LISSA- Live Interactive Social Skill Assistance

Mohammad Rafayet Ali*, Dev Crasta† , Li Jin*, Agustin Baretto*, Joshua Pachter*, Ronald D. Rogge†, Mohammed Ehsan Hoque*
*Department of Computer Science, University of Rochester, Rochester, United States
†Clinical and Social Science in Psychology, University of Rochester, United States
Email: (mali7, mehoque)@cs.rochester.edu
Email: (dev.crasta, ronald.rogge)@rochester.edu
Email: (ljin10, abaretto)@ur.rochester.edu
Email: jpachter@u.rochester.edu

***Abstract*— We present LISSA – Live Interactive Social Skill Assistance – a web-based system that helps people practice their conversational skills by having short conversations with a human like virtual agent and receiving real-time feedback on their nonverbal behavior. In this paper, we describe the development of an interface for these features and examine the viability of real time feedback using a Wizard of Oz prototype. We then evaluated our system using a speed-dating study design. We invited 47 undergraduate male students to interact with staff and randomly assigned them to intervention with LISSA or a self-help control group. Results suggested participants who practiced with the LISSA system were rated as significantly better in nodding when compared to a self-help control group and marginally better in eye contact and gesturing. The system usability and surveys showed that participants found the feedback provided by the system useful, unobtrusive, and easy to understand.**

Keywords—Real-time feedback; virtual agent; conversational skill training; speed-dating;

I. Introduction

Let us consider Bob, a college freshman who finds it difficult to engage with his peers in face-to-face conversations. Bob has heard all the “rules” for making a good first impression: maintain eye contact, speak clearly and confidently, smile often, and be animated rather than stiff, but somehow he has always had trouble putting it together. His difficulties worsen especially when he approaches Amy – one of his classmates that he finds cute. His difficulty interacting with others socially has become an obstacle in making friends on campus. As a result, Bob mainly stays in his room chatting with old high school friends online. Bob enjoys these chats and even makes new friends through social media, but he still craves face-to-face interactions. His friends encourage him to just “get out there” and get better through practice, but Bob is afraid of embarrassing himself in front of his peers. Meanwhile, dedicated social skills coaching would be costly and time consuming on a college student budget. The above scenario depicts two key points – (1) People often desire help improving their social skills, (2) Limitations, such as a lack of resources and social stigma, may impede their ability to obtain that help. In this paper we present LISSA - Live Interactive Social Skill Assistance – an online system driven by Wizard of Oz. This system would allow people like Bob to practice their social skills by having a “getting-acquainted” conversation with a human-like virtual agent. We designed an interface that provides real-time feedback on eye contact, smiling, volume, and body movement without distracting users (Fig. 1).

Fig. 1. LISSA runs on web-browser and driven by Wizard of Oz. After having a conversation with real-time feedback it gives the summery post feedback

In addition the interface provides summary feedback to help participants track their progress across sessions. We validated our system with 47 male participants with a speed-dating study using a randomized design. Participants who practiced with the LISSA system were rated as having significantly improved their nodding and marginally improving their eye contact and gesturing when compared to a self-help control group.

In particular, in this paper, we tried to address the following research questions.

1. How might a virtual agent be used to practice conversation skills?
2. How would a system provide real-time feedback in a subtle way that is useful and reflective?
3. Will such feedback elicit measurable improvements in conversational skills when practiced with other participants?

II. Background

Designing and developing an effective social skills coaching framework required integrating several disparate bodies of knowledge including research on social skills training, real-time feedback, information visualization, and intelligent virtual agents.

Since the 1970s, a wide variety of approaches have been developed to address social skills deficits in the treatment of autism [1], schizophrenia [2], and social anxiety [3]. While these approaches vary in their particulars, common elements include: alerting and discouraging inappropriate behaviors, immediate and continuous reinforcement of correct skill use, and opportunities to practice skills with a coach or therapist. Thus, an effective social skills training system would need to include all three of these elements. Although the most robust research and development of such training has been in the clinical arena, desires to improve social skills extend to a much larger population. Thus, making such a system available online would allow it to be much more accessible.

Many researchers have worked on minimizing the distraction of having real-time feedback during social interactions [4], [5]. Some past strategies include providing chat based visual feedback on language use such as proportion of agreement words and overall word count to improve group collaboration [6]. Real-time feedback was also used for training classroom teachers using bug-in-ear technology to increase teachers' rate of praise statements and their use of proven effective instructional purposes [7]. These findings motivated us to explore the possibility of using real-time feedback to better navigate a social face-to-face scenario.

Previous efforts have also used virtual agents to model social skills or provide real-time feedback. For example, in My Automated Conversation CoacH (MACH) project [8], a virtual agent mirrored smiling and head nods to give real-time feedback. The TARDIS framework [9] also worked with virtual agents to provide realistic socio-emotional interaction in the context of simulated job interviews. Cicero [10] explored the possibility of using interactive virtual audiences for public speaking training. In our work, we explored an interface using an expressive virtual agent with both real-time and summarized feedback to allow users to practice a common social scenario, namely, getting acquainted with new people.

III. SYSTEM

Designing a system that allows people to have a conversation with a human like virtual agent and get both real-time feedback and meaningful summaries introduces a host of challenges. The first challenge was providing real-time feedback to the user that would not distract them while having a conversation. A second challenge was summarizing the feedback and presenting it in a meaningful format. The third challenge was conducting a conversation using a virtual agent that would last a few minutes without making the user uncomfortable. The fourth challenge underlying our attempts was to deliver these components in a system that could function over a web browser. We have addressed all those challenges while designing our working prototype. As our primary focus was to evaluate the feasibility and effectiveness of such real-time feedback in improving users' conversational skills, we opted for a Wizard-of-Oz design where human operators controlled the behavior of the virtual agent as well as the real-time feedback. This allowed us to focus on key design aspects of the interface (e.g., feedback, virtual agent) along with an extensive evaluation of the system. The audio visual data from this prototype, will help us automate the real-time feedback and the dialogue framework in near future.

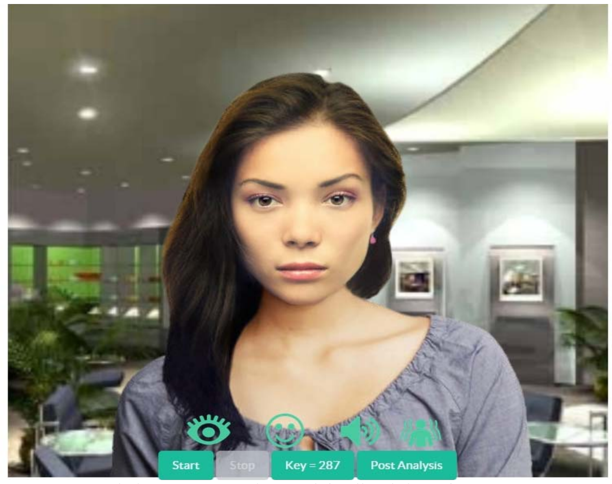

Fig. 2. LISSA real-time conversation and feedback interface.

A. *Interface*

We developed our system using a platform which is accessible via web-browser. It allowed human wizards to control the dialogue framework and the feedback system using a separate machine using TCP/IP protocol. Another advantage is that, once fully automated, our system would be ubiquitously accessible. Our system provides feedback in two formats – real-time feedback and post-session summaries. To start the conversation the user visits the LISSA webpage (available at http://tinyurl.com/lissadate) and presses the Start button in order to activate both the conversation and feedback interface. At the end of the conversation, the user can press the Post Analysis button to view their feedback summaries.

1) Virtual Agent: People may prefer speaking with a real person to paracticing social skills in front of a mirror. That's why our system has users speak with a human like virtual agent. We used a female avatar from "Site Pal"[11]. We chose a look and background environment that gave the appearance of a casual conversation with a female college student (Fig. 2).

2) Real-time Feedback: Real-time feedback is often distracting and can put cognitive load on the user, particularly when it is ambiguous [4], [6]. Among many possible nonverbal behaiors [21] we restricted ourselves to four signals with two colors in order to reduce cognitive load. While future versions of the system could include many possible behaviors for the user to select, we prioritized behaviors that were important based on prior literature in this feasibility study. We placed four simple icons at the bottom of the interface representing eye contact, smile, volume, and body movement (see Fig. 2). The concreteness of these behaviors (allowing for binary feedback as correct or incorrect) further allowed the wizards to give feedback in a consistent way. At the beginning of each session, all icons are green. During the conversation the icons turn red, prompting the user to adjust the corresponding nonverbal behavior. When the eye icon turns red it means that the user is not making enough eye contact. A red smile icon means the user should try to smile more. The volume icon turns red when the user speaks either too loudly or softly. If the body movement icon turns red it means user

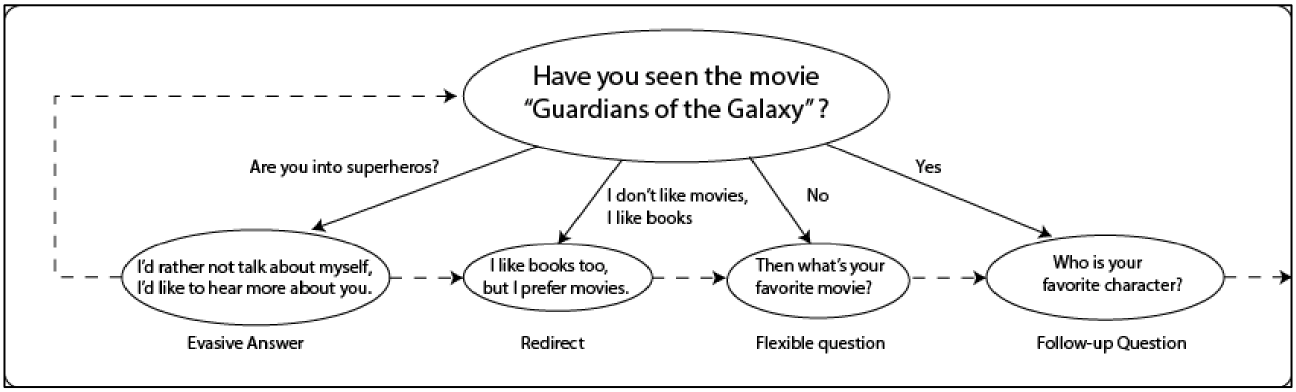

Fig. 5. Keeping conversations on track. Multiple types of responses are preloaded to help lead users to the desired follow-up questions.

need to adjust his body movements (gesture and head nods). Icons will revert to green after the user makes the appropriate adjustment. This simple two color system is as unambiguous as a traffic light, eminding users to pause and adjust before continuing. As an added reinforcer, the white background will gradually turn greener as the user maintains a perfect performance (i.e., all icons green). If a single icon turns red, the background becomes white again.

3) Post Feedback Interface: Providing summary feedback after a conversation has been demonstrated to be helpful in past systems [8]. Therefore, users could access summaries of their real-time feedback after the end of a conversation by pressing the Post Analysis button. This takes users to a different interface where they can see three charts summarizing the session (see Fig. 3). The first chart, "Reminders," represents the amount of time for which particular icon was red. "Best streaks" presents the longest time users kept each icon green. Conversely, Response Lag represents the longest time it took to fix a particular behavior after an icon turned red. Users can pull up previous feedback and see improvements. The charts shown in the post feedback interface are easy to understand and allow users to make quick comparisons across skills. We put supportive text on the charts to help users interpret the information.

B. Wizard of Oz

As our primary goal was to determine the effectiveness of real-time feedback in a conversational scenario, we used a Wizard-of-Oz design to control the real-time feedback and the dialogue framework. This allowed us to see the effect of our system on the users and highlight important points for refinement without automating all components. Secondly, running the study using this design allowed us to collect data about dialogue and feedback essential to automating our system in the future. Our system used two wizards, the Operator and the Coach, responsible for dialogue and feedback respectively. Users were kept unaware of the wizards. However, the wizards could see and hear the user in real-time through a camera to perform their roles responsively.

1) Operator: The Operator managed the virtual agent through a dedicated control panel (shown on Fig 4. And availabale at http://tinyurl.com/lissawizoz). The controller has four types of input. Going clockwise from the top left: The buttons on the top left control the virtual agent's expressions. The buttons on the right side of the panel direct the core script of the avatar, grouped by topic. The buttons on the bottom left are "Quick Replies," frequent responses applicable across topics. Finally, a chat box, lets the Operator enter free text.

Conducting an open ended, meaningful conversation with a user remains an open problem in computer science as it requires true language understanding. To circumvent this problem, we opted to use a scripted strategy. The conversation scripts were designed to lead the user with questions open-ended enough to elicit long responses where users can practice skills, but specific enough to keep users on track. The system currently has two small-talk topics (Living in the city and Movies/TV) and two deeper topics (Dream Home and Career plan). Keeping the conversation on the predefined topics served two functions – reducing the Operator's time typing responses (improving user experience) and obtaining meaningful data for automation. However, a user can respond in different ways even to the most leading stems (e.g., "Have you seen 'Guardians of the Galaxy'?"). Therefore, the script included multiple options for keeping users on topic (Fig. 5). Most deviations from the ideal course can be corrected with more flexible versions of the question or simple redirects. But irrelevant responses (e.g., "Are you into superheroes?") may take the conversation far off course, forcing the Operator to type. In that case, quick replies from the evasive answer group can help the Operator regain control of the conversation.

2) Coach: Coaches providing real-time feedback were instructed in basic principles of behavior modification in a 90 minute session with the second author and met once a month over the course of the study to maintain reliability. Given the importance of immediate responsive feedback, the Coach interface was simple (shown on Fig. 6 and available at: http://tinyurl.com/lissaexpert). By turning each toggle button on and off, the Coach activates the corresponding icons on the user's interface.

IV. EVALUATION

To best address, the primary research questions, evaluation centered on the following targets:

1. LISSA's effectiveness, as assessed through multiple staff ratings after a series of speed-dating interactions.

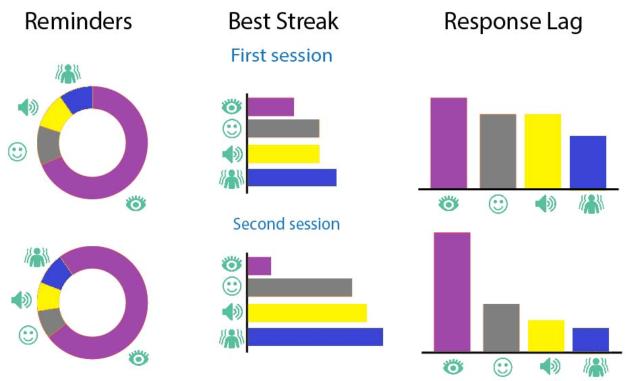

Fig. 3. Sample LISSA feedback summary for 2 sessions. This user has improved both volume and body movement, while neglecting eye contact.

2. Usefulness of our feedback, assessed through user's direct ratings of it as useful and comprehensible.
3. Realism and ease of conversation with the virtual agent, assessed through users' attitudes.

A. *Evaluation Strategy*

To evaluate our system's effect on actual behavior, we conducted a speed-dating study – an approach increasingly used in research on romantic attraction [12] – that seemed well suited to our needs. While other social skills, such as interview skills, can be readily assessed in a highly structured mock interview, how we behave in casual conversation is dependent on a wide variety of factors (e.g., topic choice, partner factors, speaker interest/motivation). Attempts to control these factors through a "standardized" conversation, would be highly artificial and may not generalize to other contexts. By having participants engage in serial 4-min interactions, speed-dating enables measurement of social skills in a scenario with high external validity while removing excess variability when ratings are averaged across all interactions.

We removed additional variability by having participants engage in two rounds of speed-dating with the same partners, yielding both a baseline rating and meaningful estimates of change. Between these two rounds, participants would be assigned to one of two 20-min skills training conditions – a

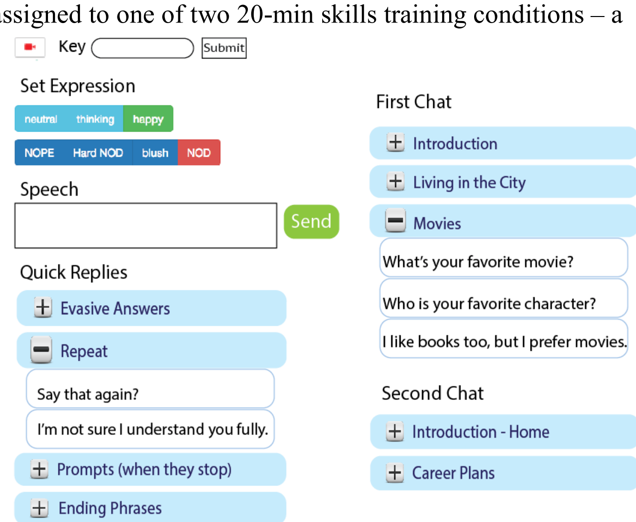

Fig. 4. Operator control panel

Control Condition where they would access readily available self-help methods for social skills, and a LISSA Condition where they engaged in two brief sessions with the system. While all participants would be expected to improve due to practice effects and increased comfort, this design would capture relative increases in performance in the LISSA condition.

For this study, we limited our recruitment to male undergraduate participants. Past research has suggested that men are less skilled at both using and decoding non-verbal behavior [13] and receive less benefit from computerized coaching systems [8]. Thus, limiting recruitment to men represents the most stringent test of LISSA's effectiveness.

B. *Participants*

We enrolled 47 male undergraduate students attending a medium sized research university. All participants were age 18 or older and native English speakers.

Our staff included 8 female research assistants (RAs), also students from that university. There were 3-5 RAs at each session to serve as speed-dating partners and raters.

C. *Setup*

The study was completed through multiple sessions of 3 to 6 participants that were scheduled on weekend afternoons. Before attending their session, participants completed a pre-screening survey, which confirmed their eligibility and included questions about their self-rated perceptions of key skills targeted by LISSA.

We divided each session into 3 parts. In the first round of speed-dating, participants went through a series of interactions with 3-5 RAs. . To reduce participant discomfort and improve reliability, RAs were trained to be encouraging, friendly, and inquisitive across interactions regardless of the participant's performance. Each interaction lasted 4 minutes, after which each male participant rotated to a new table. Before beginning

Fig. 6. Expert's Interface

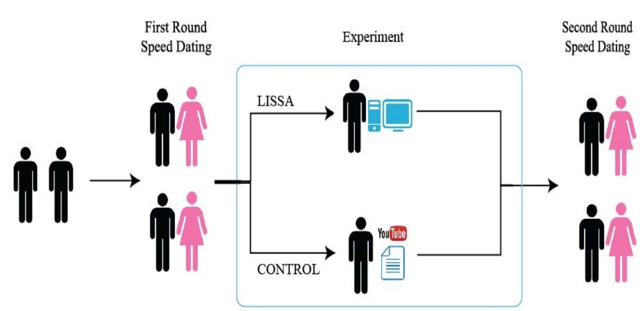

Fig. 7. Study Setup

the next conversation, participants and RAs were both given 1 minute to rate their previous conversation, with RAs rating variables of interest while speed-daters completed a number of distraction questionnaires. This design (4-min interactions and 1-min rating), is sufficient to capture meaningful, robust differences in romantic and nonromantic attraction [5], and even shorter spans of time have been demonstrated to be sufficient for observers to make accurate judgments of the social effectiveness of salespersons [14].

After this initial speed-dating round, participants were taken to another room to be separated into groups based on randomization. Participants assigned to the Control Condition (n = 24) were given time to watch a YouTube video [15] and read free articles from a social skills website [16], as these represented two easily accessible self-help methods for improving social skills targeting the same behaviors as LISSA. Participants assigned to the LISSA Condition (n = 23) had two skills-training sessions (roughly 10 and 7 minutes in length) with LISSA. Each session proceeded with the framework described above (conversation with live feedback followed by an opportunity to review feedback summaries). After this, participants were then able to give feedback about their experiences.

The third section of the study involved a second round of speed-dating in the same format as the first session. To reduce erroneous variance, participants speed-dated with the same RAs they met in the first round and were encouraged to resume their previous discussions. RAs were kept blind to participant condition, and were additionally instructed to avoid asking any questions about participants' skills training.

D. Measures

1) Conversational Skills: We selected items from the Conversational Skills Rating Scale [17] CSRS, for its reliability and because rating instructions and response options were available both for self-report and partner report versions. All items were presented on a six-level Likert scale with higher scores indicating more effective use of that skill. We selected 5 item stems from the CSRS to represent behavioral targets of LISSA, particularly *"Use of eye contact," "Smiling and/or laughing," "Vocal confidence," "Use of gestures to emphasize what is being said,"* and *"Nodding of head in response to partner statements."* These items showed strong internal consistency in RA evaluations ($\alpha = 0.87$).

2) System Evaluation: After completing their two interactions with LISSA, participants in the LISSA Condition completed a questionnaire evaluating the quality of their interaction with LISSA, the utility of the feedback, and their overall impression of the LISSA program using a 5 point scale ranging from -2 (Strongly Disagree) to +2 (Strongly Agree). Participants additionally completed the System Usability Scale (SUS) [18], a commonly used 10 item scale that ranges from 0 to 100 (with higher scores indicating a more usable system).

E. Results

The following sections are sorted by the specific evaluation targets above. Statistical details are provided for

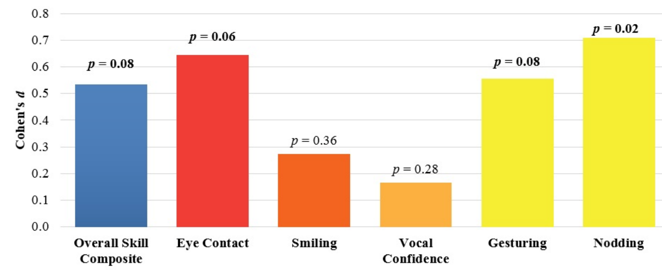

Fig. 8. Effect sizes of condition on post-treatment skills ratings after adjusting for correlation with baseline ratings

significant results. Where possible, participants' open-ended feedback of their experience is also provided to illuminate responses.

1) Effectiveness of LISSA: RA ratings for each individual CSRS item were averaged across all interactions in the first and second round of speed-dating for each participant to create a baseline and post-treatment estimate of that participant's skill level for each item. These 5 item estimates were then averaged together to create an overall skill composite. Encounters were excluded from analysis if RAs reported having met the participants prior to the study or if any possibly biasing events occurred (e.g., conversation went too long; participant revealed condition to the rater; participant behaved inappropriately). A series of t-tests on participant pre-session self-ratings and baseline RA ratings failed to reveal any significant differences between the two conditions, suggesting reasonable baseline equivalence.

As baseline and post-session ratings were strongly correlated ($r[46] = .74$; $p < 0.01$), The effect of the intervention on post-treatment RA skill ratings was assessed using analysis of covariance (ANCOVA), adjusting for baseline ratings of that skill. Using this ANCOVA approach would thus provide an estimate of the impact of LISSA when accounting for initial differences in participants' social skills [19][20]. Analysis of the 5-item composite indicated that RAs assigned individuals in the LISSA condition only marginally higher post-treatment ratings after adjusting for initial ratings ($F[1,44] = 3.28$; $p = 0.08$).

To provide a more nuanced understanding of LISSA's impact, a series of ANCOVAs were also performed on averaged RA ratings for each individual item at post-treatment adjusting for initial averages for that item. To facilitate comparisons for each skill, Fig. 8 shows the adjusted differences between the LISSA and Control condition for each skill as a Cohen's *d*, which scales each difference in terms of the pooled standard deviation for that group. All effects were in the expected direction, with participants in the LISSA Condition demonstrating gains in all targeted skills in comparison to individuals randomized into the Control Condition. In particular, participants in LISSA were rated as significantly higher with respect to nodding in response to partner statements ($F[1,44] = 5.49$; $p = 0.02$). Two other difference emerged on the skill of gesturing to emphasize what is being said ($F[1,44] = 3.34$; $p = 0.08$) and effective use of eye contact ($F[1,44] = 3.89$; $p = 0.06$), which, while

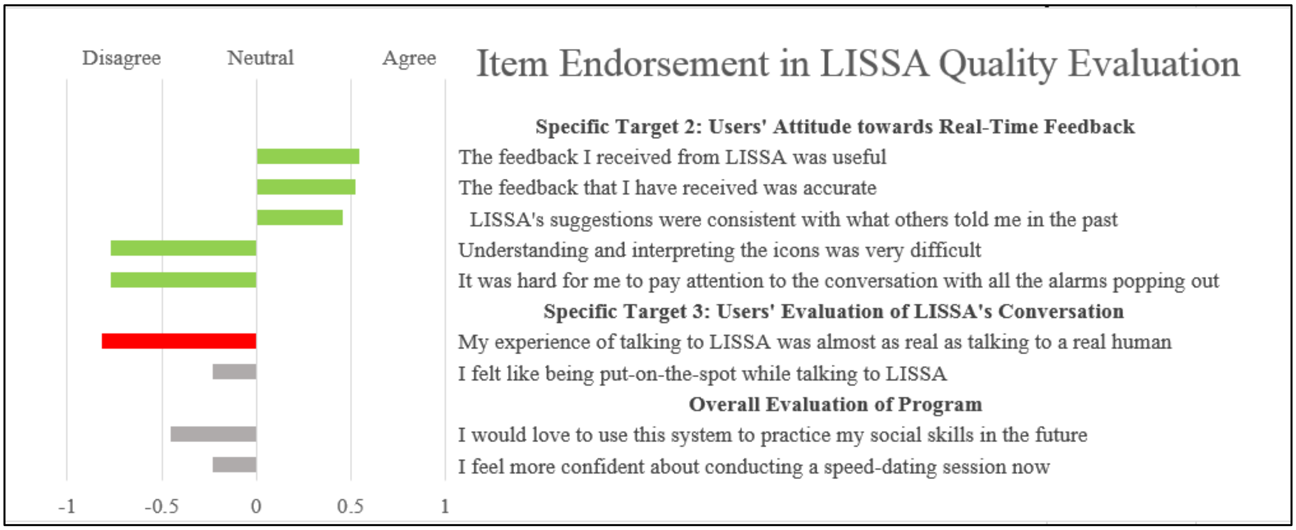

Fig. 9. System evaluation results. Green bars indicate significantly (p < .05) favorable evaluations, red bars indicate significantly unfavorable evaluations, and grey bars indicate that mean evaluations were not significantly different from the neutral option for that item (p ≥ .05). Individual responses ranged from -2 (Strongly Disagree) to +2 (Strongly Agree), figure is expanded to highlight differences between means

marginally signficant, were modest in effect size. Taken together, these effects suggest a promising efficacy of the LISSA system.

2) Evaluation of Live Feedback: As our evaluation questionnaire included separate items each assessing a unique feature of LISSA, they were evaluated separately. Responses to each item were evaluated using a one-sample t-test against a null hypothesis of 0 on the -2 to +2 scale (the "Neutral" option). Summaries for each can be found on Fig. 9.

Users tended to find LISSA's feedback useful ($t[21] = 2.53$; $p = 0.02$), accurate ($t[20] = 2.75$; $p = 0.01$), and consistent with feedback from others ($t[21] = 0.86$; $p = 0.02$). In addition, not only did users tend to deny finding the feedback difficult to interpret ($t[21] = -3.15$; $p < 0.01$) or distracting ($t[21] = 3.73$; $p < 0.01$), many open-ended responses actually highlighted the "immediacy" and "nearly instant" responses of the signals as strong advantages of the system.

3) Evaluation of LISSA's Conversation: Participants tended to find the conversation with LISSA very unrealistic, ($t[21] = -3.81$; $p < 0.01$). Open-ended criticism of the conversation seemed to center on the lack of varied, relevant responses to the unique material brought in by the user:

. *"LISSA could be improved by having more fluid natural responses that are less monotonous."*

"She should try and comment more on the content of what I say, and by that I mean that she should 'say' the same things I just said, as a form of acknowledgement."

"The conversation didn't feel like a conversation. It was far too scripted to feel realistic."

However, it is notable that users did not seem to have any concerns about the format of the conversation (with LISSA mainly asking questions and the respondent mainly answering them) and did not feel "put on the spot" by LISSA's questions. Deepening LISSA's responsiveness while maintaining the general structure of LISSA as an inquirer may greatly improve participants' experiences.

Taken together, the strength of the instant feedback system combined with the rigidity of the conversation script left participants with varied feelings about the overall usefulness and usability of LISSA. Of all the LISSA evaluation questions, participants had the widest variety of responses to the question about their interest in using the system in the future (SD = 1.22). Similarly, the average SUS ratings of participants was 69 (SD = 10), suggesting that LISSA was at an average level of usability. In addition to wide disagreement between participants' evaluations, many participants' ratings seemed to represent mixed feelings about the program's strengths and weaknesses, with one participant writing:

"It was a good way to practice body language. I think if the actual conversational abilities were more realistic then it would be greatly improved."

V. Future Work

These results indicate a promising efficacy for our interface and core concept while also providing directions for refinement. It is especially promising that these gains were observed in such a small sample after only 20 minutes with the program. As a future step, we will fully automate the real-time nonverbal feedback and benchmark the performance using the labeled data that we have collected so far. This would allow users to utilize LISSA for extended times at their own pace, increasing the benefits that can be gained from our feedback system. Even without programming an open-ended conversation, we feel that incremental improvements in our script (e.g., adding some personalized responses) could be a useful intermediate step that can still be automated. Our system would also be measurably improved by development and evaluation with a more diverse population of users (e.g., with respect to gender; age; and relevant conditions such as autism/schizophrenia). Even though our system was evaluated as a package tailored to the context of men engaging casual, non-romantic speed-dating, a similar feedback scheme could be applicable to train wider variety of users (e.g., women; individuals with autism; children) for a wider variety of contexts (e.g., job interviews, training customer service skills, and reducing social anxiety). Adapting LISSA for these contexts would also entail separating the components (e.g., immediate feedback vs. summarized feedback) in order to identify essential mechanisms of effect.

VI. CONCLUSION

We developed and evaluated a system that allows people to practice their social skills by having a conversation with a virtual agent. We designed an interface that provides real-time feedback on non-verbal skills such as eye contact, smiling, volume, and body movement without distracting the users. In addition the interface provides summary feedback to keep track of users' progress across sessions. The live feedback and dialogue framework were controlled using Wizard-of-Oz technique. We validated our system on 47 participants in the context of speed-dating and showed that participants who used our feedback interface were rated significantly higher on use of head nods, and marginally higher in effective gesturing and eye contact when compared to a simpler self-help control. These initial results show early promise in favor of our design helping individuals with real-time feedback. With full automation, our framework could be deployed online making it widely accessible.

VII. REFERENCES

[1] S. Williams White, K. Keonig, and L. Scahill, "Social skills development in children with autism spectrum disorders: A review of the intervention research," *Journal of Autism and Developmental Disorders*, vol. 37, no. 10. pp. 1858–1868, 2007.

[2] T. E. Smith, A. S. Bellack, and R. P. Liberman, "Social skills training for schizophrenia: Review and future directions," *Clin. Psychol. Rev.*, vol. 16, no. 7, pp. 599–617, 1996.

[3] R. G. Heimberg, "Cognitive-behavioral therapy for social anxiety disorder: Current status and future directions," *Biol. Psychiatry*, vol. 51, no. 1, pp. 101–108, 2002.

[4] O. Kulyk, J. Wang, and J. Terken, "Real-time feedback on nonverbal behaviour to enhance social dynamics in small group meetings," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006, vol. 3869 LNCS, pp. 150–161.

[5] T. Kim, A. Chang, L. Holland, and A. S. Pentland, "Meeting mediator: enhancing group collaborationusing sociometric feedback," in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, 2008, pp. 457–466.

[6] G. Leshed, D. Perez, J. T. Hancock, D. Cosley, J. Birnholtz, S. Lee, P. L. McLeod, and G. Gay, "Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups," *Proc. 27th Int. Conf. Hum. factors Comput. Syst. CHI 09*, p. 537, 2009.

[7] M. L. Rock, M. Gregg, B. K. Thead, S. E. Acker, R. A. Gable, and N. P. Zigmond, "Can You Hear Me Now?: Evaluation of an Online Wireless Technology to Provide Real-Time Feedback to Special Education Teachers-In-Training," *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, vol. 32, no. 1. pp. 64–82, 2009.

[8] M. (Ehsan) Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, and R. W. Picard, "Mach: My automated conversation coach," *Proc. 2013 ACM Int. Jt. Conf. Pervasive ubiquitous Comput. - UbiComp '13*, p. 697, 2013.

[9] H. Jones and N. Sabouret, "TARDIS-A simulation platform with an affective virtual recruiter for job interviews," *hazael.jones.free.fr*, 2012.

[10] L. Batrinca, G. Stratou, A. Shapiro, L. P. Morency, and S. Scherer, "Cicero - Towards a multimodal virtual audience platform for public speaking training," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8108 LNAI, pp. 116–128, 2013.

[11] "Site Pal." [Online]. Available: www.sitepal.com.

[12] E. J. Finkel, P. W. Eastwick, and J. Matthews, "Speed-dating as an invaluable tool for studying romantic attraction: A methodological primer," *Pers. Relatsh.*, vol. 14, no. 1, pp. 149–166, 2007.

[13] J. A. Hall, *Nonverbal Sex Differences: Accuracy of Communication and Expressive Style*. 1990.

[14] N. Ambady, M. A. Krabbenhoft, and D. Hogan, "The 30-Sec Sale: Using Thin-Slice Judgments to Evaluate Sales Effectiveness," *Journal of Consumer Psychology*, vol. 16, no. 1. pp. 4–13, 2006.

[15] "5 Body Language Tricks To Make Anyone Instantly Like You - Personality Development & English Lessons." [Online]. Available: http://youtu.be/TD884Dl-kLc.

[16] D. Wendler, *Improve your social skills*. 2014.

[17] A. Instructional, A. Of, and I. Competence, "Conversational Skills Rating Scale," 2007.

[18] J. Brooke, "SUS□: A Retrospective," *J. Usability Stud.*, vol. 8, no. 2, pp. 29–40, 2013.

[19] D. M. Dimitrov and P. D. Rumrill, "Pretest-posttest designs and measurement of change.," *Work*, vol. 20, no. 2, pp. 159–165, 2003.

[20] P. Dugard and J. Todman, "Analysis of Pre□test□Post□test Control Group Designs in Educational Research," *Educational Psychology*, vol. 15, no. 2. pp. 181–198, 1995.

[21] S. Siegman, Aron W.; Feldstein, *Nonverbal Behavior and Communication*, Second Edi. Psychology Press, 2014.