Alexa as Coach: Leveraging Smart Speakers to Build Social Agents that Reduce Public Speaking Anxiety

Jinping Wang
Media Effects Research Laboratory
Penn State University
115 Carnegie Building
University Park, PA, USA
jzw67@psu.edu

Hyun Yang
Media Effects Research Laboratory
Penn State University
115 Carnegie Building
University Park, USA
hzy41@psu.edu

Ruosi Shao
Media Effects Research Laboratory
Penn State University
115 Carnegie Building
University Park, USA
rus494@psu.edu

Saeed Abdullah
College of Information Sciences and Technology
Penn State University
E329 Westgate Building
University Park, USA
saeed@psu.edu

S. Shyam Sundar
Media Effects Research Laboratory
Penn State University
122 Carnegie Building
University Park, USA
sss12@psu.edu

ABSTRACT
Public speaking anxiety is one of the most common social phobias. We explore the feasibility of using a conversational agent to reduce this anxiety. We developed a public-speaking tutor on the Amazon Alexa platform that enables users to engage in cognitive reconstruction exercises. We also investigated how the sociability of the agent might affect its performance as a tutor. A user study of 53 college students with fear of public speaking showed that the interaction with the agent served to assuage pre-speech state anxiety. Agent sociability improved the sense of interpersonal closeness, which was associated with lower pre-speech anxiety. Moreover, sociability of the agent increased participants' satisfaction and their willingness to continue engagement. Our findings, thus, have implications not only for addressing public speaking anxiety in a scalable way but also for the design of future conversational agents using smart speaker platforms.

Author Keywords
Conversational agent; public speaking anxiety; sociability; experiment

CCS Concepts
•**Human-centered computing** → **Human computer interaction(HCI);** *Usability testing;*

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376561

INTRODUCTION
Good public speaking skills are essential for a person's relationship development, educational achievement, and career success. Yet, public speaking anxiety poses great challenges to developing these skills. It is estimated that approximately 15% to 30% of the general population suffer from public speaking anxiety [44], which is the most common type of social phobia. Excessive public speaking anxiety can lead to enormous stress and frustration, impaired speech performance, and further avoidance of social situations that require making public presentations. Individuals who struggle with this anxiety may be well aware of their irrational amount of fear, but it is usually difficult for them to control their responses without appropriate practice and training.

The increasing prevalence of conversational agents (CA), such as smart speakers (e.g., Amazon Alexa, Google Home) and personal assistants (e.g., Apple Siri, Microsoft Cortana), offers the potential to provide such training at scale for users with public speaking anxiety. In fact, a wide range of research efforts has explored the feasibility of using CAs to provide training sessions on social skills [43], as well as treatment and therapy to people who suffer from mental disorders including autism spectrum disorders [43], depression and anxiety [14], and Post-Traumatic Stress Disorder (PTSD) [35]. Moreover, CAs are perceived to be especially advantageous in situations where people feel uncomfortable to open up to a human trainer or therapist [46], which may be the case for people who suffer from public speaking anxiety, because they feel embarrassed talking to others about it [36]. Therefore, given the relatively low access barrier, high ease of use, and growing ability to mimic human communication, conversational agents can be particularly useful as coaches to their human users, by deliv-

ering practice and training for overcoming public speaking anxiety.

In this study, by leveraging the Amazon Alexa platform, we designed and implemented an Alexa-guided tutoring session aimed at reducing users' public speaking anxiety. The major component of the session is a *cognitive reconstructing exercise* [16] that helps participants understand their fear of public speech and address their irrational negative cognition about public speaking. In order to evaluate the effectiveness of the tutoring, we also incorporated a public speech exercise in a Virtual Reality (VR) setting to measure state speaking anxiety. Furthermore, given the competing theoretical propositions about the advantages/disadvantages of the sociability of the CA, we compared the effectiveness of two versions of Alexa, a very social one vs. a less social one, and investigated the underlying theoretical mechanisms. We collected both quantitative and qualitative data to assess users' experience interacting with Alexa. To our knowledge, this is the first study that uses a smart speaker to engage users in cognitive reconstructing exercise for reducing public speaking anxiety.

RELATED WORK

Public Speaking Anxiety

Conceptually, there are two types of public speaking anxiety, trait anxiety and state anxiety. The former refers to a general tendency to be tense when giving a public presentation, which is an enduring personal characteristic. On the other hand, state anxiety represents a transitory anxious state triggered by a specific stimulus, i.e. public speaking, at a given time and location [42]. And there is a significant positive association between these two types of anxiety [33]. In other words, individuals who are high in trait public speaking anxiety are more likely to be physiologically aroused (e.g.,sweating, racing heartbeats) and verbally dysfunctional (e.g., quivering voice) when anticipating or delivering a public speech.

Several cognitive factors offer possible explanations for public speaking anxiety. One of the main factors is more negative self-related thoughts about the anticipated speech [38]. Specifically, common negative cognitions associated with public speaking includes being afraid of making mistakes, worrying about lacking experience, concerns about being rejected or even humiliated by the audience, and anticipated negative results [8]. And for people high in public speaking anxiety, their thoughts about negative self-image, potential public criticism, and failure consequences are even more distorted and irrational [15].

From this point of view, to help individuals alleviate their psychological anxiety during one or more speaking milestones, one of the commonly used method is *cognitive reconstructing (a.k.a., cognitive modification)*, which belongs to the family of cognitive-behavioral interventions [10]. The core idea of this intervention technique is to help individuals correct their negative perceptions, by first identifying negative or irrational self-statements such as "I will sound stupid", then replacing them with more positive and rational coping statements such as "I can only improve"[16, 21]. This approach is widely adopted in public speaking classes [40] and is shown to be an effective approach for reducing public speaking anxiety compared with other techniques, such as exposure therapy, which tries to desensitize individuals from the stimulus through repeated exposure [1, 13, 4]. Therefore, cognitive reconstructing is often a major component of comprehensive cognitive-behavioral intervention programs that treat public speaking anxiety and can be delivered in various forms such as in-person [21] and via the internet [45].

While a fair amount of studies have utilized advanced technologies, in particular VR (see [34] for a review), to deliver the exposure therapy that focuses on the behavioral component of the treatment, addressing the cognitive aspect of public speaking anxiety remains a challenge. To that end, the present study is designed to test whether using conversational agents to deliver cognitive reconstruction training could help reduce the user's public-speaking anxiety effectively.

Conversational Agents and Social Anxiety

Conversational agents have shown promise to play active roles in various mental health services, such as anxiety counseling and therapy. For example, a text-based conversational agent (Woebot) was designed to deliver therapy to students with depression and anxiety symptoms and was revealed to be effective in initial experiments [14]. Another agent (Shim) based on cognitive-behavioral interventions was also found to be useful in promoting psychological well-being as well as sustaining longitudinal engagement with a non-clinical population [31]. For individuals with high social anxiety, specifically, CA could act as a psychological counselor and is known to elicit more intimate self-disclosure from users [27]. Although there are certain drawbacks worth noting, such as the current CA's limited ability to handle serious symptoms and the potential excessive attachment of users, CAs are believed to have the potential to increase the outreach and enjoyment of the therapeutic service [46].

Regarding public speaking, previous research mainly focuses on the usage of the virtual agent to promote users' skills and proficiency in public speaking, focusing less on anxiety reduction. For instance, Hoque et al. developed MATH - a human-like virtual coach providing real-time feedback for improving job interview skills of students [23]. Similar performance-focused CAs have also been employed to help students prepare for presentations [41] and to promote the willingness to communicate among English learners [3]. There is however a paucity of research on using CAs to address anxiety reduction in public speaking settings, even though techniques that focus on changing negative thoughts have been shown to be as helpful as skill training to reduce public speaking anxiety [1]. The cognitive reconstruction technique could also be used in combination with other interventions including skill training and exposure therapy to boost the overall effectiveness of the treatment program [1]. Therefore, using CAs to deliver interventions can potentially fill the gap in public speaking anxiety treatment by focusing on anxiety reduction.

The Role of Sociability

As conversational agents are inherently designed for social interactions, a certain amount of social skills, or sociability,

is required for them to function properly. A highly sociable CA can perform a variety of communication activities, such as small talk [7], turn taking [12], politeness [47], and empathetic expressions [30] in a socially appropriate manner, which helps facilitate and smooth the human-agent interaction.

The question is, when designing the CA system for social anxiety counseling and service, how much sociability should the CA demonstrate to users? From the technology acceptance perspective, the sociability of the social robot is positively associated with acceptance of the system, which is an even stronger predictor than the intelligence of the system [11]. However, in the context of social anxiety disorder treatment, the increased sociability of the CA can be a double-edged sword.

On the one hand, heightened sociability of the CA can create a sense of interpersonal closeness, through which users may feel being concerned and understood by the conversational partner. Several studies have revealed the positive effect of sociability of the agent, such as giving appropriate verbal/nonverbal feedback, on building a sense of rapport and interpersonal closeness in human-agent interaction [50, 18, 25]. This type of closeness can be especially valuable for anxiety-provoking situations. Research has shown that, for individuals who are faced with threatening situations, the mere presence of a partner is not enough. What really reduces stress and provides emotional security is the explicit demonstration of attentiveness and responsiveness by the partner [26]. Therefore, the interpersonal closeness with a highly social CA, compared with a less social CA, seems more likely to comfort users who are faced with the challenge of public speaking and reduce users' anxiety.

On the other hand, sociability is a human-like attribute in essence, which can potentially evoke greater social pressure to perform well. For individuals who are socially anxious, this outcome would be particularly undesirable. For instance, a study in a VR setting found that participants performed a novel task worse when they were co-presented with a human-controlled avatar compared with an automated agent [24]. One of the possible explanations is the evaluative nature of human, i.e., individuals are more likely to feel being judged when they are aware of attention from humans [24]. This kind of concern manifests to a larger extent in more sensitive contexts such as psychological counseling and for more socially anxious people [27]. Based on this line of thought, it is possible that users are more likely to experience fear of being judged in front of a highly social CA compared with a less social one. Moreover, the fear of being judged is often one of the main causes for a person's public speaking anxiety [8].

In terms of user experience, a CA with more social abilities is often preferred across different contexts. As shown in technology acceptance studies, an agent high in sociability could lead to a higher score on enjoyment and intention to use the system [11, 20]. To explain this, the interpersonal closeness felt by users can be the key mechanism, as a sense of closeness and intimacy resulting from a highly social conversation is often crucial to the establishment and maintenance of a long-term relationship [9].

HYPOTHESES

Based on the above literature, we propose a pair of competing hypotheses regarding the relationship between CA's sociability and users' public speaking anxiety (see Figure 1):

H1: Interacting with a highly social conversational agent (compared to a less social CA) will be positively associated with their perceived interpersonal closeness with the agent, which in turn will be negatively associated with public speaking anxiety.

H2: Interacting with a highly social conversational agent (compared to a less social CA) will be positively associated with their fear of being judged by the agent, which in turn will be positively associated with public speaking anxiety.

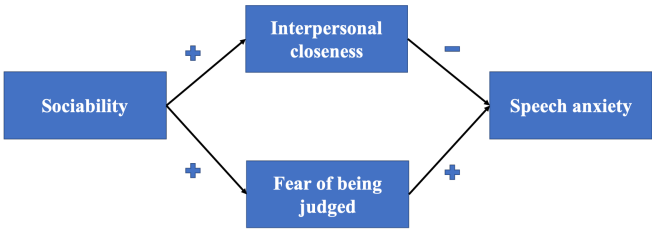

Figure 1. Proposed model (H1 & H2)

In addition, we propose another hypothesis about the effect of sociability on user experience:

H3: Interacting with a highly social conversational agent (compared to a less social CA) will be positively associated with their perceived interpersonal closeness with the agent, which in turn will be positively associated with user experience.

DESIGN OF THE CONVERSATIONAL AGENT TUTOR

We developed the tutor system on the Amazon Alexa. Compared with virtual internet-based CAs that were used in previous studies [23, 41, 3], Alexa has certain advantages. First of all, Alexa as a voice-activated smart speaker can communicate using speech modality more effectively, which allows more natural language input and output than text-based CA systems. In addition, previous internet-delivered interventions have faced with the challenge of low adherence of users [45, 46]. Building skills based on Alexa or other smart-speakers platforms can make the service more accessible and increase the adherence of users, given that the smart speaker can also perform other functions and is better integrated into users' daily lives. Therefore, it shows greater promise to help address public speaking anxiety in a scalable way.

Tutoring Session Design

The tutoring session was built based on structured conversation scripts, which were implemented by manually using the Amazon Alexa's Ask SDK v2 for node.js and then deployed through Amazon Lambda. There are four major parts of the tutoring session, including greetings, a brief introduction about public speaking anxiety, and a cognitive reconstructing exercise, followed by a closing section. During the tutoring session, Alexa would pose a series of questions as cues for

turn-taking, and would continue the conversation after getting the user's input. Below is an example interaction during the introduction session:

Alexa: Many people suffer from speech anxiety. Do you know that?

P: (participants speaking)

Alexa: Well, strong fear can change how you behave in everyday life. So, when it comes to public speaking, you may be disturbed by negative thoughts. For example, "I am going to embarrass myself in front of others". Do you?

P: (participants speaking)

Alexa: The difference between being someone who is fearful, and someone who can communicate, is being aware of your fears. Ok?

P: (participants speaking)

Alexa: Okay! Let's move on to deal with it together. Are you with me?

Cognitive Reconstructing Exercise

Following prior work [16], we adapted the cognitive reconstructing exercise as follows: first, Alexa instructs participants to imagine themselves standing on the stage and being about to deliver a speech, then Alexa asks participants how they would feel and help them identify their negative self-focused statements in that scenario. Second, Alexa teaches participants to substitute negative self-statements with more adaptive coping statements. In total, there are five pairs of negative self-statements – coping statements included in the script (see Table 1), which are known to be common self-statements during public speaking and other sources of anxiety (e.g., fear of being watched/judged, catastrophizing, performance anxiety, etc.) [22]. If participants select one of them, Alexa will continue the conversation based on the selection and teach the corresponding coping statement, which was implemented using branching logic. Below is an example paragraph of the cognitive reconstructing exercise:

P: What I say will probably sound stupid.

Alexa: Okay. For a public speech, there is no absolute right or wrong in terms of what you say. Right?

P: (participants speaking)

Alexa: Your opinions are always valued. And it will be a learning experience when you begin to talk about your opinions, correct?

P: (participants speaking)

Alexa: So please repeat after me, "There's nothing to lose. It's worth a try."

P: (participants speaking)

Alexa: Okay! Sometimes things will get better if you speak up, so try it!

High vs. Low Sociability

To investigate the question of whether a CA with high sociability is more effective in calming users with public speaking anxiety than one with low sociability, we manipulated several aspects of the conversation script to make the sociability of the CA vary across the two conditions. While the basic structure and content of the script remain equivalent, in the high social condition, Alexa is designed to have an additional self-introduction when greeting users [2], show empathy and interpersonal warmth [30], and use conversational fillers (e.g., "um", "well", "let me see") [48]. Table 2 shows examples of these differences in scripts between the two conditions.

METHOD

This study adopts a between-subject lab experiment design with two conditions (a highly social CA vs. a less social CA).

Participants

Given the focus of the study, we wanted to evaluate our system on individuals who have difficulty with public speaking. As such, we reached out to students enrolled in public speaking classes at a large university. A pre-screening question was used to identify participants who have moderate or intense fear in public speaking. Specifically, participants rated their level of agreement on the statement "I have no fear in public speaking." on a 1 (Strongly disagree) to 5 (Strongly agree) scale. Those students who indicated an agreement above 3, suggesting they had no or little fear in public speaking were not recruited.

The final sample (N = 53) consisted of 35 male and 18 female students who participated in the study in exchange for course credits. The average age was 18.23 (SD = 1.41). Most participants identified themselves as Caucasian (90.6%), followed by Asian (9.4%), African American (1.9%), Native American (1.9%), and Middle Eastern (1.9%). They were randomly assigned to either the 'high social CA' (N=26) or 'less social CA' (N=27) condition. No significant differences were found for gender or age between these two conditions.

Procedure

The experiment was conducted in a media research laboratory. When a participant arrived at the lab, s/he was guided to sit in a chair near a smart speaker–Amazon Echo. An experimenter briefly introduced the study and the procedure to the participant. After obtaining the consent from the participant, the experimenter initiated the conversation with Alexa. Depending on the randomly assigned condition (high social vs. less social), the experimenter would use different invocation phrases (high social: "Alexa, open human-computer interaction."; less social: "Alexa, open human-machine interaction.") to invoke the two custom skills. When Alexa started talking, the experimenter left the space and stayed outside of the sight of the participant in order to create a relative private space for the interaction. When participants interacted with Alexa, they knew that the experimenter was out of sight and could not see the interaction.

Following the designed scripts, Alexa delivered the tutoring session to help participants understand and overcome public speaking anxiety. The tutoring session lasted approximately 5-7 minutes. At the end of the interaction, Alexa then asked participants to use what they learned and prepare for a brief

Index	Negative Self-statements	Coping Statements
1	"Everyone is watching me. I never speak well."	"This is awkward, but I can handle it."
2	"People are judging me."	"Forget about judgement from others. I am going to say what I want to say."
3	"What I say will probably sound stupid."	"There's nothing to lose. It's worth a try."
4	"I'm a loser compared with my classmates."	"I can do it well as everyone else does."
5	"They must think I am dumb if I don't do it well."	"Even if things don't go well, it's not that bad at all."

Table 1. Five pairs of negative self-statements and matched coping statements used in the scripts

Elements of the Conversation	Alexa in High Social Condition	Alexa in less social Condition
Greetings	"Okay! I was born in California and I was trained to help people with public speaking. How are you doing today?"	"How are you doing today?"
Empathetic expressions	"For instance, I might think, like "I am going to embarrass myself in front of humans". Do you feel the same?"	"...you may disturb by negative thoughts. For example, "I am going to embarrass myself in front of others". Do you?"
Conversational fillers	"I know it sounds scary. Umm, everyone experienced some level of anxiety and fear in public speaking, right?"	"Everyone experienced some level of anxiety and fear in public speaking, okay?"

Table 2. Examples of the sociability manipulation

speech that lasts about 1-2 minutes. The speech topic (whether it is necessary for college students to purchase textbooks) was kept the same across conditions. This topic was chosen because it could be debatable and was highly relevant for undergraduate students.

When the participant felt ready for the speech, the experimenter asked them to stand up in an empty space in the lab and handed them an Android smartphone wrapped in a Google cardboard through which a virtual reality (VR) app was opened. The VR app, named public speaking simulator and developed by ancientc.com, enabled a viewer to see a 20-member virtual audience in three dimensions. A mild level of background noise was played via phone speakers to increase the realism of the simulated scenario (Figure 2). The virtual audience was pre-programmed and did not react. Prior studies have used similar VR based systems to train and assess public speaking skills [4, 34]. Using a VR device as an evaluation method also enables us to control for potential confounds, such that an actual group of audience might react differently across the duration of the study, which might impact the study participants differently, whereas a pre-programmed VR device provides consistency.

The participant was then asked to handhold the VR cardboard and stand still facing the virtual audience. Before (s)he started the speech, the experimenter would ask the participant to rate his/her distress, fear, anxiety or discomfort on a scale of 0 to 100 at that moment and recorded the number as the measure for pre-speech anxiety. Then the participant started giving the speech for about 1-2 minutes.

Right after participants put down the VR device to indicate their completion of the speech, they were asked the same question about state anxiety again. They were guided to sit

Figure 2. VR speech scenario

down and complete an online questionnaire, in which they answered questions about their perceptions of Alexa, evaluations of their experience, demographics and personal traits. Each experimental session lasted about 30 minutes.

Measures

All measures are on a 5-point scale unless otherwise indicated.

State Public Speaking Anxiety

State anxiety was measured both right before and right after delivering the speech in the VR environment with the Subjective Units of Distress Scale (SUDS) [49]. Participants were asked to rate their feeling of anxiety , distress, fear or discomfort at the moment on a scale of 0-100.

Perceptions about the CA

Sociability of the CA was measured with 12 items adopted from Powers and Kiesler's scale [37], with participants rating to what extent they perceived Alexa being "cheerful," "friendly," "likeable," etc.

Perceived interpersonal closeness was measured with 7 items [28] assessing the level of intimacy of the interaction between participants and Alexa. Sample items include "I felt close to Alexa", "Alexa created a sense of distance" (reverse-coded), and "Alexa was very impersonal in its dealings with me" (reverse-coded).

Fear of being judged was measured using 5 items adopted from previous scale [29] to assess people's fear of "looking foolish", "being criticized", and "feeling disapproval of" Alexa.

User Experience Evaluations

Usefulness was assessed with 4 items asking to what extent the tutoring session provided by Alexa was helpful (e.g., the tutoring from Alexa "could help me be more effective in overcoming my fear of public speech", "makes the public speech easier to get done", and "is useful for public speech practice".

Ease of use was measured with 3 items: "it is easy to use," "it is simple to use," and "it is user friendly."

Fun of using the system was evaluated with 2 items measuring participants' agreement on statements that "it is fun to use" and "it is pleasant to use".

We also measured participants' overall satisfaction with the tutoring session by asking them to indicate their agreement on the item "I am totally satisfied with my interaction with Alexa" and their willingness to continue engagement.

In addition, three open-ended questions were employed to further assess the user experience, including "Do you find Alexa as a public speaking tutor to be particularly useful? How?", "Do you find Alexa as a public speaking tutor NOT to be particularly useful? Why?", and "If you were to design Alexa to be a better public speaking tutor, what would you change or add?".

Control Variables

To control the effect of participants' presence in the VR setting on their state anxiety, we measured their perceived VR presence with one item, i.e., to what extent they feel "being there" in the virtual speech room on a scale of 0-100 (0 = lack of presence, 100 = level of presence in the real world) [5].

Trait anxiety of participants was measured using 6 items adapted from the Personal Report of Communication Apprehension (PRCA) scale [32]. Participants were asked to indicate their agreement on items such as "Certain parts of my body feel very tense and rigid while giving a speech," "I feel relaxed while giving a speech," (reverse-coded) and "I face the prospect of giving a speech with confidence" (reverse-coded).

We also included a question to evaluate participants' previous experience with CAs, in particular smart speakers. They were asked that "in the past several months, approximately how frequently have you used a smart speaker (e.g. Amazon Alexa/Echo, Google Home, Apple Homepod)?", from 1 (Never heard of them) to 7 (Use them all the time).

The study protocol was approved by the Institutional Review Board prior to data collection.

Variable	Mean	Standard deviation	Scale reliability
Pre-speech anxiety	40.33	22.27	NA
Post-speech anxiety	38.77	22.81	NA
Perceived sociability	3.53	0.83	0.89
Fear of being judged	1.37	0.48	0.67
Usefulness	3.37	0.73	0.80
Ease of use	4.26	0.80	0.90
Fun of use	3.78	1.0	0.93
Overall satisfaction	3.55	0.97	NA
Willingness to continue engagement	3.49	1.10	NA
VR presence	43.45	24.3	NA
Trait Anxiety	3.51	0.78	0.81
CA experience	3.71	1.10	NA

Table 3. Descriptive statistics

RESULTS

Table 3 displays the descriptive statistics of the measured variables. As can be seen, all multiple-item scales showed good reliability.

Public Speaking Anxiety

Participants' pre-speech anxiety was relatively low, with approximately 70% percent of participants reported a score lower than 50 on a 1-100 scale. In addition, the anxiety levels of more than 60% participants' anxiety did not change much over the speech– the differences between their pre-post speech anxiety scores were within a range of 10. A paired sample t-test showed that there was no significant difference between reported pre-speech anxiety and post-speech anxiety ($t(50)= 0.69$, $p = 0.49$).

Main Effects of Sociability

Before testing the main research questions and hypotheses, the sociability manipulation was checked for effectiveness: participants in the high social CA condition perceived Alexa as having higher sociability ($M = 3.96$, $SD = 0.50$) than those who were assigned to the less social CA condition ($M = 3.10$, $SD = 0.88$), $t(51)= 4.37$, $p = 0.00$.

An independent sample t-test was employed to compare the overall effectiveness of the two conditions (high social vs. less social) in reducing the public speaking anxiety of participants. The results showed no significant main effects of sociability on participants' pre-speech anxiety (High social: $M = 42.65$, $SD = 23.30$; less social: $M = 38.00$, $SD = 21.39$; $t(50)= 0.75$, $p = 0.46$).

Of note, the absence of main effects does not necessarily rule out indirect effects [19]. As we hypothesized, there may be different indirect paths canceling out the effects of each other. Therefore, it is important to investigate the underlying psychological mechanisms by conducting mediation analyses.

Psychological Mechanisms

To test the mediating effects of perceived interpersonal closeness (H1) and fear of being judged (H2) on the relationship between sociability and state anxiety concerning the upcoming

public speech, mediation analyses with 5000 bootstrapping samples were run with Model 4 of the PROCESS macro for SPSS [19]. Sociability was entered as the independent variable and pre-speech anxiety was the dependent variable, with trait public speaking anxiety, presence in VR, and previous experience using CA being control variables. Interpersonal closeness and fear of being judged were entered as parallel mediators.

The result revealed that interpersonal closeness was a significant mediator between sociability of the CA and pre-speech anxiety. A higher sociability of the CA was positively associated with perceived interpersonal closeness (b= 0.36, p = 0.00), which was negatively associated with the pre-speech anxiety (b= -9.71, p = 0.03]). The overall mediation path was statistically significant (b = -3.52, SE, 95% CI[-9.3144, -0.1258]). Our H1 was thus supported.

The mediation path via fear of being judged was not significant (b = .18, SE = 1.33, 95% CI[-2.5183, 2.9633]). In particular, the increased sociability of the CA did not lead to a higher fear of being judged by the CA (b= 0.01, p = 0.88). Nonetheless, fear of being judged by the CA was a strong positive predictor of pre-speech anxiety (b= 14.087, p = 0.02) even after controlling for the effects of participants' trait speech anxiety, VR presence and previous experience with CA. That is, the more fear participants felt of being judged by Alexa, the more anxious they were before delivering the speech. Therefore, H2 was partially supported (Figure 3).

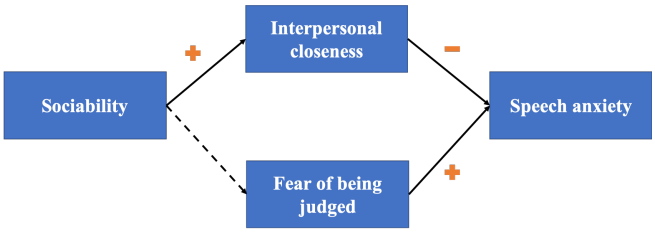

Figure 3. Mediation analyses results (H1 & H2)

User Experience

Overall, participants indicated a positive user experience interacting with Alexa. All the means of the user experience variables (usefulness, ease of use, fun of use, interaction satisfaction, and willingness to continue engagement) were above 3 on a 1-5 scale.

Main Effects of Sociability

A series of independent sample t-tests showed that the mean differences between two conditions regarding satisfaction with the interaction was almost, but not quite, statistically significant (high social condition: M = 3.81, SD = 0.90; less social condition: M = 3.30, SD = 0.99; t(51)= 1.97, p = 0.06). However, participants in the high social CA condition reported a significantly higher willingness to continue engagement (M = 3.85, SD = 0.97) than those who were in the less social condition (M = 3.15, SD = 1.13), t(51)= 2.41, p = 0.02. The ratings of usefulness, ease of use, and fun of use did not significantly differ between the two conditions (Figure 4).

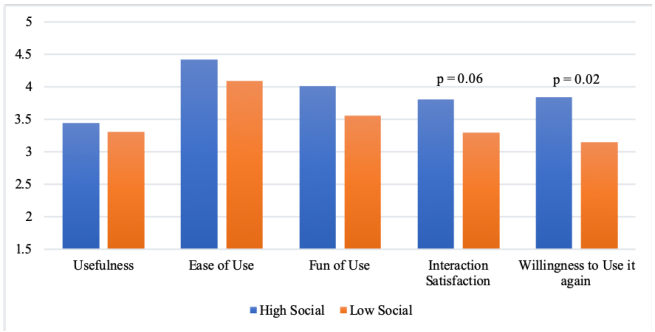

Figure 4. Comparisons of user experience (high social vs. less social)

Psychological Mechanisms

To test H3, another mediation analysis was conducted, with sociability being the independent variable, satisfaction with the interaction and willingness to engage again as the dependent variables, perceived interpersonal closeness being the mediator, and control variables. Interpersonal closeness was a significant mediator for satisfaction (b = 0.18, SE = 0.13, 95% CI[0.0251, 0.5316]) and willingness to continue engagement (b = 0.27, SE = 0.13, 95% CI[0.0698, 0.5572]). H3 was thus supported, indicating interpersonal closeness with Alexa as the underlying psychological mechanism linking sociability with user satisfaction as well as willingness to use the tutoring again.

Qualitative Feedback

Participants answered three open-ended questions focusing on the usefulness, lack thereof and future suggestions for using Alexa to provide tutoring for public speaking. Their feedback was analyzed using thematic analysis techniques, and several themes were identified related to the pros and cons of the sociability of Alexa.

Usefulness of Alexa as a Public Speaking Tutor

The usefulness of the conversational agent as a public speaking tutor was discussed by participants from two major aspects: anxiety reduction and speech preparation.

Alexa was frequently perceived to be useful for anxiety reduction by the participants in the high-social condition. In the less social condition, only one person mentioned that the conversational agent was emotionally supportive, while 11 participants in the high-social condition emphasized that Alexa comforted them, relieving their fear and anxiety about public speaking.

> *"I find Alexa as a public speaking tutor to be useful because it helped me to overcome a fear of anxiety when public speaking."* [P24, from high social condition]

The CA was also found to increase users' confidence, as one of the participants reported,

> *"I do find it to be useful because it...boosted my confidence."* [P25, from high social condition]

In addition to the emotional support, participants appreciated the opportunities to prepare for a public speech with instruction from the CA. This functional usefulness was mentioned

by 14 participants in the high social condition and 8 participants in the less social condition. In particular, participants liked the fact that Alexa helped them organize their thoughts before delivering a public speech.

> *"She is most effective with her ability to organize your thoughts."* [P22, from high social condition]

Although it was not common, the freedom of being judged also emerged as an advantage of Alexa being a speech tutor, especially for those who are assigned to the less social condition. Four participants who interacted with a less sociable Alexa specified that they benefited from the interaction because they had no concern about being judged when interacting with Alexa, compared with a human tutor.

> *"You can practice talking to something that will respond to you, without it being another person. The worry of being judged goes away when you talk to Alexa."* [P38, from less social condition]

Weaknesses of Alexa as a Public Speaking Tutor

Three themes were commonly mentioned by participants when discussing Alexa's weaknesses of being a public speaking tutor.

First, many participants complained about the lack of humanness of Alexa as a public speaking tutor. Participants perceived Alexa as being robotic and machine-like during the interaction.

> *"I do not think Alexa will be a useful public speaking tutor because she is programmed to say what she does and I look at her as a robot."* [P23, from less social condition]

Clearly, the high sociable Alexa did not meet participants' expectations of the CA being anthropomorphized and acting like a real human tutor. In fact, the lack of humanness was the most common shortcoming mentioned in the high social condition, by ten participants.

> *"There is no real human interaction. Alexa can't answer questions that someone who has physically seen the atmosphere of the speech room can."* [P21, from high social condition]

Second, the limited interaction bandwidth of Alexa, i.e., the short and restricted time for users to think and give responses, was a source of complaint. During the tutoring session, when participants were thinking for too long or giving a lengthy response, Alexa could possibly stop working or fail to respond. This caused negative user experience and was mentioned by about 5 participants from both conditions.

> *"She doesn't give me any feedback and she seems rushed with a time limit of 8 seconds to answer the question not giving much time to think."* [P33, from less social condition]

> *"As of now, Alexa simply could not process my responses in a useful manner as it kept shutting off whenever a long response was given."* [P19, from high social condition]

Participants also criticized Alexa for not providing personalized feedback. As participants were guided to share their own experience and feelings towards public speaking during the interaction, they seemed to expect individualized advice from Alexa that was based on their input.

> *"I think that people will think that her responses are not personally based they are general statements she was programmed to say."* [P25, from less social condition]

While the interaction modality of voice made Alexa sound lifelike, it also posed a difficulty in terms of memorizing information. As one participant said,

> *"It was somewhat hard to use and remember everything it said."* [P2, from high social condition]

It is worth noting that in addition to complaints, answers to this question also revealed some encouraging feedback. Eleven participants in the high social condition and 8 participants in the low-social condition expressed their satisfaction towards Alexa providing tutoring, saying that they couldn't find anything they dislike or feel useless.

Suggestions for Alexa as a Public Speaking Tutor

Both groups expected the CA's voice, tone and response to be more like a real human being.

> *"I would make her feel more natural by adding more variation in her tone while she speaks."* [P9, from high social condition]

> *"I would try to make the voice sound a little less robotic."* [P33, from less social condition]

The naturalness of the CA appeared to be lacking not only in the aspect of how it speaks but also in the aspect of what it says. In particular, participants wanted the conversation with Alexa to be more personalized, catering to what they said during the interaction.

> *"add the ability for it to give feedback back to you about a speech and what you are putting in to it.."* [P52, from less social condition]

Or as one participant pointed out, it would be a better experience if we allow users to customize the level of the tutoring, "so that a person could start at a certain comfort zone." [P34, from less social condition]

The need for richness of content from the Alexa was also frequently mentioned by both groups. Specifically, ten participants in the high-social condition and eight participants in the other condition wished the CA to give them more tips, feedback, questions, tasks and practices to prepare for their public speech.

> *"I would maybe add a variation of phrases, and possible other tips. The things she was telling me were often very similar."* [P7, from high social condition]

To address the challenge of memorizing information, three participants suggested having Alexa repeat what they have said as a recap, so that they "could hear it from another point of view" [P43]. And one person mentioned using Amazon Echo Show to add a visual aspect to the experience [P16],

which could be potentially useful for note-taking and providing summaries of the tutoring.

Interestingly, one participant in the high social condition found the use of the conversation filler distracting and intrusive:

> *"Alexa used "um". I did not like this, it feels like she is being forced to have a human like quality and it was very distracting."* [P15, from high social condition]

DISCUSSION

The Effectiveness of Alexa as a Public Speaking Tutor

We tested the efficacy of a conversational agent, Amazon Alexa, for providing tutoring to students with fear in public speaking. Based on the quantitative evaluations, it is notable that while the fear of being judged was shown to be a strong predictor of speech anxiety, the mean value of fear of being judged by the CA was quite low ($M = 1.37$, $SD = 0.48$, on a 1-5 point scale). This suggests that in both conditions, users did not feel they were being judged by Alexa, which could help keep the state anxiety low. Therefore, compared with a human tutor that can easily introduce a sense of being evaluated, a tutoring session by a CA does not seem to induce that apprehension. The qualitative data also echoed these findings: participants recognized the comfort, reassurance, and confidence building they received from Alexa, in addition to being unconcerned about being judged during the interaction.

Sociability and Public Speaking Anxiety

We hypothesized the effect of sociability of the CA on state anxiety via different theoretical mechanisms. Although no significant differences were found between the two experimental conditions (high social vs. less social) in terms of overall state anxiety, our data confirmed that the interpersonal closeness built between the user and the CA is critical. A sociable CA with empathetic expression and conversation filters leads to intimate and personal interactions as perceived by users. As a result, sociability of the CA can help to alleviate stress and anxiety during public speaking. This finding may be expanded beyond the public speaking anxiety context. Specifically, fostering a sense of interpersonal closeness with users might be particularly valuable for CAs in all kinds of therapeutic interventions.

The quantitative data also revealed that higher sociability of the CA did not generate a stronger fear of being judged, which was inconsistent with our hypothesis. It is likely that sociability (high vs. low) did not trigger great differences regarding the general human-likeness of the CA, resulting in similar levels of social pressure. Especially considering the fact that Alexa in these two conditions did not differ in the appearance, voice, or name, the manipulation at the conversation level may not be strong enough to significantly influence users' fear of being judged.

Nonetheless, the sense of freedom from being judged might still be more prominent for participants in the less social condition. Those who interacted with the less sociable Alexa were more likely to mention that their concerns about being judged went away compared with participants in the high-social condition. Like one of the participants in the low-social condition pointed out, "people who have a hard time interacting directly with others" could find it useful [P50]. This might be particularly noticeable for participants who had high social anxiety interacting with humans. Thus, while our quantitative measures did not support this, the qualitative evidence did not invalidate the potential benefits of the low-level sociability. As a result, for CAs designed for social anxiety reduction or therapy, it may be necessary to determine a moderate level of sociability for the CA that can maintain the closeness without intimidating users.

Sociability and User Experience

Both qualitative feedback and quantitative evaluations supported the advantages of having a highly sociable CA for better user experience. It is interesting that although there were no significant differences in the perceived usefulness, ease of use, or fun to use between the two conditions, participants still felt more satisfied and would like to continue engagement with a highly sociable Alexa compared with a less sociable one. The mediation path revealed that the sense of interpersonal closeness was again the key mechanism, such that participants desired to feel being closely connected and enjoyed having an engaged conversation with the CA. This type of demand for high socialness also manifests in the open-ended responses wherein participants criticized Alexa for not being very personal. Our findings thus suggest that one possible strategy for increasing the adherence of the user to the service and leading up to long-term engagement, is fostering a strong sense of interpersonal closeness during the interaction.

Implications

Our study has both theoretical importance and design implications.

The experimental and theoretical confirmation of the role of sociability in the present study contributes to the literature by highlighting the potential benefits and drawbacks of a highly sociable conversational agent. As technology advances, chances are that the social intelligence of the future agent will further increase. The challenge is to devise ways to leverage the sociability to smoothen the interaction and promote positive user experience, while ensuring that it does not introduce extra social pressure. In particular, our findings suggest careful consideration of how social aspects of agents are designed for social anxiety related conversations. Future work should continue to consider different psychological effects generated by the sociability of CA.

Practically, this work developed an interactive tutoring session that can reduce public speaking anxiety and be easily translated to home devices like smart speakers. The service is very accessible to the target population with high public speaking anxiety because it can be used at a private and safe environment for practicing speech skills with a relatively positive and social experience. Given the high prevalence of public speaking anxiety, our work shows the promise to address the issue in a scalable manner and benefit a larger population.

The conversational agent designed in our study also fills the gaps in CA technology that hitherto has focused on exposure therapy or skill training. Instead, our method paves the way for

integrating cognitive reconstruction exercise and more social conversations, which is valued by users for its emotional reassurance and support. Moreover, the interpersonal closeness built between the user and the agent can potentially increase the willingness to continue engagement. Therefore, this tutoring session has the potential to be incorporated in larger interventions or speech curricula and be used in combination with other techniques like VR exposure therapy to better address the anxiety problem. Furthermore, given that cognitive reconstruction exercise has been successfully adopted for other mental health issues [21, 15]. The ability to deliver such exercises using Amazon Alexa can be useful beyond the public speaking context.

In addition, our study confirms several effective approaches to increase the sociability of a voice-based conversational agent. One way to do this is to use warm greetings and self-introduction as the beginning of the interaction, which helps establish a quick rapport between the user and the CA, laying a good foundation for building interpersonal closeness. CAs can also employ empathetic expressions to show emotional responsiveness. When Alexa uses phrases such as "I understand" or "I can relate to you" as responses to user's expressions, users are likely to perceive it being highly social. Another design approach is to build conversational fillers into the CA system to increase its sociability. Based on our findings, the use of words such as "well," "you know," and "like" seems to be acceptable to users, whereas words like "umm" can sound too artificial. Additional design and testing can be conducted to examine the implementation of different conversational fillers.

The feedback gained in this study also highlights several ways in which we can improve the design of future conversational agents. First, participants reported an overall pleasant interactive experience, yet they believed that the system lacked human-like features. Although Alexa has a human voice and is easy to be personified [39], and we tried to make the conversation as natural as possible in the high social condition, the CA was still perceived to be less effective as a human tutor because of its robotic nature. For some participants, the use of certain human-like features (e.g., the word "umm") appeared unnatural and disingenuous. Therefore, one direction of future work can focus on how to better anthropomorphize the CA for effective and natural conversations. For instance, designers can consider enhancing the message interactivity of the human-agent conversation by increasing the degree of contingency in message exchanges (e.g., acknowledge the participant's responses to previous questions, demonstrate knowledge of interaction history, provide a conversation thread by referring to earlier questions and answers). Contingency has been shown to increase perceived humanness of the agent [17, 6]. Relatedly, how to better communicate and manage people's expectations about CA is worth exploring. Second, the rigidity of the response time set by Amazon Alexa Skills frustrated users. For many exercises and skills that require a large amount of user input, participants may need extra time to think, to pause, and then respond, while Alexa is not able to provide such flexibility. Therefore, future design improvement is needed to make the response time of the CA to be more adaptive and variable. Last, similar to a radio broadcast, the information delivered through voice-based CA can hardly be noted down or remembered by participants. This poses great challenges for the application of the CA in education domains. One promising approach would be to provide a textual summary of the training/practice in real-time and delivering the summary via the mobile application afterwards. In this way, even though the interaction with the CA is voice-based, participants also gain an opportunity to review the content and better remember what has been discussed. Additionally, designers can also consider using newer CA devices that incorporate visual modalities (e.g., Amazon's Echo Show) to present the information, which can be particularly useful in the context of skill practicing and learning.

Limitations

Our study has a number of limitations that merit note. First, our evaluation did not include a baseline condition (such as a human tutor) to compare, given that we are exploring the use of smart speakers in this novel context and our goal is to establish the preliminary efficacy of smart speakers. However, this limited the implication of this study. In subsequent studies, we will interview users to understand their expectations from an Alexa tutor and then compare with human tutors of similar levels to extend our findings. Second, we used a one-item scale to evaluate participants' state anxiety. Future research can benefit from measuring physiological data, such as heart rate, to complement self-reported anxiety. Also, the existing tutoring session lacked tailored content and individualized instruction, which weakened the usability and user experience. Future design can adopt techniques that can provide more personalized feedback based on the user's speech performance. Another limitation is that we evaluated the effects from a single laboratory session. It is necessary for future research to examine the effectiveness of the CA tutoring from a longitudinal perspective and in a field setting (e.g., a classroom), which not only can deepen our understanding of different factors that may impact user's adherence to the service, but also explore whether sustained tutoring from a CA can produce long-term and positive effects on individuals' general anxiety tendency in public speaking.

CONCLUSION

This research presents a novel approach to address individuals' public speaking anxiety by using Amazon Alexa to deliver a cognitive reconstructing intervention. We also altered the sociability of Alexa to test its effects on the success of the coaching to reduce anxiety and user experience. Based on the qualitative feedback, it appears that the interaction with the CA served to assuage pre-speech state anxiety. In addition, a highly sociable CA provided better user experience and increased the willingness to continue engagement by establishing a sense of interpersonal closeness with the user. This suggests that using a sociable smart speaker to engage users in cognitive reconstructing could be a viable solution for helping individuals manage their state anxiety, potentially benefiting a large population in the comfort of their homes.

REFERENCES

[1] Mike Allen, John E Hunter, and William A Donohue. 1989. Meta-analysis of self-report data on the

effectiveness of public speaking anxiety treatment techniques. *Communication Education* 38, 1 (1989), 54–76.

[2] Theo Araujo. 2018. Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior* 85 (2018), 183–189.

[3] Emmanuel Ayedoun, Yuki Hayashi, and Kazuhisa Seta. 2015. A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science* 60 (2015), 1433–1442.

[4] Margalit Bar-Zvi. 2011. Virtual reality exposure versus cognitive restructuring for treatment of public speaking anxiety: A pilot study. *The Israel journal of psychiatry and related sciences* 48, 2 (2011), 91.

[5] Woodrow Barfield and Suzanne Weghorst. 1993. The sense of presence within virtual environments: A conceptual framework. *Advances in Human Factors Ergonomics* 19 (1993), 699–699.

[6] Saraswathi Bellur and S Shyam Sundar. 2017. Talking health with a machine: How does message interactivity affect attitudes and cognitions? *Human Communication Research* 43, 1 (2017), 25–53.

[7] Timothy Bickmore and Justine Cassell. 2000. "How about this weather?" Social dialogue with embodied conversational agents. In *Proc. AAAI Fall Symposium on Socially Intelligent Agents*.

[8] Amy M Bippus and John A Daly. 1999. What do people think causes stage fright?: Naïve attributions about the reasons for public speaking anxiety. *Communication Education* 48, 1 (1999), 63–72.

[9] Justine Cassell and Timothy Bickmore. 2003. Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. *User modeling and user-adapted interaction* 13, 1-2 (2003), 89–132.

[10] Pim Cuijpers, Marit Sijbrandij, Sander Koole, Marcus Huibers, Matthias Berking, and Gerhard Andersson. 2014. Psychological treatment of generalized anxiety disorder: a meta-analysis. *Clinical psychology review* 34, 2 (2014), 130–140.

[11] Maartje MA De Graaf and Somaya Ben Allouch. 2013. Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems* 61, 12 (2013), 1476–1486.

[12] David DeVault, Johnathan Mell, and Jonathan Gratch. 2015. Toward natural turn-taking in a virtual human negotiation agent. In *2015 AAAI Spring Symposium Series*.

[13] Patricia Marten DiBartolo, Randy O Frost, Ashley Dixon, and Sarah Almodovar. 2001. Can cognitive restructuring reduce the disruption associated with perfectionistic concerns? *Behavior therapy* 32, 1 (2001), 167–184.

[14] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e19.

[15] Arthur Freeman, James Pretzer, Barbara Fleming, and Karen M Simon. 2004. *Clinical applications of cognitive therapy*. Springer Science & Business Media.

[16] William J Fremouw and Robert E Zitter. 1978. A comparison of skills training and cognitive restructuring-relaxation for the treatment of speech anxiety. *Behavior Therapy* 9, 2 (1978), 248–259.

[17] Eun Go and S Shyam Sundar. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97 (2019), 304–316.

[18] Jonathan Gratch, Anna Okhmatovskaia, Francois Lamothe, Stacy Marsella, Mathieu Morales, Rick J van der Werf, and Louis-Philippe Morency. 2006. Virtual rapport. In *International Workshop on Intelligent Virtual Agents*. Springer, 14–27.

[19] Andrew F Hayes. 2009. Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication monographs* 76, 4 (2009), 408–420.

[20] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2008. The influence of social presence on acceptance of a companion robot by older people. (2008).

[21] Richard G Heimberg, Robert E Becker, Karen Goldfinger, and James A Vermilyea. 1985. Treatment of social phobia by exposure, cognitive restructuring and homework assignments. *Journal of Nervous and Mental Disease* (1985).

[22] Stefan G Hofmann and Patricia Marten DiBartolo. 2000. An instrument to assess self-statements during public speaking: Scale development and preliminary psychometric properties. *Behavior Therapy* 31, 3 (2000), 499–515.

[23] Mohammed (Ehsan) Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W. Picard. 2013. MACH: My Automated Conversation Coach. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 697–706. DOI: **http://dx.doi.org/10.1145/2493432.2493502**

[24] Crystal L Hoyt, Jim Blascovich, and Kimberly R Swinth. 2003. Social inhibition in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 12, 2 (2003), 183–195.

[25] Lixing Huang, Louis-Philippe Morency, and Jonathan Gratch. 2011. Virtual Rapport 2.0. In *International Workshop on Intelligent Virtual Agents*. Springer, 68–79.

[26] Heidi S Kane, Cade McCall, Nancy L Collins, and Jim Blascovich. 2012. Mere presence is not enough: Responsive support in a virtual world. *Journal of experimental social psychology* 48, 1 (2012), 37–44.

[27] Sin-Hwa Kang and Jonathan Gratch. 2012. Socially anxious people reveal more personal information with virtual counselors that talk about themselves using intimate human back stories. *Annual Review of Cybertherapy and Telemedicine* 181 (2012), 202–207.

[28] Nanda Kumar and Izak Benbasat. 2002. Para-social presence and communication capabilities of a web site: a theoretical perspective. *E-Service* 1, 3 (2002), 5–24.

[29] G Frank Lawlis. 1971. Response styles of a patient population on the fear survey schedule. *Behaviour Research and Therapy* 9, 2 (1971), 95–102.

[30] Christine Lisetti, Reza Amini, Ugan Yasavur, and Naphtali Rishe. 2013. I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems (TMIS)* 4, 4 (2013), 19.

[31] Kien Hoa Ly, Ann-Marie Ly, and Gerhard Andersson. 2017. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. *Internet interventions* 10 (2017), 39–46.

[32] James C McCroskey. 1978. Validity of the PRCA as an index of oral communication apprehension. *Communications Monographs* 45, 3 (1978), 192–203.

[33] James C McCroskey and Michael J Beatty. 1984. Communication apprehension and accumulated communication state anxiety experiences: A research note. (1984).

[34] Katharina Meyerbröker and Paul MG Emmelkamp. 2010. Virtual reality exposure therapy in anxiety disorders: a systematic review of process-and-outcome studies. *Depression and anxiety* 27, 10 (2010), 933–944.

[35] Nasim Motalebi and Saeed Abdullah. 2018. Conversational Agents to Provide Couple Therapy for Patients with PTSD. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '18)*. ACM, New York, NY, USA, 347–351. DOI: http://dx.doi.org/10.1145/3240925.3240933

[36] Mark Olfson, Mary Guardino, Elmer Struening, Franklin R Schneier, Fred Hellman, and Donald F Klein. 2000. Barriers to the treatment of social anxiety. *American Journal of Psychiatry* 157, 4 (2000), 521–527.

[37] Aaron Powers and Sara Kiesler. 2006. The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 218–225.

[38] Charles B Pull. 2012. Current status of knowledge on public-speaking anxiety. *Current opinion in psychiatry* 25, 1 (2012), 32–38.

[39] Amanda Purington, Jessie G Taft, Shruti Sannon, Natalya N Bazarova, and Samuel Hardman Taylor. 2017. Alexa is my new BFF: social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2853–2859.

[40] Thomas E Robinson. 1997. Communication apprehension and the basic public speaking course: A national survey of in-class treatment techniques. *Communication Education* 46, 3 (1997), 188–197.

[41] Jan Schneider, Dirk Börner, Peter Van Rosmalen, and Marcus Specht. 2015. Presentation trainer, your public speaking multimodal coach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. acm, 539–546.

[42] C. D. Spielberger. 1972. *Anxiety: Current trends in theory and research*. Academic Press, Chapter Conceptual and methodological issues in anxiety research.

[43] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PloS one* 12, 8 (2017), e0182151.

[44] Vickram Tejwani, Duc Ha, and Carlos Isada. 2016. Public Speaking Anxiety in Graduate Medical Education-A Matter of Interpersonal and Communication Skills? *Journal of graduate medical education* 8, 1 (2016), 111–111.

[45] Maria Tillfors, Per Carlbring, Tomas Furmark, Susanne Lewenhaupt, Maria Spak, Anna Eriksson, Bengt E Westling, and Gerhard Andersson. 2008. Treating university students with social phobia and public speaking fears: internet delivered self-help with or without live group exposure sessions. *Depression and Anxiety* 25, 8 (2008), 708–717.

[46] Aditya Nrusimha Vaidyam, Hannah Wisniewski, John David Halamka, Matcheri S Kashavan, and John Blake Torous. 2019. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry* 64, 7 (2019), 456–464.

[47] Ning Wang, W Lewis Johnson, Richard E Mayer, Paola Rizzo, Erin Shaw, and Heather Collins. 2008. The politeness effect: Pedagogical agents and learning outcomes. *International journal of human-computer studies* 66, 2 (2008), 98–112.

[48] Noel Wigdor, Joachim de Greeff, Rosemarijn Looije, and Mark A Neerincx. 2016. How to improve human-robot interaction with Conversational Fillers. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 219–224.

[49] Joseph Wolpe and Arnold A Lazarus. 1966. Behavior therapy techniques: A guide to the treatment of neuroses. (1966).

[50] Ran Zhao, Oscar J Romero, and Alex Rudnicky. 2018. SOGO: a social intelligent negotiation dialogue system. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. ACM, 239–246.