Thomas, "The Measurement of Perceived Relationship Quality Components: A Confirmatory Factor Analytic Approach," *Personality and Social Psychology Bulletin,* vol. 26, pp. 340-354, 2000/03/01 2000.

[53] J. K. Rempel, J. G. Holmes, and M. P. Zanna, "Trust in close relationships," *Journal of Personality and Social Psychology,* vol. 49, pp. 95-112, 1985.

[54] A. Christensen and C. L. Heavey, "Gender and social structure in the demand/withdraw pattern of marital conflict," *Journal of personality and social psychology,* vol. 59, p. 73, 1990.

[55] A. J. Hawkins, V. L. Blanchard, S. A. Baldwin, and E. B. Fawcett, "Does marriage and relationship education work? A meta-analytic study," *Journal of Consulting and Clinical Psychology,* vol. 76, pp. 723-734, 2015-02-03 2008.

[56] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*: FT press, 2014.

[57] I. H. Wolcott, "Seeking Help for Marital Problems before Separation," *Australian Journal of Sex, Marriage and Family,* vol. 7, pp. 154-164, 1986/08/01 1986.