Towards Empathetically Responsive Voice Assistants

Tanuj Aggarwal
aggarwalt@student.unimelb.edu.au
University of Melbourne
Melbourne, VIC, Australia

Jorge Goncalves
jorge.goncalves@unimelb.edu.au
University of Melbourne
Melbourne, VIC, Australia

ABSTRACT

The emergence of voice assistants has resulted in a more seamless interaction between users and their devices. Existing work highlights how social features like empathy, anthropomorphism, and personification affect the interaction quality and acceptance of the device. In this exploratory study, we investigate the possibility of using voice assistants as empathetic agents that can converse socially with their users, outside their usual function, via Speech Synthesis Markup Language (SSML) features like prosody to express emotions and display empathy. A user study indicated that a combination of verbal content and prosodic features successfully helps the assistant express emotions, allowing users to perceive empathy. Moreover, while users support the emergence of empathetic voice assistants, they are also concerned about possible negative consequences, such as reduced human interaction and exploitation of emotional states. Our work contributes to the growing field of conversational user interfaces, opening doors for a new form of human-machine relationship.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*; *User studies*.

KEYWORDS

Empathetic voice assistants, conversational interaction design

1 INTRODUCTION

Voice Assistants (VA) are slowly becoming integral to many households. According to Fortune Business Insights reports, the market value of VA systems reached USD 6.9 billion in 2018, and it is predicted to go as high as USD 23.8 billion by 2026 [33]. They offer users various new functionalities and convenience, while being easily accessible through personal or home devices such as smart speakers, smartphones, and personal computers [10, 45]. Additionally,

they provide people with new means of interaction with computing systems that do not require physical touch [2].

The recent advent of voice assistants has been pivotal to HCI [17, 33]. However, the interaction is more instructional than social. Chin and Yi [10] state that users often treat VA as servants, resulting in abusive behaviour towards these agents. They also argue that this behaviour may translate into real life through poor communication skills, and even result in the mistreatment of humans and animals. Furthermore, following the Computers as Social Actors (CASA) theory, Pelau et al. [34] suggest that inclusion in users' social group plays a vital role in accepting more AI devices in our daily lives. In many real-world situations, having a more sociable VA can improve the users' experience with the device. In hospital dormitories, for instance, patients admitted in private rooms for longer duration often feel lonely and sad due to lacking social interaction. In such cases, a socially skilled VA can act as a friend and provide companionship to these patients.

Multiple factors affect the sociability of an AI device. Research suggests that anthropomorphic characteristics, such as physical appearance and mental state, along with personifying devices through gender and identity, play an essential role in developing consumer-AI relationships [9]. Furthermore, improving a device's empathy and interaction quality enhances the communication experience between humans and their VA, resulting in users interacting more with humanoid AI devices than with non-humanoid ones [34]. Empirical evidence indicates that an empathetically responsive machine can positively influence the users' emotions and behaviour, improving trust, engagement, and loyalty [10]. Reconsidering the scenario of hospital dormitories, introducing empathy in the sociable VA can enhance their companionship with the patients by acting as good listeners and supporting their mental health. However, research on using empathy to increase the sociability of VAs is limited.

In this research, we delve deeper into the concept of empathetic VAs, specifically toward the vocal emotional expression of the conversational agents, as a response to the emotions demonstrated by the users. Namely, we investigate the following research questions:

RQ1: In what ways can we effectively convey the emotional state of a voice assistant to its users?

RQ2: How well do users perceive the emotional state of a voice assistant when the assistant responds empathetically?

RQ3: What are users' opinions on the potential advent of empathetic voice assistants in everyday lives?

To answer these questions, we first implemented a voice assistant capable of changing its responses empathetically through dynamic speech synthesis and prosodic features. We then conducted a user study where participants completed representative tasks with a VA, which involved enacting a script in a specified emotion and recording their perception of the assistant's emotional state based on its

response. After the task, the users rated the assistant's capabilities on a perceived empathy scale and answered questions regarding their views on empathetic VAs.

We found that speech-based emotional expression is possible in VAs by manipulating verbal content and prosody (RQ1). When implemented precisely, users understand if the assistant is responding to them empathetically or not (RQ2). Furthermore, users support the advent of empathetic assistants but raise concerns and define boundaries that the VAs should not cross (RQ3). This research contributes valuable insights into the field of human-VA interaction and extends the work on empathetic agents.

2 RELATED WORK

2.1 Empathy in Voice Assistants

There have been many proposed definitions of *empathy* in literature. The most widely accepted one is by Hoffman [18], which describes it as "a psychological process that makes a person have feelings, that are more congruent with another's situation than with his own situation." Pelau et al. [34] define empathy as "a combination of emotional reactions and the cognitive understanding of other people's experiences and feelings." The emotional component above refers to care for others, and all the derived emotional experiences, while the cognitive component represents the individual's ability to understand and act according to others' thoughts and feelings. For their work on empathetic agents, Chin and Yi [10] categorised empathy into two types: 1) Other-oriented Empathy, which is usually driven by altruistic motives and care for others; and 2) Self-oriented Empathy, which relates to one's own performance and emotions. Yalcin and DiPaola [46] developed a computational model of empathy for interactive agents, including conversational agents, where they state that empathy is the ability to relate and respond to someone else's emotional state, and consists of emotional communication competence, emotion regulation, and cognitive mechanisms that result in a broad spectrum of behaviour. Lastly, Paiva et al. [32] define empathetic agents in two ways: 1) agents that respond emotionally to situations in a way more congruent with the users' or another agent's current state, or 2) agents that, through their actions, motivate users to respond in a way more congruent to the former's current emotional state.

In our study, we consider the perspective where the conversational agent shows empathy towards its users by manipulating voice features like prosody.

2.1.1 Benefits of Empathy in Voice Assistants. A notable use case of empathetic VAs is *mood elevation*. In a web survey, approximately 79% of the participants (total 130) liked the idea of mood-improving VAs and displayed interest in using them in the future [3]. People believe that morale-boosting through VAs might be handy in various circumstances, such as: 1) after bad days of work, 2) for demanding situations with household members, 3) during sleeplessness and insomnia, or 4) on feeling a strong undesirable reaction after an unfortunate life-event [3]. Chin and Yi [10] also observed that individuals are often left in an unfavourable affect state after behaving aggressively, which can be reduced by altering their moral emotions via an empathetic response.

Virtual personal assistants can also act as good *companions* at times of loneliness [3]. Studies reveal that in most VA-user interactions, small talk accounts for around 4-13% of the daily communication [8], which means many users may be looking for someone to talk to outside their inner circle of humans. However, many people feel that conversations with VAs are not genuine due to the assistant's robotic voice or lack of emotional expressiveness, which results in disappointment among users [10, 38]. Moreover, an emotionally understanding and expressive AI device creates a sense of embodiment, allowing users to better engage with the machine [41]. Designers also use social elements like emotional expression to reduce task load or improve the overall user experience, thus increasing trust in the technology [38].

Another benefit of empathy in VAs may include *communication training* in the form of feedback. For instance, the assistant can express disappointment when spoken to rudely, allowing users to improve their communication skills. In a real-world scenario, the emotionally supportive nature of Alexa has been seen as a plus point when it acts as a public speaking tutor [44].

2.1.2 Drawbacks of Empathy in Voice Assistants. While fostering artificial empathy adds numerous advantages to the interaction experience, it also has disadvantages. For instance, *inappropriate emotional support* may increase recipients' stress and adversely affect their relationship with the provider [10]. Additionally, Pelau et al. [34] state that an AI device replicating human intelligence and empathy can lead to various pessimistic scenarios, where machines may become self-aware and present potential threats to humanity. Hence, we need strict regulations concerning the development of artificial intelligence and its role in society [34]. Research also point out that a closer resemblance between non-human agents and humans may provoke the former to express *negative emotions*, which users may find discomforting and displeasing [8, 24]. Lastly, just like they do in humans, emotions may restrict an AI device from producing an intelligent outcome [27].

2.2 Emotion Expression in Voice Assistants

While extensive work is present on recognising emotions using AI, there is limited literature on expressing emotions by these devices, particularly in the case of VAs. Shi et al. [41] were amongst the first researchers to explore this area, where they attempted to translate the conversational states of VAs into emotional states, and displayed them on a chat screen. They exhibited seven emotions using three different modalities: 1) Facial Expressions, 2) Affective Textbox Movements, and 3) Voice Waveforms.

Feldman et al. [14] expressed emotions in AI devices through virtual 3D avatars, by giving facial expressions using Behavior Markup Language (BML), and speech output that uses both text-to-speech (TTS) conversion and pre-recorded human speech. Chin and Yi [10] attempted to mitigate verbal abuse towards conversational agents by displaying emotions using animated faces. They used two faces for different empathy orientations (i.e., other-oriented and self-oriented empathies), and toggled between fixed and varied facial expressions. Their findings confirmed that the other-oriented empathy style positively influenced users and successfully reduced abusive behaviour.

David et al. [12] designed a collaborative learning environment by integrating emotions with virtual assistants using Alexa, for which they used Amazon Skill Builder (ASB), Amazon Skills Kit (ASK), and Speech Synthesis Markup Language (SSML). Popular VAs like Apple's Siri, Alexa, and Yandex's Alice express various emotions to their users through words [8].

Most literature on emotional expression in VAs involves a display screen. However, little work has been conducted using screen-less smart speakers. Additionally, users have previously shown their disappointment towards the robotic voice of smart speakers, preferring the agent's verbal and vocal characteristics over its facial expressions [10]. This study aims to investigate the possibilities of using voice modulation and speech synthesis techniques to express emotions through a VA.

3 DESIGNING AN EMPATHETIC VOICE ASSISTANT

We used *Google Assistant* to design our VA owing to its dynamic speech synthesis and prosody manipulation features. It also provided us with a UI-based platform, namely *Dialogflow*, enabling us to conduct our user study efficiently.

3.1 Trigger-Response Mapping

Based on literature [7, 21, 40], we defined six trigger emotions to work on for our empathetic assistant, namely, *neutral*, *happiness*, *sadness*, *anger*, *fear*, and *excitement*. In this case, fear represents an individual's uneasiness while using a VA, such as a first-time user who uses a VA and is worried about making mistakes. Next, we defined appropriate empathetic responses for each trigger.

We reviewed existing work on affective empathy [4, 37, 46] and identified two types of affective responses: *parallel* and *reactive*. While parallel responses aim to maintain the current affective state, reactive responses aim to alleviate it, making the former more suitable for positive triggers and the latter for negative ones [37].

In addition to defining the emotional responses, it is essential to translate them through the assistant. Filippi et al. [15] suggests that verbal content and prosodic features are essential to express emotions in audio speech. Therefore, we must manipulate both the response emotion/mood and wordings. Please refer to Table 1 for the complete trigger-response mapping.

3.1.1 Response Emotion and Mood. We directly mapped the following emotions: neutral-neutral, happiness-happiness, and excitement-excitement (parallel responses). While we did not change the emotion for a neutral response, we used an upbeat mood to express happiness and excitement, given user expectations of VAs to express positive responses [5, 6, 25, 38]. Based on literature, we used a calm and firm demeanour, combined with gentleness and patience, as a reactive response to anger [16, 36]. Further, we used a soothing and reassuring response for fear (or uneasiness). While both anger and fear require a calm mood, anger requires an assertive yet gentle touch, and fear needs a more soothing tone. Thus, we split the calm response mood into calm/happy (soothing) and calm/firm (assertive). Lastly, while responding to sadness, instead of mimicking it, we followed a reactive empathy approach and tried cheering the individual [19, 30] using an upbeat mood.

3.1.2 Response Wording. We used words with neutral sentiment to respond to anger and assuring words to respond to fear. Further, we used words with positive sentiments to reduce stress and sadness [1]. Due to a difference in arousal levels of happiness and excitement, we distinguish their responses based on verbal content: neutral words keep the arousal levels low in happiness, and playful ones keep up with excitement's high arousal.

3.2 Prosody Mapping

Literature suggests that cheerful or upbeat voices usually have a higher pitch, faster tempo, and louder volume [39, 42, 43]. Furthermore, Scherer and Bänziger [39] also note that for upbeat or slightly elated voices, the rising and falling slopes of the pitch contour and the final falls are very steep, and the pitch peaks towards the final segments of the utterances. For a *calm* voice, Crumpton and Bethel [11] observe that there is no change in the pitch of the voice, and the pitch contour is relatively flat. Additionally, they suggest that the speech rate and volume for calm voices typically reduce. However, we slightly increased the pitch to add a *firm* grip on the voice. For a *calm/happy* scenario, we follow the happiness trends to slightly increase our calm voice's arousal levels. While happy voices usually have a higher pitch, faster speech, and louder intensity when compared to neutral or calm ones, they are much lesser than excited or typically upbeat voices [11, 20, 31]. Happiness also follows an uptrend, with the pitch peaking around the second last segment of the utterance [39].

Combining our findings from the literature with the default prosodic parameters of voice in Dialogflow, we defined the prosodic mapping for each response mood (see Table 2).

4 USER STUDY

We conducted a user study to test the effectiveness of our VA's empathetic responses, and gather participants' opinions and insights on potentially using such an empathetic agent in their everyday lives.

4.1 Study Design

We asked participants to complete five tasks using our VA. After completing the tasks, they rated the assistant's skills on a perceived empathy scale, created by an amalgamation of standard empathy surveys, and participated in an interview. The experimental design was approved by the Ethics Committee of our university. The experiment lasted approximately 60 minutes per participant.

4.1.1 Tasks. The five tasks involved conversing with our empathetic VA in five different scenarios, requiring users to enact a script with a specified emotion and record their perception of the assistant's emotional state as per its response. Each scenario represented a user's emotional state and matched the trigger emotions, as discussed in Table 1. The five scenarios with their description and representative emotions are detailed in Table 3. All the scenarios were implemented for Google Home Integration via Dialogflow. For each task, we created scenarios that were likely to be relatable, making it easier for the participants to enact the script. In addition, participants were given adequate time to practice before interacting with the assistant. To minimise social awkwardness while interacting with the VA, the researcher left the room and

Table 1: Emotional Mapping for Effective Empathetic Communication

Trigger Emotion	Response Emotion	Response Mood	Response Wording
Happy	Happy	Upbeat	Neutral
Excited	Excited	Upbeat	Playful
Sad	Cheerful	Upbeat	Positive
Fearful	Assuring/Soothing	Calm/Happy	Assuring
Angry	Understanding	Calm/Firm	Neutral
Neutral	Neutral	Neutral	Neutral

Table 2: Prosody Mapping for Response Moods

Response Mood	Pitch	Speech Rate	Volume	Contour	Source
Neutral	Default	100%	0dB	Default	N/A
Calm/Firm	-5%	95%	-3dB	Default	[11]
Calm/Happy	+5%	89%	-1dB	Default	[11, 20, 31, 39]
Upbeat	+15%	104%	+12dB	Uptrend: Pitch increases till the final segment. Very steep final fall. (0%,+30%) (30%,+50%) (60%,+70%) (100%,+50%)	[39, 42, 43]

Table 3: Task Descriptions for User Study

Task ID	Scenario Name	User Emotion	Scenario Description
T1	Customer Service	Angry	User Speaking to an automated Customer Care Service to resolve a problem.
T2	Client Meeting	Fear/Uneasy	First time using a voice assistant to book a restaurant for an important client meeting.
T3	Movie Ticket Booking	Excited	Booking tickets to a long-awaited movie.
T4	A Good Day	Happy	Employee coming home after a good day at Work.
T5	Broken Phone	Sad	User just broke his/her phone and wants to get it repaired.

communicated with the participant through a speaker in the room. The participants could only view their dialogue in the script with a cue to know when the assistant would respond. Finally, our VA responded to these scenarios based on the previously discussed mapping (see Table 2).

After every task, participants were provided with a text field where they shared what kind of emotion they perceived from the assistant's response.

4.1.2 Empathy Rating. After completing all the tasks, users rated their perception of the VA's skills on a perceived empathy scale. The scale was created by combining empathy factors from standardised surveys like the Consultation and Relational Empathy (CARE) [29], the Toronto Empathy Questionnaire (TEQ) [22], and a survey developed by Plank et al. [35]. Table 4 lists the complete set of used empathy factors. All factors were rated on a 5-point Likert Scale, starting from Strongly Disagree (1) to Strongly Agree (5). Like in the source surveys, the empathy score was calculated for each participant's response by adding the points scored by each factor on the Likert scale.

Since the original empathy surveys were designed for humans and not machines, they are insufficient to truly assess the VA's empathetic ability. Thus, we complemented this data with user interviews to gather participants' perceptions of the VA's empathetic ability.

Table 4: Factors in the Perceived Empathy Scale

ID	Empathy Factor	Source
F1	I have a good feeling when interacting with this assistant	[35]
F2	The assistant really understood my feelings about the situation	[35]
F3	The assistant seemed to feel what I needed in this situation	[35]
F4	The assistant made me feel at ease	CARE [29]
F5	I felt that the assistant was really listening to me	CARE [29]
F6	The assistant fully understood my concerns	CARE [29]
F7	The assistant showed care and compassion to my concerns	CARE [29]
F8	The assistant had a positive attitude	CARE [29]
F9	The assistant was in-tune with my mood	TEQ [22]
F10	The assistant was really interested in how I felt	TEQ [22]

4.1.3 User Interviews. We used semi-structured interviews, which were steered using the following open-ended questions:

- How was your experience using the voice assistant?
- What was your perception of the emotions the voice assistant was trying to express?
- Would you say that the assistant was empathetic? Why?

Table 5: Participants' Empathy Ratings

Factor	Min	Max	Mean	Median	SD
Scores per Empathy Factor					
F1	4	5	4.5	4.5	0.5
F2	3	5	4.36	4.5	0.72
F3	3	5	4.36	4	0.61
F4	3	5	4.14	4	0.74
F5	2	5	4.21	5	1.01
F6	4	5	4.71	5	0.45
F7	3	5	4.29	4	0.59
F8	3	5	4	4	0.76
F9	4	5	4.5	4.5	0.5
F10	3	5	4.36	4	0.61
Overall Empathy Score					
Score	36	50	43.43	44.5	4.72

- What is your belief on whether a voice assistant should exhibit any emotions or not? If yes, why, and what emotions? If not, why not?
- If a commercial empathetic voice assistant is created, would you wish to engage with it? Why or why not?

The data obtained from the interview process went through thematic analysis, a widely used approach to analyse qualitative data [13]. We identified patterns and themes across the dataset to generate more refined and usable information. This process was complemented with an affinity diagram to organise and better understand the large, unstructured, and diverse qualitative data [26]. The diagram involved sticky notes (affinity notes) for each relevant word, phrase, or sentence, sorted into hierarchical categories and sub-categories. The qualitative results section compiles the final themes, sub-themes, and their descriptions.

4.1.4 Demographics. We conducted the study with 14 participants, 6 men (43%) and 8 women (57%). All participants were adults between the ages of 22 and 33 years old. Additionally, 11 (79%) participants reported using VAs sometimes in their day-to-day lives, while 2 (14%) use them frequently and 1 (7%) participant never uses them.

4.2 Quantitative Results

This section shares the overall results for the users' perception of the assistant's response and empathy levels.

4.2.1 User Perception of Assistant's Response. As per participants, the VA's emotional reactions to *anger* varied between calm and apologetic. Responses to *uneasiness* were perceived as helpful, assuring, and neutral. Almost all participants positively acknowledged the assistant's enthusiastic responses to *excitement*, and its *happiness*-related responses were deemed fitting. While most participants found the assistant's responses to *sadness* helpful, some associated qualities like generosity, concern, worry, sadness, and cheerfulness with those responses.

4.2.2 Empathy Rating. Table 5 shows that each empathy factor, on average, scored 4 or above on a scale of 5. Overall, the mean and median scores achieved by the assistant on the perceived empathy scale are 43.43 and 44.50, respectively, out of 50 points, i.e., an 85-90% satisfaction rate. Given these results, we can safely say that our participants perceived empathy from the assistant.

4.3 Qualitative Results

Based on our thematic analysis and affinity diagramming, we identified the following themes and sub-themes for the interview data:

- Overall User Perception and Feedback
- User Feedback on Individual Emotional Responses
- User Preferences on Emotional Interactions with Voice Assistants
- User Stance on Emotions in Voice Assistants
- Impact on Society
  - Advantages and Use Cases
  - Potential Concerns
- Establishing Boundaries in Empathetic Assistants

4.3.1 Overall User Perception and Feedback. Almost all the participants had a satisfying experience using the assistant. In addition to an exciting vibe, they also commented on the novelty and potential usefulness of the empathetic VA. Some participants compared the experience with their previous usage of VAs and found our assistant more conversational.

> "To some degree, it was similar to a real-life conversation." (P5)

> "Siri does not have any emotions. I think Alexa has to some extent. But I think this has more evident emotions." (P4)

However, not everyone felt entirely positive about the assistant. While one participant (P13) did not find the responses satisfying enough, another (P6) found the assistant's responses too good to be true.

> "Had these responses come from a human, it would have made sense ... a machine cannot talk like that." (P6)

Moreover, many participants felt that while the assistant was empathetic, it was not fully empathetic.

> "The distinction was clear in positive and negative emotions. But responses to all positive and all negative emotions felt similar." (P10)

This variation in responses prompted us to dig deeper into specific areas where participants felt our assistant lacked in displaying empathy.

4.3.2 User Feedback on Individual Emotional Responses. To understand users' perception of each response given by the assistant, we asked them to think back and elaborate on their experience from each scenario.

Response to Happiness. In general, participants were content with the response provided for Happiness. Most of them believed that the assistant should share our happiness as well. However, some participants felt that sharing happiness was not enough. Instead, to be empathetic, the assistant should further elevate their happiness.

"The assistant's response was happy, but not the happy that I wanted. I want it to elevate me from happy to excited." (P11)

Response to Excitement. Similarly, almost every participant enjoyed the assistant's response of Excitement. In fact, one participant (P4) noticed a jump in the assistant's voice and tried to match his own voice with the assistant's as he continued the conversation. A few participants felt that the response was good but could be better.

"A little more exciting exclamations could be good." (P7)

Response to Uneasiness. Some participants spoke in detail about their perception of Uneasiness, as they thought it was neither good nor bad. One participant (P7), however, clearly liked the response. She believed that a calming tone was reassuring, and she felt confident enough to solve the problem herself, even if the assistant failed to fix the problem. Another participant shared a similar perspective but partly disagreed with the assistant's reply.

"The response to uneasy was good because the task was completed successfully. Had the task not been completed, the fake reassurance would have angered me ... Don't be assuring till the task is done." (P11)

Response to Anger. Only two participants (P8 and P9) felt that the response to Anger was good. All other participants found some flaw or inappropriateness in the response. For instance, one participant (P3) believed that the response to anger was very monotonic or computer-generated. Another participant (P11) preferred a neutral or less calm voice from the assistant when angry because any other response could still trigger her.

Response to Sadness. The user perception regarding the assistant's response to Sadness deviated significantly from our expectations, which probed us to delve further into the context. We found four types of responses that users would prefer hearing from an empathetic assistant when they were sad.

(1) Firstly, we had a group of users who preferred to be *cheered* by the assistant.
"If the users are going to a VA at times of sadness, they probably want to be cheered." (P6)
However, this ideology was shared by only some participants, as others considered it insincere and derisive.
"Don't like fake cheering ... Just be understanding about my situation." (P3)

(2) Next, some participants would prefer the assistant to *share the users' sadness*, as they felt the need to consume the sadness before solving the issue. However, most people disagreed with this ideology, thinking it could further ruin the mood.
"Sadness for sadness can make the user more guilty." (P14)

(3) The third type of response included *cheering or consoling the user indirectly*.
"Help the user indirectly. Maybe by getting a story worse than what happened to them." (P7)

(4) One participant (P10) considered a fourth type of response that combined multiple techniques. She suggested *resonating*

with the sadness first until the problem is fixed and then *gradually cheering* the user.

Even after such differences in opinions, participants agreed that every user might have a different experience and behaviour, and different types of responses may be required. So they suggested using *personality and situations* to decide the type of response. Some even considered giving users the power to choose their response type.

"The users can choose in the start the kind of response they prefer, or maybe the assistant can gauge it over time." (P9)

A significant diversity in responses makes it difficult for us to present a straightforward model to address sadness in an individual. Thus, a better understanding of users' preferences and beliefs is needed.

4.3.3 User Preferences on Emotional Interactions with VAs. While discussing the type of emotional responses users would expect from a voice assistant, we found some interesting insights. In line with our expectations, the overall bias was towards *positive emotions*. For instance, participant P3 suggested that the assistant be supportive, empathetic, and comforting. Other participants also suggested keeping all the negative emotions away.

One participant suggested that the assistant should show either positive or neutral emotions, and did not support any harmful speech or hatred coming from a machine.

While most participants agreed with her statement, two participants suggested otherwise. For instance, participant P9 suggested that the assistant should judge the situation and respond positively or negatively.

"Well! The assistant can be harsh if the user is about to do something dangerous." (P9)

P10 also supported this statement as she believed that the end goal is to reach a positive outcome, and the assistant should try whatever it can to reach it.

4.3.4 User Stance on Emotions in Voice Assistants. Most participants believed having emotions in VA is beneficial, but everyone had different reasons and expectations. For example, P3 believed VA should have emotions because it is a step towards disguising the machine as a human. Participant P14, who had a similar mindset as P3, added that as long as there are limitations, emotional machines and humans can coexist.

Another participant (P7) had a different reason for allowing VA to express emotions. She believed that emotions would be constant in humans and that machines are built to help them. Therefore, if expressing emotions helps humans, the machine should do it.

"The assistant should know how to handle human emotions, not display its own." (P7)

Lastly, three participants (P1, P9, and P13) did not explicitly support machines emoting, as they felt emotions were human characteristics. However, they did not speak against it.

4.3.5 Impact on Society. In this theme, we discuss the impact of emotional or empathetic VA on society regarding benefits, concerns, and use cases, as stated by our participants.

Advantages and Use Cases. Most participants talked about the advantages of having a voice assistant that can express emotions. The first and most mentioned advantage is that users would have someone to talk with and share their feelings. They also suggested that empathetic assistants would be good company during boredom, and could be great companions for lonely older adults.

> "I am an introvert and spend most of the time alone. It'll be nice to have someone to talk to." (P2)

Contrarily, other participants liked having emotions in machines because it added a human feel to the experience. Two participants (P4 and P11) even mentioned that adding emotions to a voice assistant's response will increase convenience and enhance the task's process.

Another example was the use of these assistants in customer care centres (P1, P3, and P6). Participant P1 extended this to any tasks that humans find frustrating.

One use case, which more than one participant agreed upon, was in psychotherapy and medical fields. P5, P6, P10, and P14 found the assistant comfortable and convenient enough to be used in counselling, and handling users with depression and anxiety.

> "I would like to have different kinds of conversations with this assistant. Not the ones I usually have with humans … Maybe like the ones I have with my counsellor? … I don't expect it to be judgemental about what I tell it." (P5)

Lastly, one participant (P9) suggested that an empathetic voice assistant may impact user behaviour positively by training people to become more emotionally intelligent.

Potential Concerns. While the participants saw many benefits of introducing emotions in VA, they shared potential concerns and worries. A common concern that most participants revealed was that a machine entering human territory like this could create a distance among humans. They were also worried about being unable to distinguish between humans and machines.

> "Distinction should be clear … A machine is a machine." (P13)

Another significant concern that some participants (P3, P4, and P12) shared with us was regarding the assistant's emotion recognition accuracy. They believed that if the emotion recognition accuracy of the voice assistant is not 100%, it can make things worse. However, two participants (P10 and P11) did not agree with this concern as they advocated that even humans cannot recognise every emotion perfectly, so how can we train a machine to do that?

Some other concerns included increased dependency on technology, increased laziness, and machines exploiting our vulnerabilities.

As the participants revealed their concerns, we looked further into where they think the line should be drawn to distinguish between humans and machines. We did get some engaging and helpful opinions from them on the matter.

4.3.6 Establishing Boundaries in Empathetic Assistants. The first boundary the participants suggested was ensuring that the machine was not trying to replace humans. For instance, P5 believes that a machine can never and should never replace human touch or warmth. She even suggested a boundary that could clear the distinction for her.

> "Multiple emotions are confusing to adapt. Maybe this could be where machines and humans differ." (P5)

Another participant (P10), who earlier suggested that a voice assistant could be used as a replacement for humans, clarified that she intends the replacement to be only temporary and only when she seeks it or initiates it.

The next suggested boundary was avoiding any form of negativity from the voice assistant's perspective.

> "Don't be judgemental … Avoid giving unasked advice and opinions." (P10)

> "Avoid any inappropriate responses or harshness … Won't go to a machine for lecture or judgement." (P11)

Two participants (P4 and P13) proposed limiting the empathetic capabilities of the assistant to specific situations or emotions. P4 further advocated that the interaction between humans and machines should also be reduced, as he does not want robots to become a part of his life.

> "I want less interaction with robots … It is not my wife or girlfriend. Won't spend hours talking to it." (P4)

5 DISCUSSION

Our study showed that participants effectively perceived the empathetic responses of our voice assistant. We also gathered in-depth insights regarding their perceptions of using empathetic VA, which we discuss next.

5.1 Tackling Negative Emotions

While tackling positive emotions using an empathetic voice assistant is somewhat straightforward, it can be rather challenging when dealing with negative emotions. If negative emotions are not handled well, this has the potential to impact users' mental well-being. One participant even recommended not emoting at all in case of negative emotions, as handling negative emotions can be dangerous. Not only is it essential to respond with the appropriate emotion, but it is also vital that the timing of the emotion is accurate. The Client Meeting scenario was one such scenario where a user felt the assistant could fail. We intended to ensure that our assistant reassured the users when they felt uneasy or slightly worried while trying the assistant for the first time. However, premature reassurance may negatively impact the experience if the task is not completed as per expectation. Additionally, emotional responses depend on the situational context, users' personalities, and moods [37].

Another reason responding to negative emotions could have negative consequences is that everyone has a different preference about how others should react (even when interacting with humans). This is especially true in the case of sadness, for which we saw very different user preferences. While some considered the assistant to respond cheerfully, others wanted it to either sympathise with their situation, solve the problem, or resonate with their sadness. Neither preference can be generalised to a larger population, as the perfect response depends on situations as well as the personality of the user.

However, many users expect social machines with empathetic capabilities, to be highly skilled in social interactions [38], which does not give the assistant much space to make mistakes. A discussion with the participants on handling variations of responses in sadness generated some helpful ideas, including personality selection or gauging the user and the situation. One challenging, yet attractive, solution involves gradually moving from one emotion to another through the conversation, such that, the assistant resonates with the users' sadness until the problem is fixed and then cheers them up. Nevertheless, more research is needed on tackling negative emotions, including the types of responses and the damage control if the suggested response fails.

5.2 Advent of Empathetic Assistants: Boon or Bane

An empathetic assistant can be a boon to the users, be it to improve usability experience or have innumerable use cases in everyday life. As expected, most participants agreed that empathetic assistants would be great companions for older adults or anyone who feels lonely. In the UK, people around the ages of 18-34 (same as the age group of our participants) form the highest proportion of people who feel lonely, worry about loneliness, and seek help for the same [28]. Finding companionship in VA may be helpful in such cases.

Another interesting use case suggested by a participant included using empathetic VA for emotional intelligence training. As established earlier, VAs have faced abusive behaviour from users, and condoning it may result in poor communication skills [10, 34]. Looije et al. [25] claim that therapists often use empathetic behaviour to change users' attitudes and behaviours. So possibly, empathetic reactions may increase a voice assistant's capabilities to train individuals to be more polite with the assistant and improve their communication skills. However, many users worry about social approval when talking to machines, and some also dislike giving the assistants control of the conversations [38]. Therefore, not all users may support the development of empathetic assistants for this purpose.

5.3 Concerns and Boundaries

The first concern many users shared regarding empathetic assistants is the extent to which the assistants can recognise human emotions. They believed that if the emotional perception accuracy of the assistant is not 100%, it may end up responding inappropriately, leading to undesirable consequences. Participants also suggested that an empathetic assistant should only enter the mass market if the emotion recognition aspect is perfected. From a user perspective, it makes sense to introduce these assistants to the public only when we ensure smooth and flawless execution.

One suggestion is to check the confidence levels in the assistant's prediction. If the confidence level is higher than a certain threshold value, the assistant may respond emotionally or default the emotion expression to neutral. We should also endeavour to create well-defined boundaries of what the assistants are allowed and not allowed to do. As some participants suggested, we must ensure that no opinions, judgements, offensive remarks, or negativity should come from the AI's responses, which is already in the works for VAs [8].

Several participants also felt that machines should not replace humans. The concern arose from thinking that users may become more comfortable with the assistants and avoid human contact altogether. Due to these reasons, participants suggested maintaining distance from artificial intelligence and limiting its empathetic capabilities. However, where to limit these capabilities? Some participants suggested limiting the empathetic ability of the assistants to specific situations, but this specificity could be different for different people. While some recommended avoiding emotions for task-based usage, others supported it by saying it improves the overall interaction experience. Nevertheless, all users unanimously agreed that they did not wish for the assistants to feel and emote all the time. Thus, we should consider allowing personalised empathy, which lets users select their preferred communication style (emotional or non-emotional). It may also support increased and reduced levels of empathy based on the users' moods, personalities, and situations. The number of personalisation options users can and should be allowed, opens doors for future research in empathetic conversational agents.

A final concern expressed by our participants was giving machines the power to exploit users' emotional states and vulnerabilities to deceive them. This privacy and trust issue towards AI is genuine, and implementation and actions on relevant privacy policies are still weak [23]. While the privacy policies must be robust, users could also be given the power to decide what information can be shared beyond the machine's local database. They must also be informed of the consequences of sharing and not sharing any particular information.

5.4 Limitations and Future Work

There are several limitations of this study. First, we did not evaluate a fully working prototype in this study. While a plethora of emotion recognition algorithms exist, the aim of this paper was restricted to the emotional expression abilities of a voice assistant and related user perceptions. Future work should conduct an end-to-end evaluation to further examine the robustness of empathetic VA.

Furthermore, we conducted our evaluation in a controlled laboratory experiment with fixed scripts for the user-assistant interaction. This study does not represent the natural conditions of its typical usage and is based on a one-off experience. Conducting a longitudinal evaluation in users' naturalistic settings would yield additional insights into users' emotional perception of empathetic VA. It will also help validate if there are any direct or indirect benefits to humans of using empathetic VA over time.

Lastly, the study was conducted with a relatively small sample and does not represent the general population. Future work should explore a wider population (e.g., older adults) and more user scenarios in their studies.

6 CONCLUSION

In this paper, we designed an empathetic voice assistant and conducted a laboratory-based study, which included task perceptions, empathy rating, and qualitative interviews, to better understand how participants perceived our assistant and the impact of empathetic assistants in everyday lives.

Based on our study, we recommend manipulating verbal content and prosodic features to elicit emotions from a voice assistant's perspective. Secondly, we successfully validated that users perceived empathy from our designed voice assistant. However, as agreed unanimously, the assistant is not fully empathetic yet, and more work is needed to improve its empathetic skills. Lastly, our interviews with the participants generated many insights regarding ways to handle negative emotions, the merits and demerits of empathetic assistants in society, and concerns and boundaries to consider while moving forward in this area.

The contributions of this research are many-fold. Firstly, in the academic domain, it extends the work on empathetic voice-assistant interaction. This paper investigates the possibilities of using speech synthesis to modulate the assistant's voice and tone to elicit the desired emotions. Furthermore, it opens doors for more use cases of virtual assistants, such as mood elevation, communication training, education, and emotional intelligence training. Other scenarios, such as companionship, tutoring, and storytelling, may also be considered. Finally, this research moves towards shifting from a functional to a social interaction style with VA, to create a better connection between humans and machines, which can increase the trust and acceptance of AI devices in our daily lives. To conclude, we must acknowledge that while empathetic assistants may benefit humans in various scenarios, they should not aim to replace or inhibit human contact.

ACKNOWLEDGMENTS

REFERENCES

[1] Sarah F. Allen, Mark A. Wetherell, and Michael A. Smith. 2020. Online writing about positive life experiences reduces depression and perceived stress reactivity in socially inhibited individuals. *Psychiatry Research* 284 (2020), 112697. https://doi.org/10.1016/j.psychres.2019.112697

[2] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottridge. 2018. Understanding the Long-Term Use of Smart Speaker Assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3 (2018), Article 91. https://doi.org/10.1145/3264901

[3] R Bernhaupt, C Murko, G Pottier, and A Battut. 2019. User acceptance of emotion-aware mood-improving voice assistants. In *International Broadcasting Convention (October 2019)*, Vol. 7. IBC, ruwido, Austria, 8 pages.

[4] R. J. R. Blair. 2005. Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition* 14, 4 (2005), 698–718. https://doi.org/10.1016/j.concog.2005.06.004

[5] Michael Bonfert, Maximilian Spliethöver, Roman Arzaroli, Marvin Lange, Martin Hanci, and Robert Porzel. 2018. If You Ask Nicely: A Digital Assistant Rebuking Impolite Voice Commands. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18)*. Association for Computing Machinery, Boulder, CO, USA, 95–102. https://doi.org/10.1145/3242969.3242995

[6] Michael Braun, Anja Mainz, Ronee Chadowitz, Bastian Pfleging, and Florian Alt. 2019. At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, Glasgow, Scotland Uk, 1–11. https://doi.org/10.1145/3290605.3300270

[7] Michael Braun, Jonas Schubert, Bastian Pfleging, and Florian Alt. 2019. Improving Driver Emotions with Affective Strategies. *Multimodal Technologies and Interaction* 3, 1 (2019), 21. https://www.mdpi.com/2414-4088/3/1/21

[8] Daria Bylieva, Victoria Lobatyuk, Dmitry Kuznetsov, and Natalia Anosova. 2021. How Human Communication Influences Virtual Personal Assistants. In *Knowledge in the Information Society*, Daria Bylieva, Alfred Nordmann, Olga Shipunova, and Violetta Volkova (Eds.). Springer International Publishing, Cham, 98–111.

[9] Julia Cambre and Chinmay Kulkarni. 2019. One Voice Fits All? Social Implications and Research Challenges of Designing Voices for Smart Devices. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (2019), Article 223. https://doi.org/10.1145/3359325

[10] HyoJin Chin and Mun Yong Yi. 2022. Voices that Care Differently: Understanding the Effectiveness of a Conversational Agent with an Alternative Empathy Orientation and Emotional Expressivity in Mitigating Verbal Abuse. *International Journal of Human–Computer Interaction* 38, 12 (2022), 1153–1167. https://doi.org/10.1080/10447318.2021.1987680

[11] Joe Crumpton and Cindy Bethel. 2014. *Conveying emotion in robotic speech: Lessons learned*. IEEE, Edinburgh, UK. 274–279 pages. https://doi.org/10.1109/ROMAN.2014.6926265

[12] B. David, R. Chalon, B. Zhang, and C. Yin. 2019. Design of a Collaborative Learning Environment integrating Emotions and Virtual Assistants (Chatbots). In *2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, Porto, Portugal, 51–56. https://doi.org/10.1109/CSCWD.2019.8791893

[13] Ghada Refaat El Said. 2017. Understanding How Learners Use Massive Open Online Courses and Why They Drop Out:Thematic Analysis of an Interview Study in a Developing Country. *Journal of Educational Computing Research* 55, 5 (2017), 724–752. https://doi.org/10.1177/0735633116681302

[14] Sara (Salevati) Feldman, Özge Nilay Yalcin, and Steve DiPaola. 2017. Engagement with Artificial Intelligence through Natural Interaction Models. In *Proceedings of the Conference on Electronic Visualisation and the Arts (EVA '17)*. BCS Learning & Development Ltd., London, United Kingdom, 296–303. https://doi.org/10.14236/ewic/EVA2017.60

[15] Piera Filippi, Sebastian Ocklenburg, Daniel L. Bowling, Larissa Heege, Onur Güntürkün, Albert Newen, and Bart de Boer. 2017. More than words (and faces): evidence for a Stroop effect of prosody in emotion word processing. *Cognition and Emotion* 31, 5 (2017), 879–891. https://doi.org/10.1080/02699931.2016.1177489

[16] Gary Hankins and Carol Hankins. 2000. *Prescription for anger: Coping with angry feelings and angry people*. Barclay Press, USA.

[17] Danula Hettiachchi, Zhanna Sarsenbayeva, Fraser Allison, Niels van Berkel, Tilman Dingler, Gabriele Marini, Vassilis Kostakos, and Jorge Goncalves. 2020. "Hi! I Am the Crowd Tasker" Crowdsourcing through Digital Voice Assistants. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376320

[18] Martin L. Hoffman. 2001. *Toward a comprehensive empathy-based theory of prosocial moral development*. American Psychological Association, Washington, DC, US, 61–86. https://doi.org/10.1037/10433-003

[19] Christopher K. Hsee, Elaine Hatfield, John G. Carlson, and Claude Chemtob. 1990. The effect of power on susceptibility to emotional contagion. *Cognition and Emotion* 4, 4 (1990), 327–340. https://doi.org/10.1080/02699939008408081

[20] Tao Jianhua, Kang Yongguo, and Li Aijun. 2006. Prosody conversion from neutral speech to emotional speech. *IEEE Transactions on Audio, Speech, and Language Processing* 14, 4 (2006), 1145–1154. https://doi.org/10.1109/TASL.2006.876113

[21] Sunyoung Kim and Abhishek Choudhury. 2021. Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study. *Computers in Human Behavior* 124 (2021), 106914. https://doi.org/10.1016/j.chb.2021.106914

[22] Ntina Kourmousi, Eirini Amanaki, Chara Tzavara, Kyriakoula Merakou, Anastasia Barbouni, and Vasilios Koutras. 2017. The Toronto empathy questionnaire: reliability and validity in a nationwide sample of Greek teachers. *Social Sciences* 6, 2 (2017), 62.

[23] Song Liao, Christin Wilson, Long Cheng, Hongxin Hu, and Huixing Deng. 2020. Measuring the Effectiveness of Privacy Policies for Voice Assistant Applications. In *Annual Computer Security Applications Conference (ACSAC '20)*. Association for Computing Machinery, Austin, USA, 856–869. https://doi.org/10.1145/3427228.3427250

[24] Erin Chao Ling, Iis Tussyadiah, Aarni Tuomi, Jason Stienmetz, and Athina Ioannou. 2021. Factors influencing users' adoption and use of conversational agents: A systematic review. *Psychology & Marketing* 38, 7 (2021), 1031–1051. https://doi.org/10.1002/mar.21491

[25] Rosemarijn Looije, Mark A. Neerincx, and Fokie Cnossen. 2010. Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies* 68, 6 (2010), 386–397. https://doi.org/10.1016/j.ijhcs.2009.08.007

[26] Andrés Lucero. 2015. Using Affinity Diagrams to Evaluate Interactive Prototypes. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 231–248.

[27] Juan Martınez-Miranda and Arantza Aldea. 2005. Emotions in human and artificial intelligence. *Computers in Human Behavior* 21, 2 (2005), 323–341.

[28] Timothy Matthews, Andrea Danese, Avshalom Caspi, Helen L. Fisher, Sidra Goldman-Mellor, Agnieszka Kepa, Terrie E. Moffitt, Candice L. Odgers, and Louise Arseneault. 2019. Lonely young adults in modern Britain: findings from an epidemiological cohort study. *Psychological Medicine* 49, 2 (2019), 268–277. https://doi.org/10.1017/S0033291718000788

[29] Stewart W Mercer, Margaret Maxwell, David Heaney, and Graham CM Watt. 2004. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation

process measure. *Family Practice* 21, 6 (2004), 699–705. https://doi.org/10.1093/fampra/cmh621

[30] Jannay Morrow and Susan Nolen-Hoeksema. 1990. Effects of responses to depression on the remediation of depressive affect. *Journal of personality and social psychology* 58, 3 (1990), 519.

[31] Mumtaz Mustafa, Raja Ainon, Roziati Zainuddin, Zuraidah Mohd Don, Gerry Knowles, and Salimah Mokhtar. 2010. Prosodic Analysis And Modelling For Malay Emotional Speech Synthesis. *Malaysian Journal of Computer Science* 23 (2010), 102–110. https://doi.org/10.22452/mjcs.vol23no2.3

[32] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. 2017. Empathy in Virtual Agents and Robots: A Survey. *ACM Trans. Interact. Intell. Syst.* 7, 3 (2017), Article 11. https://doi.org/10.1145/2912150

[33] Daehee Park, Heesung Park, and Scott Song. 2020. A Method for Increasing User Engagement with Voice Assistant System. In *Design, User Experience, and Usability. Design for Contemporary Interactive Environments*, Aaron Marcus and Elizabeth Rosenzweig (Eds.). Springer International Publishing, Cham, 146–157.

[34] Corina Pelau, Dan-Cristian Dabija, and Irina Ene. 2021. What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior* 122 (2021), 106855. https://doi.org/10.1016/j.chb.2021.106855

[35] Richard E Plank, Ann P Minton, and David A Reid. 1996. A short measure of perceived empathy. *Psychological reports* 79, 3_suppl (1996), 1219–1226.

[36] Richard S Podemski and John H Childers Jr. 1986. How to Deal with Angry People. *Principal* 66, 1 (1986), 55–58.

[37] Zeeshan Rasool, Naoki Masuyama, Md. Nazrul Islam, and Chu Kiong Loo. 2015. *Empathic Interaction Using the Computational Emotion Model*. IEEE, Cape Town, South Africa. 109–116 pages. https://doi.org/10.1109/SSCI.2015.26

[38] Ronda Ringfort-Felner, Matthias Laschke, Shadan Sadeghian, and Marc Hassenzahl. 2022. Kiro: A Design Fiction to Explore Social Conversation with Voice Assistants. *Proc. ACM Hum.-Comput. Interact.* 6, GROUP (2022), Article 33.

https://doi.org/10.1145/3492852

[39] Klaus R Scherer and Tanja Bänziger. 2004. Emotional expression in prosody: a review and an agenda for future research. In *Speech prosody 2004, international conference.* International Speech Communication Association, Nara, Japan, 359–366.

[40] Daniel B. Shank, Christopher Graves, Alexander Gott, Patrick Gamez, and Sophia Rodriguez. 2019. Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior* 98 (2019), 256–266. https://doi.org/10.1016/j.chb.2019.04.001

[41] Yang Shi, Xin Yan, Xiaojuan Ma, Yongqi Lou, and Nan Cao. 2018. Designing Emotional Expressions of Conversational States for Voice Assistants: Modality and Engagement. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, Montreal QC, Canada, 1–6. https://doi.org/10.1145/3170427.3188560

[42] Laurel J. Trainor, Caren M. Austin, and Renée N. Desjardins. 2000. Is Infant-Directed Speech Prosody a Result of the Vocal Expression of Emotion? *Psychological Science* 11, 3 (2000), 188–195. https://doi.org/10.1111/1467-9280.00240

[43] Jürgen Trouvain, Sarah Schmidt, Marc Schröder, Michael Schmitz, and William J. Barry. 2006. Modelling personality features by changing prosody in synthetic speech. In *Proc. Speech Prosody 2006.* Saarland University, Dresden, Germany, paper 088.

[44] Jinping Wang, Hyun Yang, Ruosi Shao, Saeed Abdullah, and S. Shyam Sundar. 2020. *Alexa as Coach: Leveraging Smart Speakers to Build Social Agents That Reduce Public Speaking Anxiety*. Association for Computing Machinery, Honolulu, HI, USA, 1–13. https://doi.org/10.1145/3313831.3376561

[45] Jing Wei, Weiwei Jiang, Chaofan Wang, Difeng Yu, Jorge Goncalves, Tilman Dingler, and Vassilis Kostakos. 2022. Understanding How to Administer Voice Surveys through Smart Speakers. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 548 (nov 2022), 32 pages. https://doi.org/10.1145/3555606

[46] Özge Nilay Yalcin and Steve DiPaola. 2018. A computational model of empathy for interactive agents. *Biologically inspired cognitive architectures* 26 (2018), 20–25.