

Can Social Robots Improve People's Attitudes towards Individuals who Stutter?

JULE KÖRNER, University of Waterloo, Canada

SHRUTI CHANDRA, University of Waterloo and University of Northern British Columbia, Canada

KERSTIN DAUTENHAHN, University of Waterloo, Canada

Public attitudes towards stuttering are rooted in stereotypes and misconceptions, leading to negative reactions and discrimination against individuals who stutter. Previous research highlights the positive impact of educational interventions on people's attitudes towards stuttering. The potential of social robots as an educational tool in the context of stuttering awareness remains unexplored. In the present study, we investigate whether a social robot can improve public attitudes when giving an interactive presentation on the topic. We compare its impact with a tablet-only condition. Additionally, we differentiate between two robot conditions - one in which the robot imitates stuttering and another where the robot has fluent speech. In the robot conditions, visuals are shown on a tablet. We used a co-design approach and incorporated the perspectives and experiences of two individuals with lived experiences of stuttering into our study design. A user study with 69 participants reveals significant improvements in attitudes across all three conditions, with no significant difference between conditions. However, participants perceived the robot as significantly 'warmer', more 'attractive' and 'novel' when compared to the tablet. These findings provide valuable insights into the potential of social robots as intervention techniques for improving attitudes in the field of stuttering.

CCS Concepts: • **Human-centered computing** → Touch screens; **Interaction devices**; User interface design; **Interaction design process and methods**.

Additional Key Words and Phrases: Human-robot interaction, Stuttering awareness. Robot-assisted education, Social robots, Autonomous systems, People with Stuttering (PWS).

ACM Reference Format:

Jule Körner, Shruti Chandra, and Kerstin Dautenhahn. 2018. Can Social Robots Improve People's Attitudes towards Individuals who Stutter?. *J. ACM* 37, 4, Article 111 (August 2018), 35 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Stuttering affects approximately one percent of the global adult population [7, 80]. People who stutter (PWS) face daily challenges caused by negative public attitudes, including stereotypes and misconceptions. These challenges affect PWS of all ages, leading to social exclusion, discrimination, and psychological distress [14, 47, 48]. However, public knowledge about stuttering remains limited. Consequently, stuttering has evolved into a stigmatized disorder[75]. Therefore, raising awareness, educating society about stuttering, and providing guidance on appropriate behaviour is crucial.

Authors' addresses: Jule Körner, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada, N2L 3G1, julemkoerner99@gmail.com; Shruti Chandra, University of Waterloo and University of Northern British Columbia, 200 University Ave W, Waterloo, ON, Canada, N2L 3G1, julemkoerner99@gmail.com; Kerstin Dautenhahn, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada, N2L 3G1, kerstin.dautenhahn@uwaterloo.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 0004-5411/2018/8-ART111

<https://doi.org/XXXXXXX.XXXXXXX>

The social environment significantly impacts the well-being of PWS. Previous research successfully demonstrated that education, through methods such as presentations, documentaries and podcasts, positively influences people's attitudes toward PWS. [1, 33, 62, 73, 78], underscoring the importance of raising awareness about the topic. With the rise of technology, especially social media, education via digital platforms has increased [22, 34]. Social robots are one of the advanced technologies [40]. Due to their embodiment, they offer engaging and relatable interactions [27, 50]. In Human-Robot Interaction, social robots have shown promising results in education and health-care [6, 15]. Notably, social robots, have been successfully incorporated into stuttering therapy [81]. However, to the best of our knowledge, to date, none of the existing studies have explored the effectiveness of social robots in educating and changing people's attitudes towards stuttering.

We aimed to investigate if a social robot could improve people's attitudes towards PWS by giving an interactive presentation on the topic. Furthermore, we compared the effectiveness of the robot intervention with a tablet-based intervention. This comparison helped us to assess whether the robot's embodiment, gestures and facial expressions have an additional impact on people's attitudes. Additionally, we examined differences in attitude changes among people exposed to a stuttering robot versus a non-stuttering robot. Next to evaluating attitude changes, we also investigated people's emotional reactions towards PWS. Our goal was to explore if different intervention techniques led to differences in empathetic responses and attitudes towards PWS.

This work has three main contributions. First, it is the first study that tested the use of social robots in the context of stuttering awareness. Second, we developed two fully autonomous technical systems: 1) comprised of a social robot and a tablet; and 2) a tablet. Third, the study involved two collaborators with lived experience of stuttering, who contributed to the educational content and robotic system. The overarching goal of this project is to use social robots to create a more inclusive, supportive and educated environment for PWS.

The remainder of this article is structured as follows. Section 2 presents background and related work, and Section 3 outlines the research questions. Section 4 presents our design choices. Our co-design procedure is presented in Section 5. The system development is described in Section 6. Section 7 outlines the user study, including experimental design, measures and study procedure. The corresponding results are presented in Section 8, and discussed in Section 9. Limitations and future works are described in Section 10. Section 11 concludes the article.

2 RELATED WORK

2.1 Developmental Stuttering

Stuttering is a speech disorder. People who Stutter (PWS) know exactly what they want to say, but their fluent speech is disrupted by involuntary repetitions, prolongations and blocks [37]. Stuttering is classified into developmental and acquired stuttering [44]. Acquired stuttering often results from psychological trauma or brain damage [17, 76]. In contrast, developmental stuttering occurs during child development. It is the most common form of stuttering and will be the focus of this work [64]. When talking about stuttering in the following, we refer to developmental stuttering. Stuttering typically appears between the ages of 2 and 5. While many children naturally recover without receiving therapy, some individuals experience persistent stuttering throughout their lives [37]. There is currently no cure for persistent developmental stuttering. Speech therapy is the primary treatment option to improve speech fluency and coping strategies [23]. Developmental stuttering is more prevalent in males, with a 3:1 male-to-female ratio [37]. Despite ongoing research, the cause of stuttering remains largely unknown.

2.2 Challenges Faced by PWS in our society

Communication plays a vital role in social interactions. It is essential for building relationships and the expression of thoughts and emotions. Stuttering can significantly hinder the ability of PWS to communicate, thereby creating barriers to effortless social interaction [30]. PWS encounter various challenges in their daily lives, including bullying, exclusion and negative peer reactions during childhood and adolescence [25, 48]. Furthermore, stuttering often affects their educational performance and achievements, as school participation requires verbal communication and presentations [41, 63]. It can also affect their social life, making it challenging to form friendships or romantic relationships [39, 41]. Additionally, stuttering can influence job opportunities and career choices [46]. In a survey of 332 interviewed PWS, the majority reported avoiding jobs involving phone work or frequent presentations, and some were overlooked for promotions due to their stutter [39].

2.3 Public Attitude Towards PWS

An underlying problem of these challenges is that stuttering is a stigmatized disorder [8]. Negative reactions are rooted in stereotypes and misconceptions. PWS are frequently perceived as less intelligent, shy, less confident, self-conscious, anxious, introverted, and insecure [11, 24]. These negative beliefs persist across various groups, including family members, speech-language pathologists, teachers, college professors, and employers [2, 3, 18, 29]. Constant exposure to these negative perceptions can cause PWS to internalize such beliefs, leading to self-stigmatization and self-blame [9, 10, 46, 47]. As a result, PWS may underestimate their own abilities and conform to societal stereotypes.

2.4 Intervention Techniques to Change People's Attitude towards Stuttering and PWS

Previous research has aimed to improve the situation of PWS by educating various groups, including teachers, university and college students [1, 33, 45, 55, 73]. These studies used methods such as live presentations, documentaries, videos, and podcasts. They all reduced stigma by providing accurate information about the complexity of stuttering.

2.4.1 Live Presentations and Videos: In one study, Flynn and Louis [33] educated 83 health class students about stuttering. The study was comprised of two conditions: first, participants listened to a live presentation given by a PWS. In the second condition, people watched a video presentation featuring three PWS. Both approaches significantly improved attitudes, but the live presentation had a greater impact. Similarly, Williams et al. [78] used live presentations by a speech-language pathologist (SLP) to educate pre-service teachers about stuttering. They compared two conditions: one where the SLP presented alone and another where a child with stuttering co-presented certain parts of the presentation. They identified significant attitude improvements in both conditions. Nonetheless, there was no difference between the two conditions. A similar study was conducted by Abdalla and St. Louis [1]. In-service school teachers and pre-service trainees watched a documentary video that included factual information about stuttering and tips on how teachers should interact with students who stutter. Furthermore, it featured three students who stuttered. They talked about emotional experiences they had had at school. A positive change in attitude after the intervention was found in pre-service trainees (teachers in training) but not in in-service teachers.

2.4.2 Podcasts: In a study by Nelson [62], thirty-nine college students were divided into two groups. One group listened to a humorous, educational and emotional podcast featuring a PWS. The second group read the transcript. Both groups showed positive attitude shifts without significant differences between them.

2.4.3 Content - Anti-Stigma Strategies: Boyle et al. [11] focused on the content of different interventions. They investigated how different anti-stigma approaches show an effect on people's knowledge of stereotypes, negative emotional reactions, social distance, discriminatory intentions, and empowerment of PWS. Their study consisted of three conditions in which participants watched videos implementing the different anti-stigma approaches. The first condition, labelled "contact", featured a person with lived experiences. In the second condition, "education", myths about stuttering were presented and replaced by factual information. The last condition, "protest" condemned morally unjustifiable attitudes towards PWS. In all conditions, participants exhibited improved attitudes, showing reduced negative emotions towards PWS, and decreased stereotypes and discriminatory intentions. Notably, "contact" had the most positive effect in promoting affirming attitudes towards PWS. In a follow-up study, Boyle and colleagues [12] highlighted the importance of factual information, including dispelling myths and stereotypes, in the "education" and "protest" approaches. In the "contact" approach, having a PWS to share lived experiences and daily challenges and promote positive behaviors was especially impactful.

2.5 Technology Comparisons

Comparisons between technologies are often made to evaluate their effectiveness in different contexts. For instance, voice-based technologies have been compared with screen-based ones to assess their impact on empowering employees at work [65]. In Human-Robot Interaction, similar comparisons focus on the benefits of a robot's physical presence versus screen-based alternatives. These comparisons have been explored in various domains. In smart home control, social robots were compared with other interfaces like voice-control loudspeakers, wall-mounted touch-screens, and a mobile application [57]. While more conventional devices, like touch-screen and mobile applications, were noted for their ease of use and comfort, social robots, with their ability to perform gestures, were seen as more enjoyable and engaging.

In education, social robots have also been preferred over tablets. They were effective in facilitating group work by introducing tasks, managing time, and encouraging group discussions [68]. Similarly, when teaching children new alphabets, social robots offered a more enjoyable learning experience compared to tablet-only conditions [83]. However, learning outcomes were the same in the study. In mental health, participants preferred interacting with a social robot over a laptop during guided breathing exercises [66]. Likewise, in the care of older adults, human-like robots were viewed as more engaging and intelligent compared to character-like robots and tablets [61]. Preference for embodied technology was also evident in comparing physically present social robots for engaging in physical exercises, compared to virtual ones [32].

In summary, across various fields, physically embodied robots are often preferred for their social and engaging qualities [28, 51, 77]. However, to our knowledge, no comparative studies have yet explored social robots' effectiveness in stuttering education, a gap this study aims to address.

2.6 Social Robots for Speech-Language Therapy

The use of social robots in Speech-Language Therapy has been explored, particularly in therapy with children. Spitale et al. [70] found that 20 children significantly improved their linguistic skills after engaging with a physical or virtual robot. While some Speech-Language Pathologists are sceptical about incorporating robots into therapy sessions, they recognise their benefits in enhancing enjoyment and engagement. Other studies also reported positive effects on learning outcomes and engagement when robots assist in therapy [31, 67, 82]. Additionally, social robots have also been specifically tested for stuttering therapy. Yoshikawa et al. [81] found that PWS experience fewer psychological symptoms related to stuttering when talking to a robot compared to a human and when reading in tandem with a robot than reading alone.

3 RESEARCH QUESTIONS

In our study, we are interested in exploring the impact of two different technologies on people's attitudes towards stuttering when delivering a presentation on the topic. The project was guided by the following set of Research Questions (RQs):

- RQ1.** Can the robot intervention improve people's attitudes towards PWS?
- RQ2.** Is there a significant difference in the effectiveness between the robot intervention and the tablet intervention in improving people's attitudes towards PWS?
- RQ3.** Does using a robot that stutters help people more in improving their attitude towards PWS compared to a non-stuttering robot?
- RQ4.** Is there a significant difference in people's empathy towards people who stutter between the intervention techniques (robot and tablet)?
- RQ5.** Does the stuttering robot make people more empathetic towards PWS compared to the non-stuttering robot?

Due to the exploratory nature of our study and the lack of prior research on using social robots in the education about people who stutter, we did not formulate hypotheses for the above-mentioned research questions.

4 DESIGN CHOICES

Our study design involved three distinct conditions: 1) a robot with a tablet, 2) a stuttering robot with a tablet, and 3) a tablet alone. The comparison between the robot and the tablet aimed to assess whether the robot's embodiment has a greater effect on attitude change. This builds on similar comparisons from previous research (see Section 2.5). Additionally, we included the stuttering robot condition to provide participants with a more realistic context of stuttering. We aimed to explore if this imitation further impacts our results. Lastly, we decided to deliver a presentation on stuttering using our technology. This was inspired by previous studies that successfully shifted people's attitudes by raising awareness and reducing stereotypes (See section 2.4).

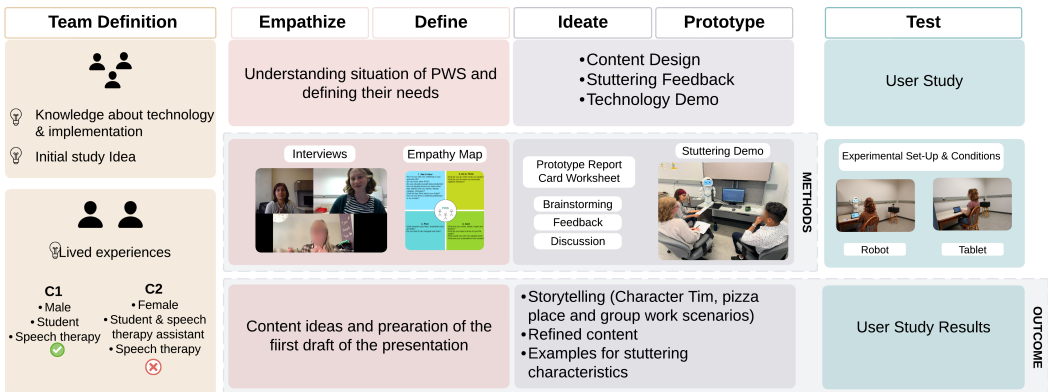


Fig. 1. Co-Design procedure

5 CO-DESIGN

Co-design is a methodology that promotes active collaboration among diverse stakeholders to jointly identify problems and create solutions [74]. This falls within the framework of human-centred design. Our research team included five members: three primary investigators and two collaborators who stutter. Two investigators specialize in Human-Robot Interaction (HRI), while

the third has a background in Cognitive Science. Together, they defined the initial idea of using social robots for stuttering education. Furthermore, based on their expertise in HRI, they handled the technical implementation of the robot and tablet.

We set two criteria for our collaborators: 1) they must have lived experience of stuttering, and 2) they must be able to collaborate with us in person. Given the stigma surrounding stuttering and the vulnerability associated with it, finding collaborators was challenging. We recruited two collaborators through word of mouth, reaching out to contacts within our university and a speech and language therapy centre. Both collaborators have developmental stuttering. One is a female student and pediatric speech therapy assistant, contributing insights both as a person who stutters and as a therapist. The other is a male undergraduate student at the University of Waterloo.

We intended to involve them for several reasons. First, we need to ensure that our educational technology accurately represents the complexity of stuttering. Second, their lived experience should help us better understand the social and emotional challenges faced by PWS in our society and how others should respond to them. Finally, their feedback on the robot's stuttering imitation should help us ensure it was accurate and non-offensive.

Throughout the co-design process (see Figure 1), the two collaborators provided valuable lived experiences, contributing to the development of the educational content for the presentation. Additionally, they played a crucial role in refining the robot's stuttering imitation to be both respectful and realistic.

5.1 Co-Design Procedure

We used the “Design Thinking” methodology to structure our co-design procedure. It comprises five stages: Empathize, Define, Ideate, Prototype, and Test. It was originally conceptualized by Brown et al. [16]. Figure 2 visually presents this “Design Thinking” process.

Due to different schedules and locations, we conducted separate sessions with each collaborator. We scheduled three sessions with both of them, each lasting approximately 60 to 90 minutes. We completed all sessions with our first collaborator (C1) before moving into the second co-design loop with the second collaborator (C2). This allowed us to already incorporate changes in the content and implementation based on the feedback of the first collaborator. Nevertheless, the overall structure and content of the sessions were the same for both collaborators. Before starting the user study, both provided us with their final feedback in written form for our study. Figure 1 gives an overview of the entire co-design process.

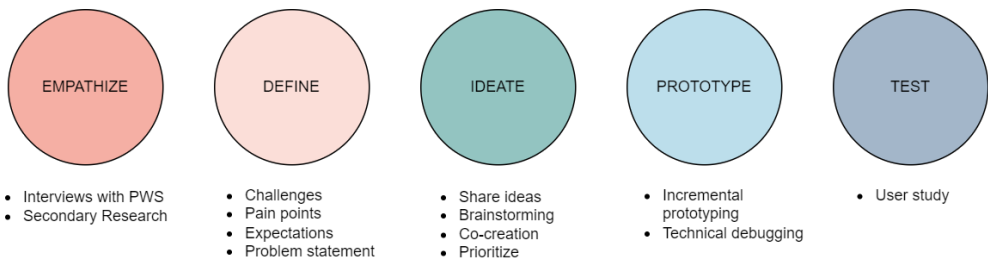


Fig. 2. Five phases of “Design Thinking” Process. Redrawn and adapted from Brown et al. [16]. Secondary research during the Empathize phase includes literature reviews [20, 38] and interviews with Speech-language pathologists [21].

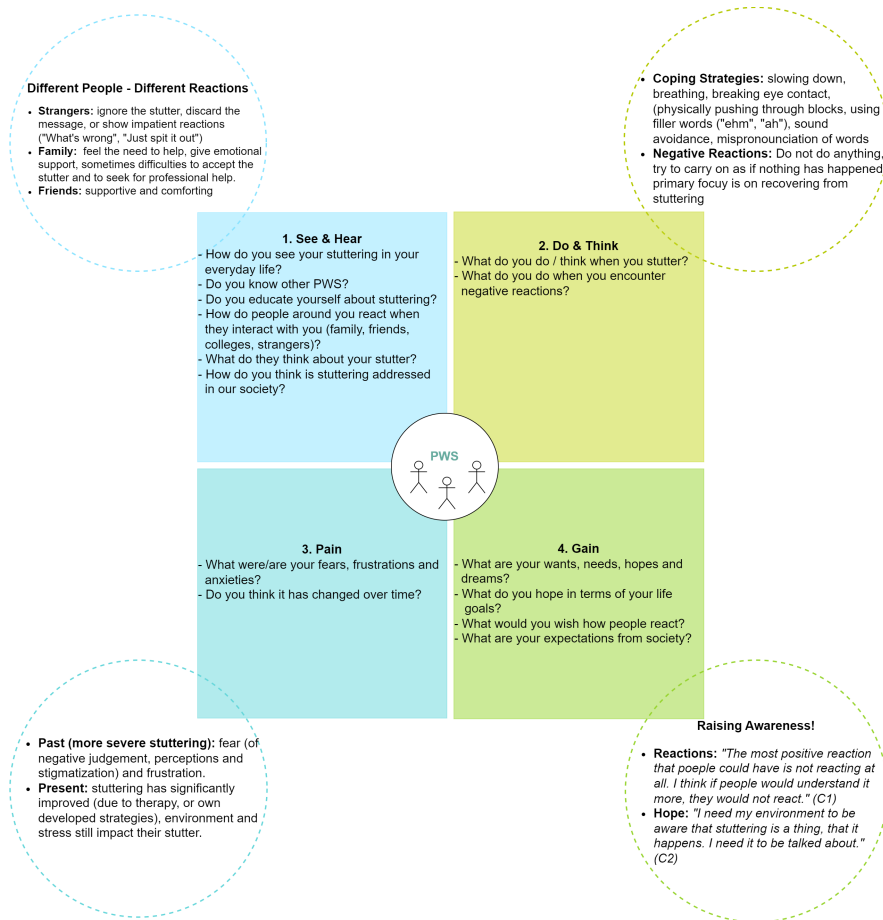


Fig. 3. Empathy Map with the questions we used for interviewing our collaborators. The empathy map was redrawn and adapted from an online source²

5.1.1 Empathize. In the first stage, we aimed to gain deeper insights into our collaborators' personal experiences as individuals who stutter. We conducted one-to-one interviews to understand their perspectives and identify key elements to include in the presentation. To structure these interviews, we used the empathy map¹. It is a UX design tool for understanding and empathizing with a target group. In the context of our study, we aimed to improve the situation of PWS. The empathy map helped us explore the following areas: What does the target group *see*? What do they *do*? What do they *hear*? What do they *think*? What are their *pains* (What challenges do they face)? What are their *gains* (What do they hope for)? Based on these questions, we crafted our interview questions. Figure 3 presents our interview questions together with the collaborators' responses. Their insights were crucial for determining the key aspects to include in our presentation for the user study. Specifically, it led us to include factual information, personal experiences, and practical advice on interacting with PWS.

¹<https://www.nngroup.com/articles/empathy-mapping/>

²<https://medium.com/@davegray/updated-empathy-map-canvas-46df22df3c8a>

5.1.2 Define. In the Define phase, we summarized key insights from the collaborators in the Empathize phase. Complementing this, we incorporated challenges faced by PWS and their expectations from their environment. We also Incorporated findings from our initial literature review, including stuttering characteristics, secondary behaviours, and the prevalence of developmental stuttering. This led to the formulation of the problem statement: “How can we develop technologies that can transmit the voice of people who stutter to mitigate stigma around them?”

Using this, we drafted the initial content for the presentation, setting the stage for subsequent co-design sessions.

5.1.3 Ideate. We had three objectives during the Ideate phase: 1) gather feedback on the stuttering imitation, 2) brainstorm ways to make the educational content more engaging, and 3) decide where to incorporate stuttering imitation in the presentation. First, we demonstrated the robot and its technical capabilities. Second, we presented examples of stuttering imitation to our collaborators. They did not perceive the robot’s stuttering as offensive, so we decided to include it. Next, we engaged in discussions and brainstorming on how to make the content both educational and engaging.

We decided to develop a Tablet Application (App) that ran synchronized with the presentation. It supplemented the robot’s spoken content with visual aids, and incorporated quizzes to maintain user engagement. A comprehensive overview of the App’s design and content is described in greater detail in Section 6.3. Furthermore, we decided to craft a fictional character, ‘Tim’, to illustrate lived experiences. During the presentation, Tim faces common challenges of PWS in two scenarios: at a pizza place and during a group project at university. The scenarios are described in more detail in Section 5.2.

Moreover, we discussed how to effectively include stuttering imitation in the presentation. Our collaborators provided us with sentences to showcase the characteristics of stuttering. To convey repetition, we used the phrase: “I wa, wa, wa want coffee”. Prolongations were demonstrated through the sentence: “I am e mmm motional”, and blocks were represented by the phrase: “You look b (pause) beautiful”. It is noteworthy that the spelling of the repetitions aligns with our implementation on the robot in order to achieve results similar to actual stuttering speech.

Table 1. Final feedback from the two collaborators for the final stuttering robot implementation.

Category	Feedback
Overall Impression	- Informative presentation that gives an excellent demonstration of stuttering. - The robot’s movements paired with the speech and tablet is engaging. - Positive impression of interactive quizzes and visual scenes
Speed	- “The speed is fine it could be a little faster but that could just be me.” - Good speech rate and appropriate pauses between the sentences - Good change in voice tones for the different talkers
Content	- “I think the content is the best part, integrating stories and information makes the presentation some of the best content I’ve seen for non-stutters.” - Gives a complete picture of stuttering.
Robot’s Gestures	- Similar to natural movements (during presentations and stuttering)
Robot’s stuttering	- “I was actually pretty impressed with it, especially how the robot makes the sound of someone attempting to finish a word during a speech block.”
General Suggestions	- Moving storytelling parts to the beginning to give a more concrete idea of stuttering before getting into the facts.

5.1.4 Prototype. We integrated all the input gathered during the Ideate phase into our content for the presentation and implemented it on the robot and the App. The final version, which synchronized the robot and App, was demonstrated to our collaborators. We sought feedback using the "Prototyping Report Card" ³ which consists of two key parts. The first part defines specific aspects that we aim to test: the robot's speed, voice, speech, content, and stuttering imitation. Based on the feedback, we adjusted the robot's speed and replaced certain words that were not pronounced well. The overall content was positively received. Concluding our iterative refinement process, we shared a video of the final version of the stuttering robot with both collaborators. We received feedback on four areas: 1) the overall impression, 2) the speed of the App and robot speech, 3) the presentation content, and 4) the robot's gestures. This feedback is summarized in Table 1.

5.2 Final Content

To summarize, the content across all conditions encompassed three main elements: educational (information to promote awareness about stuttering), storytelling, and tips. To enhance engagement and educational outcomes, we also integrated different interactive components⁴. The elements are outlined below:

- (1) **Educational Elements:** The educational elements provided a comprehensive definition of stuttering, covering its characteristics, prevalence and general factual knowledge.
- (2) **Storytelling Elements:** We used a character named 'Tim', a person who stutters, to illustrate two scenarios. In the first, Tim's stuttering becomes apparent while ordering a pizza, leading to impatience and rude comments from the vendor and customers. In the second, during university group work, Tim's stuttering causes his fellow students to dismiss his idea. To enhance the storytelling experience, we varied the narrator's voice with different pitches to represent different characters.
- (3) **Tips Elements:** We concluded with practical tips for interacting with PWS, emphasizing the importance of patience, avoiding interruptions or finishing their sentences, staying relaxed, and showing genuine interest.
- (4) **Interactive Components:**
 - *"Pen & Paper" exercise:* Following the storytelling elements in the presentation, a brief exercise encouraged participants to suggest more appropriate reactions for individuals presented in the scenarios. They were asked to write down their thoughts on paper.
 - *App:* We implemented an App that displayed accompanying slides, as well as quizzes to complement the spoken content and to maintain user engagement. The App was running in sync with the verbal presentation. The App design and content will be described in greater detail in Section 6.3

5.3 Robot Behavior

We used the QTrobot⁵ from LuxAI ⁶.

- (1) **Voice:** We purchased the "Ella" voice from LuxAI and used it for both the robot and tablet. The volume was set to 90% (0-100 QT volume setting) to ensure clarity. The authors together

³https://design-kit-production.s3-us-west-1.amazonaws.com/Design+Kit+Method+Worksheets/DesignKit_prototypereportcard_worksheet.pdf

⁴See the supplementary material for the full interaction script.

⁵<https://luxai.com/19>

⁶<https://luxai.com/>

reviewed several voice options and decided on a human-like but clearly robotic voice, consistent with the robot's robotic appearance. Research has shown that how a robot's voice is being perceived depends on factor such as task, and robot embodiment [60].

- (2) **Stuttering Imitation:** The stuttering imitation was only used in the stuttering robot condition. With our two collaborators, we decided that the robot should only stutter in selected situation, not throughout the entire presentation, as we were unsure how participants would perceive it. In order to avoid that participants would perceive the robot's stuttering as a 'malfunction', in the robot stuttering condition, the robot states in the introductory comments of the session: "Since you may not be familiar with stuttering, I will try to imitate characteristics of stuttering at some points during this presentation. But I'm a robot. So I may not be the best at doing so. But I think you can still get an idea." The robot first demonstrated the three stuttering characteristics (repetition, prolongation, block) with the previously mentioned sentences. Then applied these during Tim's speech in the scenarios.
- (3) **Gestures and Facial Expressions:** To enhance the robot's role as a presenter, we implemented a range of simple gestures, such as raising and lowering its arms, moving them from side to side while speaking, or pointing to the tablet to guide participants' attention to the screen for quizzes. It also nodded or shook its head in response to input. We prioritized gestures over facial expressions. The robot maintained a natural smile throughout the presentation to create a friendly atmosphere.

6 SYSTEM DEVELOPMENT

6.1 Robot and other Materials

The QTrobot is a table-top humanoid social robot and has 14 degrees of freedom, enabling the execution of different gestures through head and arm movements. Furthermore, the robot's face is a display which can convey various facial expressions and emotions. The robot's pitch modulation enabled distinct character portrayals.

Additionally, we developed an App using the Kivy framework⁷. In each of the three experimental conditions, a tablet served as the interface for executing the App, displaying visuals, and interacting with the robot through button presses. To ensure uniformity between robot and tablet conditions, we recorded and used the robot's voice in the tablet condition. Thus, there was no difference in terms of the voice used between the robot and tablet conditions. More details on the App are provided below (Section 6.3).

During the interaction, participants engaged in an exercise, providing answers using pen and paper. We created an online survey using Qualtrics⁸ to administer pre- and post questionnaires. Participants used a computer in the lab for this assessment. Finally, the session was recorded using a camera.⁹

6.2 System Architecture

The system architecture of our study is shown in Figure 4. The QTrobot consists of two computers: QTRP (Raspberry PI) and QTPC (Intel NUC i5/i7 PC). They are internally connected via LAN and run on Ubuntu Linux. The Robot Operating System (ROS) manages the communication between QTPC and QTRP and enables to control the robot services like text-to-speech or face and emotion detection. All computational processes run in nodes that can communicate with each other by publishing and subscribing to ROS topics. We also developed an interactive node for the Kivy app,

⁷<https://kivy.org/> based on Python language

⁸<https://www.qualtrics.com/>

⁹Video of a full session of the stuttering robot condition: https://youtu.be/LcjLXAkS_fY

enabling users to engage with the robot and providing visual components. The tablet connects to the robot via the QTrobot hotspot. The whole system operated fully autonomously without any experimenter intervention.¹⁰

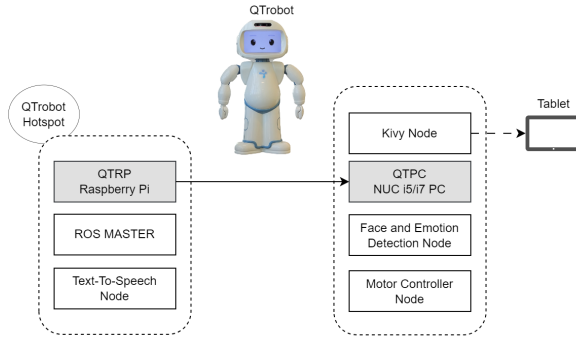


Fig. 4. System Architecture



Fig. 5. Examples of the three different components of the Kivy App used in the study.

6.3 App Development

We developed a Tablet App using Kivy to display presentation slides and enable user interaction with the robot. The App aligns with the presentation and includes the following three main components. Examples are shown in Figure 5:

- (1) **Information Display:** This component presented written information that reinforced the verbal content of the presentation. See Figure 5a.
- (2) **Illustrations:** We included illustrations to visually support storytelling parts. One illustration is shown in Figure 5b. Slides were designed by using assets from Freepik¹¹.
- (3) **Interactive Quizzes:** The App included multiple-choice quizzes. Users selected answers by clicking labeled buttons. The system (robot or tablet) provided feedback and corrected incorrect responses. There were two types of quizzes: educational and “filler” questions. The educational quizzes focused on stuttering-related knowledge, with questions such as “How many people worldwide are stuttering?” and “What do you think is the most common stereotype related to stuttering?” These questions were asked before providing information to ensure active engagement rather than passive listening. The question-answer pairs were created by the research team, based on what made sense in the context of the presentation.

¹⁰Note, we will share the code repository if researchers contact us.

¹¹<https://www.freepik.com/>

The “filler questions” aimed to further engage users, including “Have you ever met someone who stutters” and “Do you like working in a group”. All quiz screens are depicted in Figure 6.

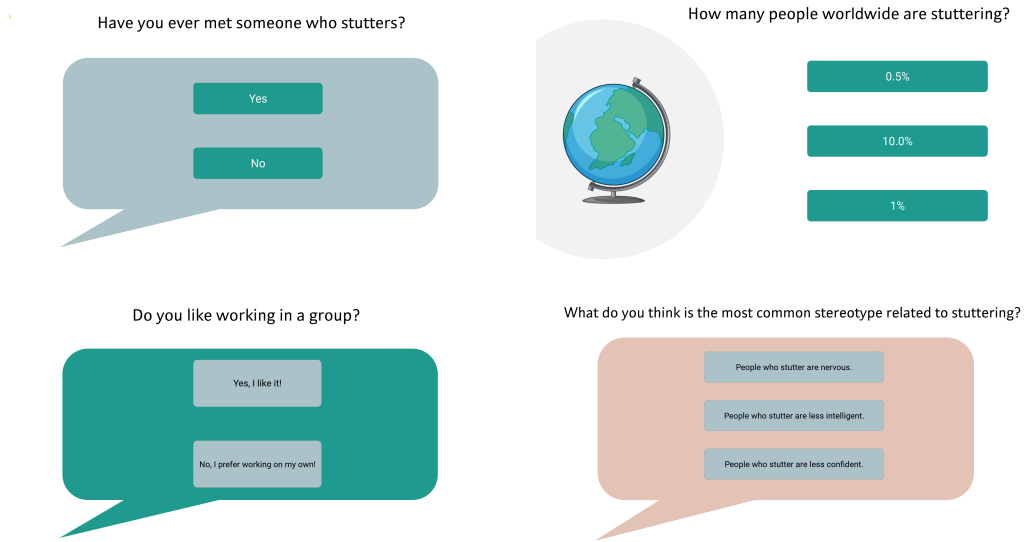


Fig. 6. Interactive question screens. The left column shows “filler questions”, the right column “educational questions”.

7 USER STUDY

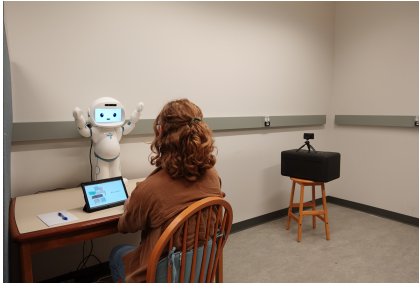
7.1 Experimental Design

The study used a between-participant design with three conditions: two involving a social robot and a tablet, and one using only a tablet. In the first robot condition (R) the Qtrobot served as the narrator for the presentation, supplemented by slides on a tablet. In the second robot condition (RS), the Qtrobot delivered the presentation while mimicking stuttering at specific points, also accompanied by tablet slides. In tablet condition (T), the presentation was delivered solely through the tablet, with visual slides complementing the audio.

The content across all three conditions was the same, differing only in the technology (Qtrobot with a tablet vs Tablet-only) and the inclusion of stuttering imitation in RS).

7.2 Experimental Setup

The study was conducted in one experimental room of the Social and Intelligent Robotics Laboratory at the University of Waterloo. Figure 7 presents the experimental setup for both robot conditions and the tablet condition. Each session included one-to-one interaction between a participant and the robot or the tablet. In the robot conditions, the Qtrobot was positioned on a table, facing the participants (see Figure 1). Participants sat in front of it at an approximate distance of 50cm. The tablet was placed in front of the robot. Participants were allowed to adjust the distance to the tablet according to their preferences. In the tablet condition, the tablet was placed in front of the participants. Participants could again adjust the distance to the tablet themselves. The sessions were video recorded with a camera positioned to capture participants’ profiles and the technological



(a) Robot conditions



(b) Tablet condition

Fig. 7. Experimental set-up

setup. The experimenter set up the technology and positioned the camera at the start of each session.

7.3 Participants

A total of 69 participants were involved in the study. We recruited participants from among students and staff at the University of Waterloo. Participants came from different backgrounds: Twenty-seven were pursuing a Bachelor's degree, 34 were enrolled in Master's programs, 5 PhD students, one postgraduate student, one summer research student, and one staff member. Participation was voluntary. All participants received a 15 CAD gift card for participating in the study, which took approximately one hour. The study received full ethics clearance from the University of Waterloo Human Research Ethics Board. We advertised our study by posting flyers on campus and sent emails to both graduate and undergraduate students. Participants were assigned to one of the three experimental conditions, ensuring a balance in gender across the three conditions. Condition R was composed of 14 female ($M_{age} = 23.07$, $SD = \pm 1.592$), and 9 male ($M_{age} = 25.44$, $SD_{age} = \pm 1.592$) participants. Condition RS also composed 14 female ($M_{age} = 24.57$, $SD_{age} = \pm 4.071$), and 9 male ($M_{age} = 25.33$, $SD_{age} = \pm 4.472$) participants. Condition T composed 13 female ($M_{age} = 23$, $SD_{age} = \pm 3.215$), 9 male ($M_{age} = 23.22$, $SD_{age} = \pm 2.87$) and one non-binary ($age = 18$) participant.

7.4 Measures

To answer our research questions on people's attitudes towards stuttering before and after the intervention, we assessed their empathy levels, their general attitudes towards technology, their perception of the robot's social attributes, and participants' user experience using standardized questionnaires. We also included open-ended questions. Figure 8 outlines when each measure was administered during the session. The questionnaires are as follows:

7.4.1 Attitude Towards Stuttering. To assess participants' attitudes towards stuttering before and after listening to the presentation, we purchased the Public Opinion Survey of Human Attributes-Stuttering (POSHA-S) by St. Louis [71]. It is a standard tool that was developed to measure public attitudes towards stuttering. Questionnaire items are summarized under two *subscores*: Beliefs about PWS and Self Reactions to PWS. These subscores reflect the average of multiple *component scores*. Beliefs components are: "Traits/Personality", "Help from", "Cause", and "Potential". Self-Reactions components are: "Accommodating/Helping", "Social Distance/Sympathy", "Knowledge/Experience", and "Knowledge Source". An *Overall Stuttering Score* (OSS) determines people's overall attitude towards stuttering; it is the mean of the two subscores. Items subsumed under the Beliefs subscore are "external" to the respondent [43, 54, 56]. They reflect the respondent's thoughts or beliefs about,

for example, the cause of stuttering or personality traits of PWS. In contrast, Self Reaction items involve the respondent personally, asking how participants would react or feel when interacting with a PWS.

7.4.2 Empathy. To determine participants' empathy levels, we used the Toronto Empathy Questionnaire (TEQ) by Spreng* et al. [71]. Following the intervention, participants watched a video showing PWS of different ages and genders talking. We measured their emotional response to this exposure using 24 adjectives reported by Batson et al. [5]. They can be grouped into adjectives reflecting feelings of distress and empathy [4]. Distress-related adjectives included: *Alarmed, Grieved, Troubled, Distressed, Upset, Disturbed, Worried, Perturbed, Worried*. Empathy-related adjectives included: *Sympathetic, Softhearted, Warm, Compassionate, Tender, Moved*. We used a five-point Likert scale (1 = not at all, 5 = extremely) to measure the elicited emotional response [69]. In this paper, we call this questionnaire "Emotional Response Questionnaire".

7.4.3 Perception of the Technology. We used the Robotic Social Attributes Scale (RoSAS) to measure the social perception of the technology used (robot, robot with stuttering, and tablet) [19]. This scale consists of 18 adjectives that describe various social attributes. Respondents rate each adjective on a nine-point likert scale. Additionally, we included the User-Experience Questionnaire (UEQ) to evaluate participants' experiences with the technology [49].

7.4.4 Open-Ended Questions. We also incorporated open-ended questions into our study. Prior to the intervention, we gathered information about participants' experiences with social robots and their knowledge about stuttering.

- (1) **Prior Encounter with Social Robot:** Participants were asked if they had ever seen a social robot and where they encountered it (e.g. in-person, internet search, film/television or magazines).
- (2) **Interaction with Social Robot:** We assessed if they had interacted with a social robot before and asked them to specify the nature and setting of the interaction.
- (3) **Programming Experience with Social Robot:** Participants who have programmed a robot before were asked to describe the context in which they did.
- (4) **Knowledge about Stuttering:** Participants were asked if they already have some knowledge about stuttering and asked them to elaborate on what they knew.
- (5) **Prior Interaction with People who Stutter:** Participants were asked if they have interacted with a person who stutters, the duration of the interaction and their experience.

Following the intervention, we asked questions about participants' perception of the intervention and their main takeaways.

- (1) **Effectiveness of the Intervention:** We wanted to know if participants perceived the intervention as effective in helping them learn about stuttering. They were then asked to specify which aspects of the intervention (e.g. about technology or content) they found most effective.
- (2) **Main Takeaways:** Participants were asked to identify their 2-3 main takeaways from the intervention regarding stuttering.
- (3) **Important Points to remember when Interacting with People who Stutter:** Participants were asked to describe the three most important things to remember when interacting with individuals who stutter.
- (4) **Challenges Faced by People who Stutter:** They were asked whether they think that people who stutter face regular challenges in their daily life.

7.5 Procedure

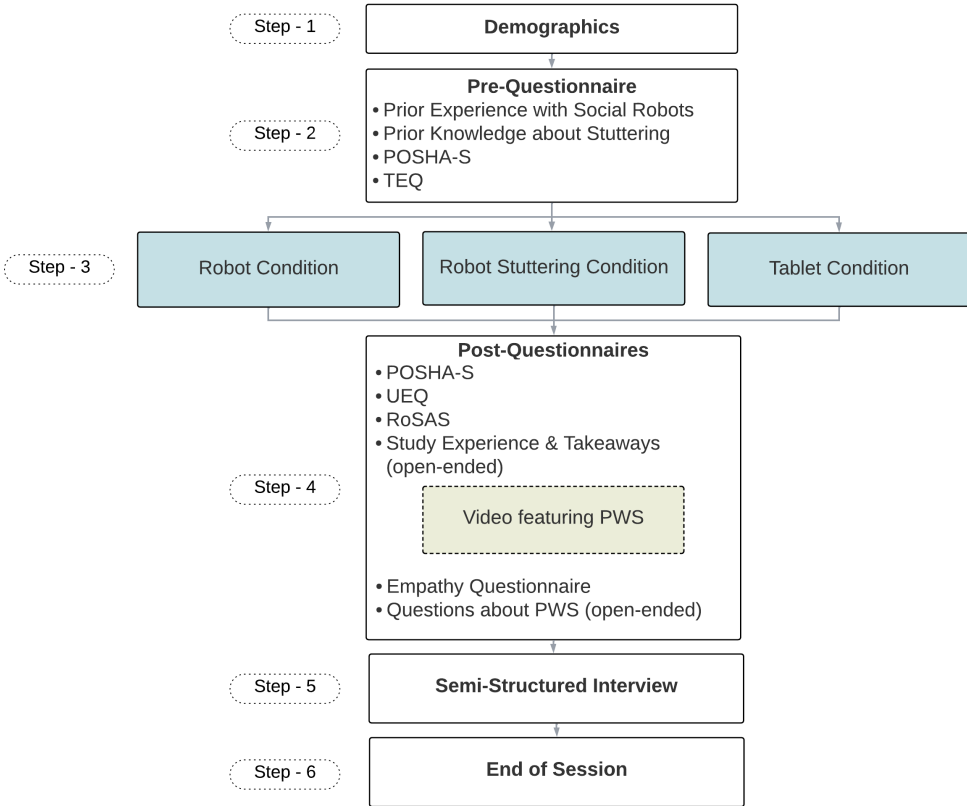


Fig. 8. A flowchart diagram illustrating the user study procedure. The dashed box only applies to condition R and RS.

7.5.1 Semi-Structured Interview. The last instrument was a short semi-structured interview at the end of the session conducted by the first author in the lab. Participants were asked to share their overall impression and provide feedback on the study. For the RS condition, we explicitly asked participants about their perception of the stuttering imitation since, as mentioned above, this has been one of the research team's concerns regarding the use of robots in this application domain that has not been studied before. The duration of the interviews varied individually based on each participant's response. All responses were audio recorded for subsequent analysis.

The study was conducted over four weeks. Figure 8 provides an overview of the whole study procedure.

Step 0 - Introduction Participants were first welcomed by the researcher and the study procedure was explained.

Step 1 - Demographics Questionnaire: After reading and accepting the consent form, participants were requested to respond to demographic questions. Predefined options were provided for participants to specify their gender and current education/occupation. Additionally, participants

had the opportunity to enter their own responses in a textbox. Age was accessed via a textbox. Age was collected through a textbox. They had the option of not answering any of these questions.

Step 2 - Pre-Experimental Questionnaire: Following the demographics, the pre-questionnaire was conducted. We started by asking questions about participants' prior experience with social robots (see subsub-section 7.4.4). Following this, we wanted to get some insights into their familiarity with stuttering (see subsub-section 7.4.4). After that POSHA-S was conducted. Finally, we administered the Toronto Empathy Questionnaire.

Step 3 - Intervention/Presentation about Stuttering: After completing the pre-questionnaire, participants were assigned to one of the three experimental conditions. They were seated and introduced to the technology (either robot and/or robot). The instructor verbally asked participants to carefully listen to the instructions provided by the technology during the intervention. After that, the experimenter started the video recording and left the room. Participants could start the presentation themselves by pressing the "Start Presentation" button on the tablet. The presentation lasted approximately 20 minutes, depending on how long participants took to answer the questions.

Step 4 - Post-Questionnaires: After the presentation, participants were asked to complete the post-experimental questionnaire. We administered the POSHA-S questionnaire once more to measure changes in stuttering attitudes before and after listening to the presentation. Following this, participants answered the UEQ and RoSAS. Additionally, we asked open-ended questions about the study experience and main takeaways (see subsub-section 7.4.4).

As an additional component to measure elicited empathy among participants due to the exposure to different technologies, we embedded a 2:17 minute video that showed five short clips^{12,13,14,15,16} of PWS talking. The video featured PWS of varying ages and gender (three females and two males). Our aim was to measure the emotional responses elicited in participants towards PWS after the presentation. To achieve this, we asked participants to fill out the Emotional Response Questionnaire immediately after they had watched the video. Afterwards, we asked a few questions related to PWS (see subsub-section 7.4.4).

Step 5 - Semi-structured interview: In the final step, the first author conducted a short semi-structured interview.

Step 6 - End of the Session: Finally, we expressed our gratitude to the participants and provided them with a gift card as an appreciation of their time. The whole study session lasted on average between 45 minutes and 1 hour.

7.6 Data Collection and Analysis

We used three different forms to collect our data: 1) Survey - We used Qualtrics to collect our questionnaire data; 2) Hard copy / Paper records - For the "Pen and Paper" exercise during the intervention; and 3) video recordings.

7.6.1 Quantitative Data. The quantitative data includes the POSHA-S data to evaluate participants' attitude changes towards stuttering (RQ1, RQ2, RQ3). We analysed the Toronto Empathy Questionnaire (TEQ) and the Emotional Response Questionnaire data to investigate the research questions related to empathy (RQ4, RQ5). Specifically, we focused on the six adjectives- sympathetic, softhearted, warm, compassionate, tender and upset since they were defined as indicators of feelings

¹²The Post Athens, "Living with a Stutter" (0:13-0:30): https://www.youtube.com/watch?v=BUMV_f8OfXQ

¹³Special Books by Special Kids, "The Stuttering Advocate (Ignoring Mainstream but not by us)" (7:36-8:04): <https://www.youtube.com/watch?v=2EUdPhzAPtM&t=83s>

¹⁴Matice Ahnjamine, "Q&A about Stuttering" (1:13-1:36): https://www.youtube.com/watch?v=OWL5LT_U7uE&t=78s

¹⁵STAMMA, "Becoming a YouTuber when you Stammer" (2:33-2:50): <https://www.youtube.com/watch?v=YC5aQNOIFkM>

¹⁶Special Books by Special Kids, "Life with a Stutter and Social Anxiety (10:05-10:28 & 12:08-12): <https://www.youtube.com/watch?v=Bqfwc5Ebx1Q>

of empathy [4]. Furthermore, we used UEQ and RoSAS data to assess participants' perceptions of the technology used in our study.

We applied a Shapiro-Wilk statistical test for all of our data to examine whether it follows a normal distribution. To assess statistically significant differences from pre to post in POSHA-S scores, we used a two-sided paired t-test for normally distributed data and a Wilcoxon Signed-Rank test for non-parametric data. For all questionnaires (POSHA-S, TEQ, Empathetic Response Questionnaire, RoSAS, and UEQ), to determine if there are statistically significant differences between conditions, we applied a one-way ANOVA test for normally distributed data and a Kruskal-Wallis test for non-parametric data. Post-hoc analyses included Dunn's test for pairwise comparisons to identify conditions with significant differences.

To measure the effect size of statistically significant results (pre/post score differences between the three conditions), we applied Cohen's d (d) for a two-sided paired t-test: $d = 0.2$ (small effect), $d = 0.5$ (moderate effect), $d = 0.8$ (large effect). We used $r = \frac{z}{\sqrt{N}}$ for the effect size calculation of Wilcoxon Signed-Rank tests. The values are interpreted as the following: $0.10 < 0.3$ (small effect), $0.3 < 0.5$ (moderate effect) and ≥ 0.5 (large effect). We applied the eta squared (η^2) measure for significant differences detected by Kruskal-Wallis. The values range from 0 to 1: $0.01 < 0.06$ (small effect), $0.06 < 0.14$ (moderate effect) and ≥ 0.14 (large effect).

7.6.2 Qualitative Data. To analyze the qualitative data from the interviews and "Pen & Paper" exercise, we employed a thematic analysis (TA) to identify emerging themes [13]. In this study, we adhered to an inductive or "bottom-up" approach. Meaning, we derived themes directly from the data itself without considering predefined themes. Following the six-step framework proposed by Braun and Clarke [13], our analytical process involved: 1) Familiarizing with the data, 2) Generating initial codes, 3) Searching for themes, 5) Defining and naming themes, and 6) Producing the report. A code was accepted if it appeared in at least five comments. The qualitative data analysis was done by two of the authors, who cross-checked and reached an agreement on all the data.

8 RESULTS

8.1 Prior Experiences with Social Robots and Stuttering

As an initial part of our data analysis, we investigated participants' prior experiences with Social Robots and Stuttering.

8.1.1 Prior Experiences with Social Robots. These results only include the 46 participants who were assigned to one of the robot conditions. In our study, 26 out of 46 participants (57%) have seen a social robot before our intervention, while 20 individuals (43%) have not seen a social robot before. Among those who have seen a social robot, the majority (22 people, 81%) have reported seeing the robot in person. A small percentage (7.8%) has encountered the robot through films, television, or internet searches. Three participants (12%) specified other settings, such as encountering a robot during a friend's research or as part of a robotics class at university.

Considering the appearance of the observed robots, the predominant form was humanoid, with 21 participants (81%). Two participants (4%) reported encounters with animal-like robots, while another four mentioned cartoon-like robots and the Fetch robot [79]. 20 out of 46 participants (43%) in the robot conditions had interacted with a social robot prior to our study. Only 2 participants (4%) out of 46 participants had experience in programming robots.

8.1.2 Prior Experiences with Stuttering. Across all three experimental conditions, the majority of our participants indicated prior knowledge about stuttering. 41 (60%) out of 69 individuals already had prior knowledge, while 26 (38%) participants did not. Two respondents (3%) provided nuanced responses by selecting "other". One of the participants expressed familiarity solely through media

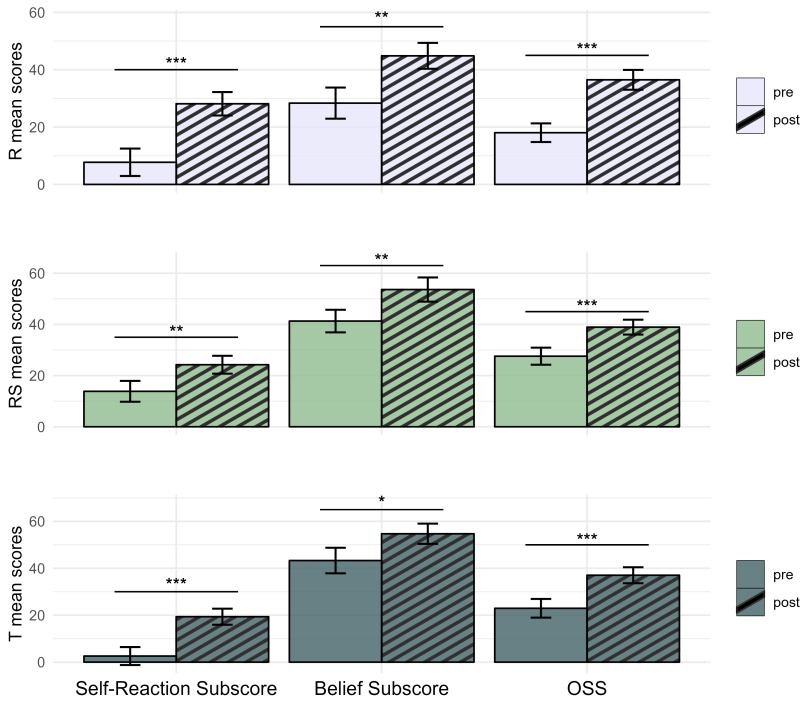


Fig. 9. Average Scores for Self-Reaction, Beliefs, and Overall Stuttering Score (OSS) before and after the intervention. The scale goes from -100 to +100. Asterisks indicate significant difference between pre- and post-scores. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

exposure, without formal education on the causes of stuttering. The other one characterized their knowledge as “very little”. The distribution of responses regarding this question was fairly even across all three conditions. In the RS condition, 12 individuals replied with “yes”, while in the R condition, 15 did, and in the T condition, 14 did so. Participants in all experimental conditions also exhibited similar levels of previous interactions with PWS. In total, 52 (75%) respondents reported prior encounters with PWS, with 18 in RS, 17 in R, and 17 in T. Conversely, 17 participants (25%) had no prior encounters, with five in RS, six in R, and six in T.

8.2 POSHA-S: Attitude towards PWS

8.2.1 Differences in Attitude Changes between Conditions. We further tested for significant differences in the effectiveness of the three conditions in changing people’s attitudes towards PWS. A two-way mixed ANOVA was conducted to examine the effect of time (pre-test and post-test) and condition (R, RS and T) on overall stuttering scores. Results revealed a significant main effect of time, $F(1, 66) = 94.21, p < .000, \eta^2 = .588$, indicating a significant increase in overall stuttering scores from pre-test to post-test across all conditions (see Table 2 and Figure 9). However, the main effect of the condition was not significant, $F(2, 66) = .507, p < .605, \eta^2 = .015$, suggesting that there were no significant differences in stuttering scores across the three conditions. In addition, there was no significant interaction effect between time and condition, $F(2, 66) = 2.89, p < .062, \eta^2 = .081$. These findings confirm that groups’ attitudes toward PWS did not differ before the intervention.

Similarly, a two-way mixed ANOVA was conducted to examine the effect of time (pre-test and post-test) and condition (R, RS and T) on the Belief and Self-reaction sub-scores. Results revealed a significant main effect of time for both subscores; Belief subscore: $F(1, 66) = 36.31, p < .000, \eta^2 = .35$ and Self-reaction subscore: $F(1, 66) = 74.14, p < .000, \eta^2 = .52$, indicating a significant increase in both scores from pre-test to post-test across all conditions (see Table 3 and Figure 9). However, the main effect of condition for both subscores was not significant, Belief subscore: $F(2, 66) = 1.93, p < .15, \eta^2 = .056$ and Self-reaction subscore: $F(2, 66) = 1.44, p < .24, \eta^2 = .042$, suggesting that there were no significant differences in stuttering scores across the three conditions. In addition, for the Belief sub-score, there was no significant interaction effect between time and condition, $F(2, 66) = .78, p < .46, \eta^2 = .023$. However, we found a significant interaction effect between time and condition for self-reaction subscore; $F(2, 66) = 3.45, p < .037, \eta^2 = .095$.

Table 2. Average scores of POSHA-S Beliefs subscore with corresponding components and items before and after the intervention.

POSHA-S Items	R		RS		T	
	Pre	Post	Pre	Post	Pre	Post
Overall Stuttering Score	18.04	36.48***	27.58	38.93***	22.95	37.03***
Beliefs	28.35	44.84**	41.30	53.62**	43.29	54.71*
Caused by	34.06	55.07*	36.96	46.38	40.58	42.03
Genetic inheritance	34.78	52.17	47.83	26.09	34.48	-4.35
Learning or habits ^r	4.35	26.09	0	8.69	17.39	-8.69
A frightening event ^r	-8.69	30.43	-8.60	34.78	-4.35	8.69
An act of God ^r	52.17	69.57	60.87	69.57	60.87	86.96
A virus or disease ^r	39.13	60.87	43.48	52.17	39.13	69.57
Ghost, demons, spirits ^r	82.60	91.30	78.26	86.96	86.96	100
Helped by	15.94	21.74	14.49	24.64	30.43	20.29
SLP	91.30	91.30	100	100	95.65	82.61
Other PWS	34.78	52.17	0	0	26.09	8.69
Medical doctor ^r	-78.26	-78.26	-56.52	-26.08	-30.43	-30.43
Potential	70.65	70.65	80.43	86.96	76.09	78.26
Can make friends	91.30	95.65	95.65	100	95.65	95.65
Can lead normal lives	95.65	86.96	91.30	100	82.60	91.30
Can do any job they want	34.78	34.78	82.60	91.30	78.26	78.26
Should have jobs requiring good judgment	60.87	65.21	52.17	56.52	47.82	47.82
Traits/Personality	-7.25	31.88**	33.33	56.52*	26.09	78.26**
Nervous or excitable ^r	-34.78	17.30	8.69	56.52	17.39	78.26
Shy or fearful ^r	-43.48	-4.35	13.04	39.13	-4.35	69.57
Have themselves to blame ^r	56.52	82.61	78.26	73.91	65.22	86.96

* indicates significant difference between pre and post scores. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

^r inverted scoring such that positive scores reflect more positive attitude

8.2.2 Effects of the Technological Intervention on Participants' Attitude towards PWS. Table 2 shows the pre-and post-intervention average scores for the OSS, the Beliefs score, the Beliefs components and their corresponding items. In Table 3, we display the pre-and post-intervention average scores for the Self Reactions subscore, components and their items. To test for statistical significance difference between pre-and post-intervention for each condition, we applied two-sided paired

t-tests/Wilcoxon Signed-Rank test to the overall scores and subscores. We applied the Bonferroni correction of $p \leq .00416$ to correct for inflated Type 1 errors. We divided the alpha level of .05 by 12, the mean number of items per prompt [53]. The results are as follows:

Table 3. Average scores of POSHA-S Self Reactions subscore with corresponding components and items before and after the intervention.

POSHA-S Items	R		RS		T	
	Pre	Post	Pre	Post	Pre	Post
Self Reactions	7.72	28.12***	13.85	24.24**	2.59	19.36***
Social Distance/Sympathy	14.94	38.65***	28.44	45.53***	12.02	37.44***
Feel comfortable or relaxed	34.78	82.61	73.91	100	52.17	95.65
Feel pity ^r	-4.35	52.17	13.04	60.87	-13.04	47.83
Feel impatient ^r	60.87	86.96	65.22	95.65	60.87	86.96
Concerned about my doctor ^r	26.09	60.87	65.22	78.26	17.39	60.87
Concerned about my neighbor ^r	86.96	91.30	86.96	86.96	69.57	91.30
Concerned about brother or sister ^r	-39.13	13.04	-17.39	4.35	-39.13	0
Concerned about me ^r	39.13	13.04	17.39	-4.35	39.13	0
Impression of PWS	6.82	28.26	7.5	36.96	-6.82	10.87
Want to have stuttering	-78.26	-54.35	-68.18	-54.55	-75	-56.52
Knowledge/Experience	-9.28	3.26	-10.72	-8.55	-5.39	3.91
Amount known about stuttering	-59.09	-15.22	-58.69	-39.13	-47.73	-21.74
PWS known	4.57	4.76	4.78	5.0	5.43	4.75
Personal experience (me, my family, friends)	21.74	17.39	21.74	13.04	21.74	21.74
Knowledge Source	-13.91	-2.61	-8.69	-5.22	-29.57	-26.96
Television, radio, films	56.52	52.17	39.13	26.09	8.69	17.39
Magazines, newspapers, books	-52.17	-30.43	-69.57	-60.87	-69.57	-78.26
Internet	39.13	52.17	43.48	43.48	17.39	13.04
School	-30.43	-4.34	21.74	34.78	-26.09	-17.39
Doctors, nurses, other specialists	-82.61	-82.61	-78.26	-69.57	-78.26	-69.57
Accommodating/Helping	39.13	73.19**	46.38	65.21**	33.33	63.04**
Try to act like the person was talking normally	65.22	100	91.30	100	86.96	91.30
Helped by person like me	-26.09	8.69	-43.48	-21.74	-65.22	-34.78
Fill in person's word ^r	0	86.96	39.13	86.96	30.43	86.96
Tell person "slow down" or "relax" ^r	34.78	47.83	21.74	43.48	34.78	56.52
Make joke about stuttering ^r	91.30	100	95.65	100	73.91	100
Should try to hide their stuttering ^r	69.57	95.65	73.91	82.61	39.13	78.26

* indicates significant difference between pre and post scores. * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

^r inverted scoring such that positive scores reflect more positive attitude

- (1) **Overall Stuttering score:** A two-sided paired t-test revealed a statistically significant difference between pre and post-OSS scores for condition R ($t(22) = 6.3914$, $p = 0.00005$, $d = .7571$) and RS ($t(22) = 4.964$, $p = .00005$, $d = 1.1458$). Applying a Wilcoxon Signed-Rank test to the OSS also showed a significant difference for the T condition ($V = 255$, $p = .0001$, $r = .742$).

- (2) **Beliefs subscore:** A Wilcoxon Signed-Rank test found a significant difference in the Beliefs subscore for the R condition ($V = 198$, $p = 0.0041$, $r = .598$) and for the RS condition ($V = 224.5$, $p = .0015$, $r = .663$) (see Figure 9). In contrast, no significant difference was found for the T condition ($V = 203.5$, $p = .0195$, $r = .521$), given the adjusted p-value (.0041).
- (3) **Self Reactions subscore:** Figure 9 presents the results for the Self Reactions subscores. According to a two-sided paired t-test, the scores significantly differ in both robot conditions, R ($t(22) = 5.6669$, $p = .00001$, $d = 0.9564$) and RS ($t(22) = 3.2535$, $p = .0036$, $d = 0.5715$). As well as in the tablet condition ($t(22) = 5.4333$, $p = .00005$, $d = .748$).

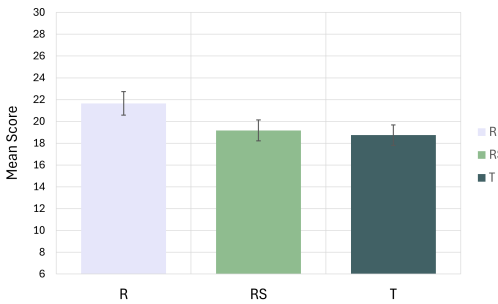


Fig. 10. Average scores for the six empathy adjectives (sympathetic, soft-hearted, warm, compassionate, tender, moved) of the Emotional Response Questionnaire. (Scale range: 6 to 30).

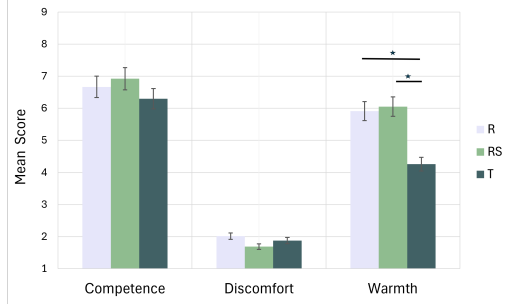


Fig. 11. Average Scores for three RoSAS subscales: Competence, Discomfort and Warmth. Asterisks indicate a significant difference between pre and post-scores. $* = p < .05$ (Dunn's test). (Scale range: 1 to 9).

8.3 Participants Perception of the Intervention's Effectiveness in Teaching about Stuttering

In addition to the improvement in people's attitudes towards stuttering, a significant majority of participants expressed the belief that the intervention effectively helped them learn about stuttering. In the post-questionnaire (see Subsubsection 7.4.4), within the RS condition, 22 individuals (96%) affirmed the effectiveness of the intervention. Similarly, in the R condition, 21 participants (91%) and in the T condition, 20 participants (87%) perceived the intervention as valuable in facilitating their learning about stuttering.

8.4 Empathy

8.4.1 Participant's Level of Empathy. We used the TEQ to measure participant's level of empathy. The average scores were similar in all three conditions (R: $M = 46.22$, $SD = \pm 7.95$; RS: $M = 50.0$, $SD = \pm 6.26$; T: $M = 46.43$, $SD = \pm 7.47$). Applying a one-way ANOVA did not show significant differences in participant's empathy levels between groups ($F(2,66) = 1.967$, $p = .148$). Collectively, the findings suggest that participants demonstrated similar levels of empathy in all three experimental conditions before the intervention.

8.4.2 Empathetic Response to the Video. To assess participants' empathetic response to the video, we summed the scores of six adjectives: *sympathetic*, *soft-hearted*, *warm*, *compassionate*, *tender*, *moved*. The total scores can range from 6 to 30. The mean scores across participants are depicted in Figure 10. To test for differences in the elicited empathy between conditions, a one-way ANOVA was applied. No significant difference between groups was found ($F(2,66) = 1.682$, $p = .194$). Thus, participants in different groups showed, on average, similar empathetic responses to the video.

8.4.3 TEQ and Empathetic Response Correlation. We run a Pearson correlation to investigate the connection between participants' levels of empathy as measured by the TEQ and their empathetic responses to the video presented. Our findings revealed a weak positive correlation ($r = .2529$, $n = 67$, $p = .03602$). Suggesting that individuals with higher TEQ scores also tended to show greater empathetic responses while watching the video.

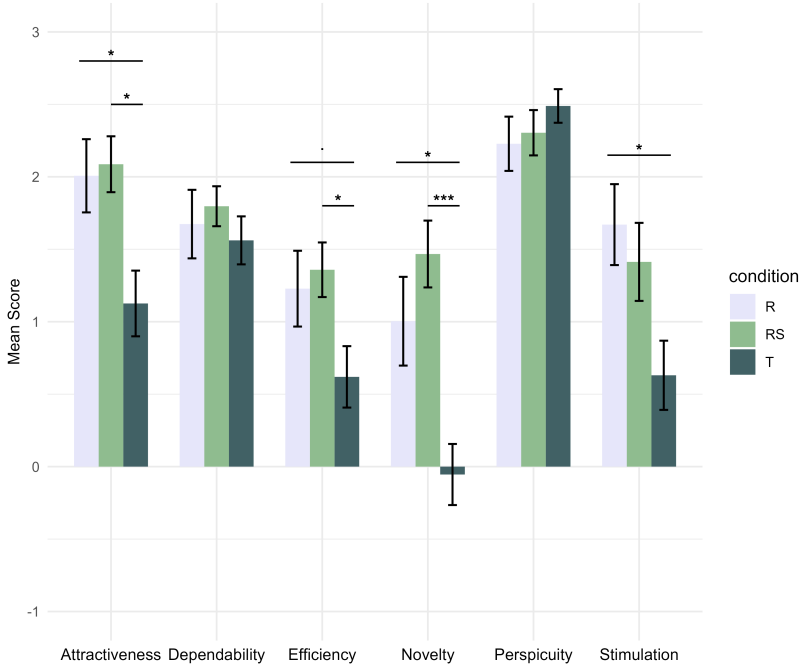


Fig. 12. Average Scores for the five UEQ subscales: Attractiveness, Dependability, Efficiency, Novelty, Perspicuity, Stimulation. Asterisks indicate significant differences between pre and post-scores. * = $p < .05$; ** = $p < .01$; *** = $p < .001$; · = 0.064 (Dunn's test).

8.5 Perception of Technology

8.5.1 RoSAS. We used RoSAS to measure participant's social perception of the technology used. We tested for differences in the three dimensions *Competence*, *Discomfort*, *Warmth*. The scores for all dimensions can range from 1 to 9. First, we took the average score for all dimensions, as shown in Figure 11. All three conditions scored highest for the *Competence* dimension (R: $M = 6.67$, $SD = \pm 1.62$; RS: $M = 6.92$, $SD = \pm 1.78$; T: $M = 6.29$, $SD = \pm 1.59$). They scored second highest for the *Warmth* dimension (R: $M = 5.91$, $SD = \pm 1.83$; RS: $M = 6.05$, $SD = \pm 1.97$; T: $M = 4.26$, $SD = \pm 1.72$), and lowest for the *Discomfort* dimension (R: $M = 2.01$, $SD = \pm 1.87$; RS: $M = 1.69$, $SD = \pm 1.36$; T: $M = 1.88$, $SD = \pm .80$).

We applied a Kruskal-Wallis test to measure for differences between conditions. We found a significant difference in the *Warmth* dimension between conditions ($\chi^2(2) = 11.484$, $p < .005$) with an effect size of $\eta^2 = 0.144$, indicating moderate-large effect. Pairwise comparisons using Dunn's test indicated that the difference is significant between the R and T condition ($p = .017$), as well

as between RS and T ($p = .0063$) condition. Thus, both robots were rated as being significantly 'warmer' than the tablet.

8.5.2 UEQ. Additionally, we used the User Experience Questionnaire to measure participant's experience with the technologies used. We tested for significant differences between conditions for the six scales: *Attractiveness*, *Dependability*, *Efficiency*, *Novelty*, *Perspicuity*, and *Stimulation*.

A Kruskal-Wallis test detected significant differences between conditions in four UEQ dimensions: First, significant results were found for *Attractiveness* ($\chi^2(2) = 11.547$, $p = .00311$). Eta square measure found a moderate-large effect ($\eta^2 = 0.1446$). A post-hoc Dunn's test showed that differences were significant between R and T condition $p = .01024$, as well as between RS and T $p = .00931$ condition. Thus, the robot was rated as being more attractive than the tablet. Second, we also detected significant results for the *Efficiency* subscale ($\chi^2(2) = 7.447$, $p = .02415$) with $\eta^2 = 0.0825$. Pairwise comparison, using Dunn's test indicated that differences were significant between RS and T condition ($p = 0.047$). Furthermore, the R and T conditions showed a trend approaching a significant difference ($p = .064$). Third, *Stimulation* showed significant differences ($\chi^2(2) = 8.9704$, $p = .01127$). A moderate effect size was measured ($\eta^2 = 0.1056$). Dunn's test indicated significant differences between R and T conditions ($p = .011$). The last one, *Novelty*, showed significant results ($\chi^2(2) = 16.438$, $p = .0002695$) with $\eta^2 = 0.2188$ (large effect) between R and T condition ($p = 0.009$), and between RS and T ($p = 0.0003$) condition when applying a Dunn's test.

8.6 Thematic Analysis: Interview Data

We performed a Thematic Analysis (TA) on our interview data, to identify prevalent feedback from our participants. Due to the different nature of the interaction between the robot conditions and the tablet condition, we analyzed them separately. Leading to slightly different themes. In the following sections, we present the results of the thematic analysis of the interview data. The themes and codes will be further described and supported by representative quotes.

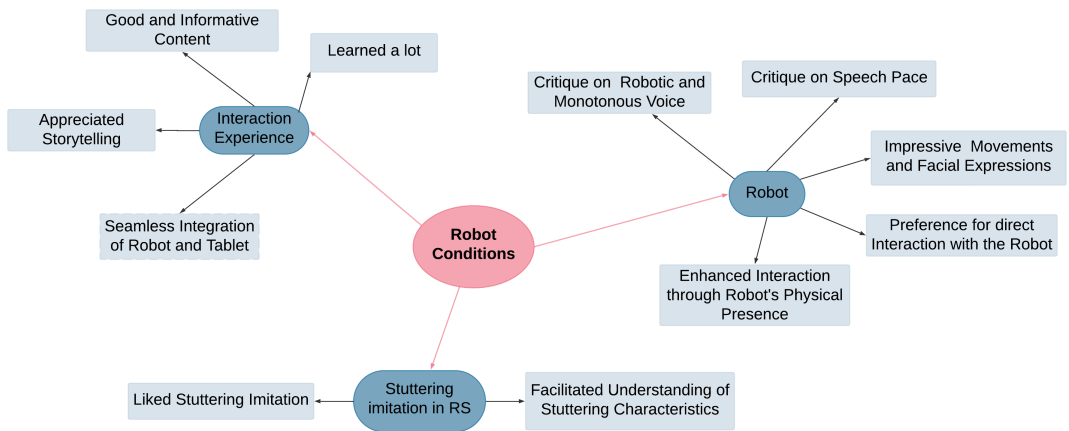


Fig. 13. Results of Thematic Analysis for the interview data of both robot conditions. Themes are shown in dark blue, codes in light blue.

8.6.1 Robot Conditions. We identified two distinct themes within the interview data for both robot conditions. An additional theme emerged specifically for the RS condition, revolving around the stuttering imitation. Figure 13 shows these themes along with their corresponding codes.

- (1) **Theme 1 - Interaction Experience:** This theme comprises four codes, all capturing various aspects of participants' experience. These codes include feedback on the content, their learning outcomes, and the use of technology.
 - (a) **Good and Informative Content:** 19 participants expressed satisfaction with the content and found the shared knowledge about stuttering to be informative.
"It provides you with knowledge about stuttering, the knowledge it provides it knowledgeable."
 - (b) **Learned a lot:** 13 participants commented on the learning outcomes. They emphasized that they gained significant information during the presentation, especially regarding strategies for interacting with PWS.
"It was a great experience in terms of understanding what should a person's behavior be while interacting with someone with stutters."
 - (c) **Appreciated Storytelling:** 8 participants expressed appreciation for the storytelling components and the introduction of the character 'Tim' to share lived experiences, noting that it also increased engagement and understanding.
"[The scenarios] are well designed and they can actually happen in their daily lives. (...). It was really helpful in terms of helping people understand [their] feelings and how our actions can affect them."
 - (d) **Seamless Integration of Robot and Tablet:** 5 comments praised the good alignment of tablet and robot, as well as the responsiveness of the robot to the tablet.
"The robot works well with the tablet."
"I like that the robot was responsive to the tablet."
- (2) **Theme 2 - Robot** The second theme includes five codes that represent participants' feedback regarding the robot.
 - (a) **Impressive Movements and Facial Expressions:** 15 participants provided positive feedback regarding the robot's gestures and facial expressions. They also highlighted that they helped them to maintaining their attention.
"The robot was pretty awesome! I was pretty impressed by how his movements were". "The robot's movements were good, and make you more attentive towards what the robot is speaking".
 - (b) **Critique on Robot's Speech Pace:** 12 participants expressed concerns regarding the slow pace of the robot's speech. They expressed a preference for faster speech.
"When I am listening to audios with robotic voices, I listen to it in times two speed."
"I lost interest at some points because it was so slow."
 - (c) **Critique on Robot's Robotic and Monotonous Voice** 5 participants found the voice to be "too monotonous" and "robotic".
"The voice was robotic, it would be nice if it would be more human like."
"It could have been more interesting, the monotone tone of speech was boring."
 - (d) **Enhanced Interaction Through Robot's Physical Presence:** 19 participants highlighted that they liked the robot as it added engagement to the pure educational content.
"The whole perception of the presentation was better because of the robot".
"The content may be kind of boring but the robot with its movements made it more interesting."
"Feeling the robot's presence was better than just the tablet, for attending more".

- (e) **Preference for Direct Interaction with the Robot:** According to 5 participants expressed desire to verbally interacted with the robot instead of giving their answers via button press on the tablet.
- "I wish the robot could verbally understand users"*
- "It would have been more fun if I could directly interact with the robot."*
- (3) **Theme 3 - Stuttering Imitation in RS** This theme specifically relates to the robot's imitation of stuttering in condition RS.
- (a) **Liked Stuttering Imitation:** 12 participants responded positively to the imitation, finding it natural and providing a "real life experience".
- "It was like talking to a person."*
- (b) **Facilitated Understanding of Stuttering Characteristics:** 8 participants mentioned that the imitation helped them understand the three characteristics (repetition, prolongation, blocks). This was particularly helpful for people that do not have prior experience with stuttering.
- "I did not know [anything] about stuttering, so it was good that [the robot] showed it to me."*
- "I did not know the three characteristics of stuttering and the robot imitated all of them. That was good!"*
- "I clearly understood the three different types of stuttering".*

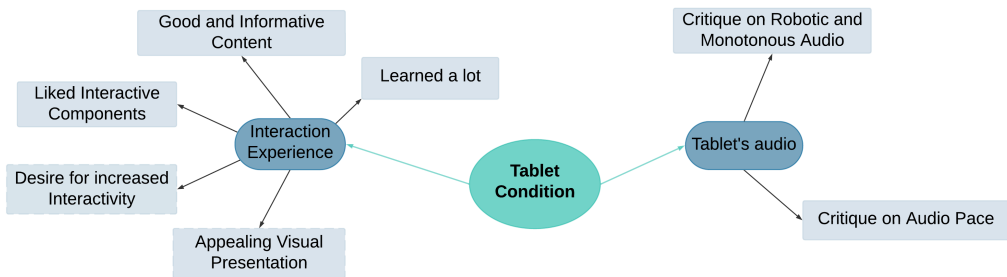


Fig. 14. Results of the Thematic Analysis for the interview data of the tablet condition. Themes are shown in dark blue, codes in light blue.

8.6.2 Tablet Condition. We determined two themes in the tablet condition. The themes are similar to the robot conditions, however they slightly differ in their codes. The themes and corresponding codes are presented in Figure 14.

- (1) **Theme 1 - Interaction Experience** Similar to the robot conditions, this theme captures different aspects of the overall interaction experience. It includes five codes.
- (a) **Code 1 - Good and Informative Content:** As in the robot conditions, participants liked the content of the presentation. Fifteen participants perceived it as interesting.
- "It was informative, now I know what [stuttering] is."*
- (b) **Code 2 - Learned a lot:** Five participants appreciated that they learned a lot through the presentation delivered by the tablet. 2 of them additionally mentioned that it changed their perception of PWS.
- "I accomplished my goal of learning how to interact with people who stutter."*

- (c) **Code 3 - Appealing Visual Presentation:** Ten participants praised the visual presentation in the App, including the presentation of information as well as the graphics supporting the storytelling parts.
"I liked the visuals. It was very statically pleasing, which I enjoyed, like all the graphics. It made it easy to pay attention to."
 - (d) **Code 4 - Liked Interactive Components:** Eight participants mentioned that they liked the interactive parts, like quizzes and Pen&Paper exercise.
"The questions were good for a little bit of interaction. I think it was a good balance for the length of presentation."
"I liked the written questions to put my ideas out on the page."
 - (e) **Code 5 - Desire for increased Interactivity:** Eleven participants suggested increasing the amount of interaction during the presentation.
"I was expecting a couple of more questions."
- (2) **Theme 2 - Tablet's audio:**
- (a) **Code 1 - Critique on Audio Pace:** Similar to the robot condition, 21 participants also perceived the tablet's audio as too slow and their preference to listen to faster audio.
I think I don't have the attention span, especially with social media (...). Taking information is usually at faster pace, but it didn't loose my attention."
 - (b) **Code 2 - Critique on Robotic and Monotonous Audio:** According to five participants, the voice was too robotic and monotonous.

8.7 Thematic Analysis: Pen & Paper Exercise

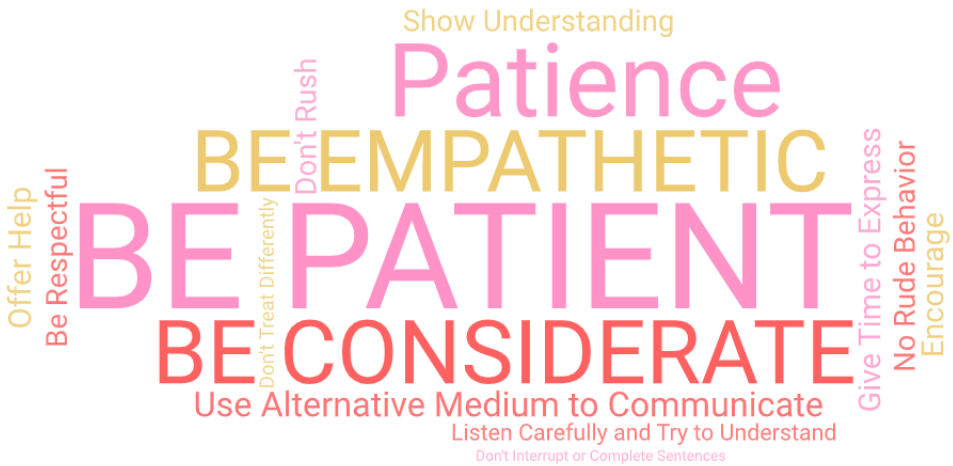


Fig. 15. Thematic Analysis results of “Pen and Paper” exercise. Themes are shown in capital letters, the corresponding codes belong to the same color-coding. The bigger the font size, the more participants commented.

We employed Thematic Analysis (TA) to determine which behaviors towards PWS were suggested by participants in the “Pen & Paper” exercise. 68 out of 69 participants engaged in the exercise. Our

TA revealed three themes. These themes, along with their corresponding codes, are depicted in the word cloud in Figure 15. Further insights are provided in the following:

- (1) **Theme 1 - Be Patient:** The first theme combines five codes, all capturing aspects of a patient attitude.
 - (a) **Give Time to Express:** Thirty participants emphasized the importance of giving PWS the necessary time to express their thoughts.
"The vendor and people at the pizza place could have waited until Tim finishes ordering since stuttering problem can happen with anyone."
"[People at the pizza place] could have given him time which indirectly gives him confidence to speak."
 - (b) **Don't Rush** Five participants stressed the importance of refraining from rushing PWS.
"Pressuring him to talk does not help. I believe it may be more efficient for Tim to speak in front of a patient person. Tim would feel less intimidated."
 - (c) **Don't Interrupt or Complete Sentences:** Thirteen participants suggested to avoid the interruption or completion of sentences for individuals who stutter.
"I feel like the best approach was to let Tim finish his sentence rather than cutting him off."
 - (d) **Patience:** The majority (50 participants) stressed the general importance of patience.
"I believe that people at the pizza place should have handled Tim's stuttering with patience and kindness."
- (2) **Theme 2 - Be Considerate:** This theme encompasses four codes, reflecting descriptions of considerate behavior when encountering PWS.
 - (a) **No Rude Behavior:** According to seventeen participants, people should refrain from rude behavior, both verbal and non-verbal.
"People should avoid negative comments like 'Maybe his opinion is not important' or 'Did he forget to speak.'"
"Not yelling at Tim / not making derogatory comments. These hurtful comments made Tim feel embarrassed for something that he was born with."
 - (b) **Listen Carefully and Try to Understand:** Thirty participants said that people should listen carefully and try to understand the message PWS are trying to convey.
"Tim's fellow students could have taken the time to understand his idea, instead of dismissing them, they could have put in more effort in making him feel more included in the group project."
 - (c) **Use Alternative Medium to Communicate:** Thirty-two participants further suggested the use of alternative communication methods, such as written communication instead of spoken.
"The vendor could ask Tim to write down the toppings he want."
"They can try to give alternative way of solving communication problem - texting or writing note! But it should be polite ask about it!"
 - (d) **Be Respectful:** According to sixteen participants, people should treat PWS respectfully.
"Fellow students should listen to Tim's talking more respectfully."
- (3) **Theme 3 - Be Empathetic:** Five codes are summarized in this theme.
 - (a) **Offer Help:** Nine participants commented that people should offer help if they experience a PWS is having a hard time.

“They could ask him if he needs help first.”

“As for the students, I feel that it was very disrespectful of them to ignore him. They should have listened to his ideas and helped him to convey his thoughts.”

- (b) **Encourage:** Fourteen participants suggested to encourage PWS to speak when they realize they have a hard time getting their message across.

“As a group, no matter what Tim goes through, they should respect him, and encourage Tim to speak more.”

“People should have been more encouraging to Tim, that would boost the confidence of Tim (...).”

- (c) **Show Understanding:** Eighteen participants mentioned that people should show understanding and empathy for the situation of PWS, recognizing the challenges they face.

“They could have been more understanding of his stutter and communication issues.”

“Understanding that stuttering is not a choice and have more patience and empathy to those who stutter.”

- (d) **Don’t Treat Differently:** According to ten participants, individuals who stutter should not be treated differently.

“Tim’s fellow classmates should have understood what Tim was suggesting and not treat him any less intelligent as clearly his idea was good.”

9 DISCUSSION

In our study, we developed two autonomous technical systems, one comprised of a social robot integrated with a tablet and the other one only a tablet. Through a co-design approach, we developed stuttering awareness content and integrated it into the systems. We aimed to investigate the effect of these technologies on changing people’s attitudes towards PWS. Additionally, we compared the two robot conditions. In one condition, the robot was stuttering, while in the other condition, the robot had fluent speech. Next, to examine the attitude variable, we assessed the impact of these interventions on participants’ empathetic response when watching a video featuring PWS talking. In the following section, we discuss the impact of interventions on participants’ attitudes and empathy towards PWS.

9.1 Impact of Intervention on People’s Attitude Towards Stuttering

9.1.1 Impact of the Social Robot on People’s Attitude towards PWS (RQ1). Addressing RQ1, our results show that the QTrobot positively influences participants’ attitudes toward PWS. In both robot conditions (R and RS), we observed significant increases in the Overall Stuttering Score, as well as improvements in the Beliefs and Self Reactions subscores. For the first time in HRI, our findings show that social robots can be used as an educational tool to successfully teach people about stuttering and help reduce the associated stigma. This opens a new and exciting field in HRI research. Furthermore, across all conditions participants suggested desired behavior like “Be Patient”, “Be Considerate”, and “Be Empathetic” towards PWS. This suggests that beyond delivering information, social robots could potentially not only lead to changes in attitudes but possibly also to behavioural changes towards PWS. Similarly, previous studies have shown that social robots hold potential to raise awareness and promote behavioural change across pressing social issues. For instance, social robots have been designed and employed to promote environmental conservation efforts [26] and to raise awareness about gender inequality among children [59]. Additionally, social robots have been employed to normalize discussions surrounding mental health and reduce

associated stigma [58]. Furthermore, studies have highlighted their role in increasing awareness of both mental and physical health issues, encouraging individuals to adopt healthier lifestyle choices [42, 52].

9.1.2 Comparison of Social Robot vs. Tablet Conditions (RQ2). Both the tablet and robot conditions were equally successful in improving people's attitudes. Thus, the embodiment of the robot did not yield any additional impact on the change in people's attitudes. This suggests that the content of the intervention, rather than the mode of delivery, was the key driver for attitude changes. These findings align with research by Gordon et al. [35], which tested children's curiosity levels after interacting with a social robot vs a tablet showing curious behaviour and found no significant difference between the two modalities. Similarly, a study by Zhexenova et al. [83] examined the learning of a Latin-based alphabet among elementary school children across three conditions: a social robot, a tablet-only, and a teacher-led condition. Results showed that children acquired comparable knowledge in all three groups, further emphasizing that content quality, rather than mode of interaction, is often the primary driver of learning and attitudinal outcomes.

A distinct pattern emerged when it comes to the social perception of technology. In both robot conditions, participants consistently perceived the robot as significantly more "warm" than the tablet condition, as measured with RoSAS. Furthermore, in the UEQ, participants rated the robots as more "attractive", "efficient", "novel," and "stimulating" when compared to the tablet. Our Thematic Analysis further revealed that participants considered the "Robot's Physical Presence", as well as its ability to execute gestures and facial expressions, as a notable addition to the whole educational experience. These results suggest that while embodiment may not directly influence attitude change in this context, it plays a critical role in enhancing engagement and overall user satisfaction. Moreover, these results align with prior HRI studies where social robots were compared to tablets in various contexts, like smart-home [28], educational [83], or healthcare settings. Findings from these studies suggest that a social robot's physical embodiment positively influences perceptions of sociability and creativity, making them appear more engaging and personable than tablets [28, 51, 77]. Lastly, these findings highlight a promising avenue for future research that will be discussed in greater detail in Section 8.

9.1.3 Comparison of Robot Conditions (RQ3). We did not find statistically significant differences in stuttering scores between our two robot conditions. The results suggest that exposure to stuttering imitation did not lead to a significantly greater improvement in participants' attitudes towards stuttering. However, participants in the RS condition reported that the imitation provided richer contextual information (see Subsection 8.6.1).

Furthermore, no differences were observed in the social perception of the robot between RS and R, as measured by RoSAS. This indicates that the stuttering imitation was neither distracting nor did it negatively influence participants' perception of the robot. Additionally, it had no detrimental impact on attitude outcomes, as POSHA-S scores did not show a significant difference between RS and R. Furthermore, although the difference was not significant, the stuttering robot scored higher in the UEQ "Novelty" dimension. This observation could hold value for future research, particularly in exploring how robotic stuttering imitation impacts the educational experience and whether further improvements in imitation may lead to more positive attitude outcomes.

9.1.4 Impact of the Educational Content about Stuttering. Participants' attitudes towards stuttering improved significantly across all conditions. Indicating that our content, successfully provided participants with accurate information about stuttering. This finding is consistent with previous research, which highlights the overall effectiveness of educating individuals about stuttering and shows that targeted interventions can positively change attitudes [1, 33, 62].

Our Thematic Analysis revealed that participants valued the storytelling elements which incorporated lived experiences. This finding supports observations made in previous studies by Flynn and Louis [33] and Boyle et al. [11], both of which highlighted the positive effect of including personal experiences in educational sessions. However, unlike these studies that involved actual PWS delivering live presentations or engaging in direct interactions, we used a fictional character to convey lived experiences through storytelling. In the past, storytelling has proven to be an effective tool for shifting perspectives [36]. This approach could be further explored in the field of stuttering awareness.

9.2 Impact of Interventions on Participants' Empathy (RQ4 & RQ5)

Despite the emergence of the “be empathetic” theme in the Thematic Analysis of the “Pen and Paper” exercise, indicating that several participants felt empathy towards Tim, we did not observe any significant differences in the empathetic response between the three conditions after participants watched the video showing PWS. This shows that the type of intervention did not have a differential impact on participants' empathy toward PWS, addressing both RQ4 and RQ5. We speculate that one of the reasons for this result could be that the video featuring real PWS was far more emotionally charged when compared to the presentation's content. Therefore, the intervention did not affect participants' elicited empathy. Another reason may be that, despite the presence of real PWS in the video, all of them displayed a strong sense of confidence, even in the face of severe stuttering. Consequently, the interventions may not have generated higher levels of empathy among participants.

10 LIMITATIONS AND FUTURE WORK

This study represents the first exploration of using social robots in stuttering education. In addition to the significant results achieved in our study, it also paves the way for future research.

We conducted the study within a university environment, with the majority of participants being students. This limits our sample diversity in terms of age and educational background. Future studies could benefit from conducting the experiment with a more diverse population. This approach would allow for testing the robot's potential on a broader target group and determining if the robot's effect may differ between different demographic groups.

Moreover, the study only assessed attitude changes immediately after the intervention. We did not measure whether the education had long-lasting effects on people's perception of PWS. In a follow-up study, St. Louis and Flynn [72] identified that positive attitude changes remained consistent over seven years, with only minor declines, in a group of high school students who had received education about the topic. Future research could investigate whether the robot also has enduring effects on individuals' attitudes. Additionally, exploring potential long-term differences between technologies and the impact of stuttering imitation may be worthwhile.

In our study, participants did not perceive the stuttering imitation as disturbing. Future studies may want to investigate this further. It may be interesting for the robot to simulate a ‘first-person perspective’ as a PWS. The robot could take the role of a PWS and share personal experiences. This approach may be interesting for two reasons. First, it would enable testing if the first-person perspective is more effective in changing people's attitudes compared to the narrator's role in the current study. Second, it would allow for a comparison with studies that involve actual PWS. Related to this, it is worth noting that stuttering was only imitated at some specific points during the presentation. Future research may consider increasing the amount of stuttering to more rigorously assess its impact and create a clearer distinction between stuttering and non-stuttering conditions.

Regarding the technology used, participants perceived the robotic voice as too monotonous. Future studies may consider using more expressive robot voices to enhance engagement, as voice can impact the perception of a robot McGinn and Torre [60].

Also, the content of the presentation successfully educated people about stuttering and improved their perception. The content included different components: general educational facts about stuttering, tips on appropriate behavior, as well as storytelling parts. It remains open to determine if some content components are more effective in educating about stuttering. This may be interesting for future work.

Finally, future work, on a related but separate line of research, can investigate how to develop robot-assisted therapeutic application for children who stutter, which our research group has been working on, see [20, 21, 38].

11 CONCLUSION

Stuttering is a highly stigmatized disorder in many societies. Increasing awareness about it is essential to create a more inclusive and supportive environment for individuals who stutter. Robots have been used for the past 25+ years in many assistive and educational applications, e.g. for children with autism spectrum disorder, showing promising results. However, people who stutter, and how society perceives them, has been overlooked to date in social robot applications. The aim of this study was to investigate this novel application area for using social robots that could benefit PWS, in terms of education others about the condition. We have demonstrated, for the first time, that a social robot can effectively improve people's attitudes towards individuals who stutter. We showed that while a robot might not be better in delivering educational content, compared to an app running on a tablet, participants showed more positive attitudes towards the robot conditions. This implies that indeed, we could have non-robot based educational interventions who can deliver successfully similar educational content, but, in our study, social robot interventions were perceived more positively. Educating people about stuttering also requires them to have an interest in, and enjoyment of the interaction, which might impact how meaningful this intervention is for them in their future interactions with PWS. Thus, our findings open up a promising, novel area of inquiry within the domain of stuttering education, whose potential needs to be further investigated in the future. Research into robot-assisted therapy for children with autism, started 2 decades ago, is now being focused on in numerous research labs world wide. However, we are now on the verge of diversifying research, tackling new frontiers in assistive technology applications.

ACKNOWLEDGMENTS

This research was undertaken, in part, thanks to funding from the Canada 150 Research Chairs Program. The work was also partially supported by a PROMOS ("Programme to increase the Mobility of Students from German Universities") scholarship.

In addition we would like to thank our two collaborators, Shrinjay Mukherjee and Alexandria Scenna, for their valuable support and dedication during the project.

REFERENCES

- [1] Fauzia Abdalla and Kenneth O St. Louis. 2014. Modifying attitudes of Arab school teachers toward stuttering. *Language, Speech, and Hearing Services in Schools* 45, 1 (2014), 14–25.
- [2] Fauzia A Abdalla and Kenneth O St Louis. 2012. Arab school teachers' knowledge, beliefs and reactions regarding stuttering. *Journal of Fluency Disorders* 37, 1 (2012), 54–69.
- [3] Hayley S Arnold, Jian Li, and Kathryn Goltl. 2015. Beliefs of teachers versus non-teachers about people who stutter. *Journal of Fluency Disorders* 43 (2015), 28–39.
- [4] C Daniel Batson, Karen O'Quin, Jim Fultz, Mary Vanderplas, and Alice M Isen. 1983. Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of personality and social psychology* 45, 3 (1983),

706.

- [5] C Daniel Batson, Marina P Polycarpou, Eddie Harmon-Jones, Heidi J Imhoff, Erin C Mitchener, Lori L Bednar, Tricia R Klein, and Lori Highberger. 1997. Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of personality and social psychology* 72, 1 (1997), 105.
- [6] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science robotics* 3, 21 (2018), eaat5954.
- [7] Oliver Bloodstein, Nan Bernstein Ratner, and Shelley B Brundage. 2021. *A handbook on stuttering*. Plural Publishing.
- [8] Michael P Boyle. 2017. Personal perceptions and perceived public opinion about stuttering in the United States: Implications for anti-stigma campaigns. *American Journal of Speech-Language Pathology* 26, 3 (2017), 921–938.
- [9] Michael P Boyle and Gordon W Blood. 2015. Stigma and stuttering: Conceptualizations, applications, and coping. In *Stuttering meets stereotype, stigma, and discrimination: An overview of attitude research*. West Virginia University Press, 43–70.
- [10] Michael P Boyle, Madeline R Cheyne, and Amy L Rosen. 2023. Self-Stigma of Stuttering: Implications for Communicative Participation and Mental Health. *Journal of Speech, Language, and Hearing Research* (2023), 1–18.
- [11] Michael P Boyle, Lauren Dioguardi, and Julie E Pate. 2016. A comparison of three strategies for reducing the public stigma associated with stuttering. *Journal of fluency disorders* 50 (2016), 44–58.
- [12] Michael P Boyle, Lauren Dioguardi, and Julie E Pate. 2017. Key elements in contact, education, and protest based anti-stigma programs for stuttering. *Speech, Language and Hearing* 20, 4 (2017), 232–240.
- [13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77.
- [14] Patrick M Briley, Hope Gerlach, and Molly M Jacobs. 2021. Relationships between stuttering, depression, and suicidal ideation in young adults: Accounting for gender differences. *Journal of Fluency Disorders* 67 (2021), 105820.
- [15] Joost Broekens, Marcel Heerink, Henk Rosendal, et al. 2009. Assistive social robots in elderly care: a review. *Geron-technology* 8, 2 (2009), 94–103.
- [16] Tim Brown et al. 2008. Design thinking. *Harvard business review* 86, 6 (2008), 84.
- [17] Christian Büchel and Martin Sommer. 2004. What causes stuttering? *PLoS biology* 2, 2 (2004), e46.
- [18] Ayhan Çağlayan and Ramazan Sertan Özdemir. 2022. The attitudes of employers towards stuttering and people who stutter: Descriptive research. *Türkiye Klinikleri Journal of Health Sciences* (2022).
- [19] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The robotic social attributes scale (RoSAS) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. 254–262.
- [20] Shruti Chandra, Garima Gupta, Torrey Loucks, and Kerstin Dautenhahn. 2022. Opportunities for social robots in the stuttering clinic: A review and proposed scenarios. *Paladyn, Journal of Behavioral Robotics* 13, 1 (2022), 23–44.
- [21] Shruti Chandra, Torrey Loucks, Gerardo Chavez Castaneda, and Kerstin Dautenhahn. 2022. Social Robots in the Stuttering Clinic: A Human-Centred Exploration with Speech Language Pathologists. In *International Conference on Social Robotics*. Springer, 299–313.
- [22] Anita L Cloete. 2017. Technology and education: Challenges and opportunities. *HTS: Theological Studies* 73, 3 (2017), 1–7.
- [23] Daniel Costa and Robert Kroll. 2000. Stuttering: an update for physicians. *Cmaj* 162, 13 (2000), 1849–1855.
- [24] Ashley Craig, Yvonne Tran, and Magali Craig. 2003. Stereotypes towards stuttering for those who have never had direct contact with people who stutter: A randomized and stratified study. *Perceptual and motor skills* 97, 1 (2003), 235–245.
- [25] Stephen Davis, Peter Howell, and Frances Cooke. 2002. Sociodynamic relationships between children who stutter and their non-stuttering classmates. *Journal of Child Psychology and Psychiatry* 43, 7 (2002), 939–947.
- [26] Berardina De Carolis, Francesca D’Errico, Nicola Macchiarulo, and Veronica Rossano. 2019. Investigating the social robots’ role in improving children attitudes toward recycling. The case of peppercycle. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 301–306.
- [27] Jan de Wit, Arold Brandse, Emiel Krahmer, and Paul Vogt. 2020. Varied human-like gestures for social robots: Investigating the effects on children’s engagement and language learning. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 359–367.
- [28] Andrea Deublein and Birgit Lugin. 2020. (Expressive) social robot or tablet?—on the benefits of embodiment and non-verbal expressivity of the interface for a smart environment. In *Persuasive Technology. Designing for Future Change: 15th International Conference on Persuasive Technology, PERSUASIVE 2020, Aalborg, Denmark, April 20–23, 2020, Proceedings* 15. Springer, 85–97.
- [29] Michelle Dorsey and R Kim Guenther. 2000. Attitudes of professors and students toward college students who stutter. *Journal of fluency disorders* 25, 1 (2000), 77–83.

- [30] Shane Erickson and Susan Block. 2013. The social and communication impact of stuttering on adolescents and their families. *Journal of fluency disorders* 38, 4 (2013), 311–324.
- [31] Alireza Esfandbod, Zeynab Rokhi, Ali F Meghdari, Alireza Taheri, Minoo Alemi, and Mahdieh Karimi. 2023. Utilizing an emotional robot capable of lip-syncing in robot-assisted speech therapy sessions for children with language disorders. *International Journal of Social Robotics* 15, 2 (2023), 165–183.
- [32] Juan Fasola and Maja J Matrić. 2013. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction* 2, 2 (2013), 3–32.
- [33] Timothy W Flynn and Kenneth O St Louis. 2011. Changing adolescent attitudes toward stuttering. *Journal of fluency disorders* 36, 2 (2011), 110–121.
- [34] Benjamin Gleason and Sam Von Gillern. 2018. Digital citizenship with social media: Participatory practices of teaching and learning in secondary education. *Journal of Educational Technology & Society* 21, 1 (2018), 200–212.
- [35] Goren Gordon, Cynthia Breazeal, and Susan Engel. 2015. Can children catch curiosity from a social robot?. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 91–98.
- [36] Melanie C Green and Timothy C Brock. 2000. The role of transportation in the persuasiveness of public narratives. *Journal of personality and social psychology* 79, 5 (2000), 701.
- [37] Barry Guitar. 2013. *Stuttering: An integrated approach to its nature and treatment*. Lippincott Williams & Wilkins.
- [38] Garima Gupta, Shruti Chandra, Kerstin Dautenhahn, and Torrey Loucks. 2022. Stuttering treatment approaches from the past two decades: Comprehensive survey and review. *Journal of Student Research* 11, 2 (2022).
- [39] Rosemarie Hayhow, Anne Marie Cray, and Pam Enderby. 2002. Stammering and therapy views of people who stammer. *Journal of Fluency disorders* 27, 1 (2002), 1–17.
- [40] Anna Henschel, Guy Laban, and Emily S Cross. 2021. What makes a robot social? a review of social robots from science fiction to a home or hospital near you. *Current Robotics Reports* 2 (2021), 9–19.
- [41] Stephanie Hughes, Rodney Gabel, Farzan Irani, and Adam Schlagheck. 2010. University students' perceptions of the life effects of stuttering. *Journal of communication disorders* 43, 1 (2010), 45–60.
- [42] Matouš Jelínek and Kerstin Fischer. 2021. The role of a social robot in behavior change coaching. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 434–438.
- [43] Lejla Junuzovic-Zunic, Eman Mostafa, and Kenneth O St. Louis. 2023. Comparing public school teachers' attitudes toward stuttering in Bosnia & Herzegovina and Egypt. *Speech, Language and Hearing* (2023), 1–14.
- [44] Lejla Junuzovic-Zunic, Osman Sinanovic, and Blazenka Majic. 2021. Neurogenic stuttering: Etiology, symptomatology, and treatment. *Medical Archives* 75, 6 (2021), 456.
- [45] Lejla Junuzovic-Zunic, Mary E Weidner, Isabella K Reichel, Susanne Cook, Kenneth O St Louis, and Mercedes B Ware. 2015. Effects of fluency disorders coursework on students' stuttering attitudes in two countries. (2015).
- [46] Joseph F Klein and Stephen B Hood. 2004. The impact of stuttering on employment opportunities and job performance. *Journal of fluency disorders* 29, 4 (2004), 255–273.
- [47] Michelle Klompas and Eleanor Ross. 2004. Life experiences of people who stutter, and the perceived impact of stuttering on quality of life: Personal accounts of South African individuals. *Journal of fluency disorders* 29, 4 (2004), 275–305.
- [48] Marilyn Langevin, Ann Packman, and Mark Onslow. 2009. Peer responses to stuttering in the preschool setting. (2009).
- [49] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*. Springer, 63–76.
- [50] Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. 2006. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International journal of human-computer studies* 64, 10 (2006), 962–973.
- [51] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (2015), 23–37.
- [52] Sara L Lopes, Aristides I Ferreira, and Rui Prada. 2023. The use of robots in the workplace: Conclusions from a health promoting intervention using social robots. *International Journal of Social Robotics* 15, 6 (2023), 893–905.
- [53] Kenneth O St Louis. 2012. Research and development on a public attitude instrument for stuttering. *Journal of Communication Disorders* 45, 2 (2012), 129–146.
- [54] Kenneth O St Louis, Hilda Sønsterud, Lejla Junuzović-Žunić, Donatella Tomaioli, Francesca Del Gado, Emilia Caparelli, Mareen Theiling, Cecilie Flobakk, Lise Nesbakken Helmen, Ragnhild R Heitmann, et al. 2016. Public attitudes toward stuttering in Europe: Within-country and between-country comparisons. *Journal of communication disorders* 62 (2016), 115–130.
- [55] Kenneth O St Louis, Katarzyna Węsierska, Aneta Przepiórka, Agata Błachnio, Chelsea Beucher, Fauzia Abdalla, Timothy Flynn, Isabella Reichel, Ann Beste-Guldborg, Lejla Junuzović-Žunić, et al. 2020. Success in changing stuttering attitudes: A retrospective analysis of 29 intervention studies. *Journal of Communication Disorders* 84 (2020), 105972.

- [56] Kenneth O St Louis, Mandy J Williams, Mercedes B Ware, Jacqueline Guendouzi, and Isabella K Reichel. 2014. The public opinion survey of human attributes-stuttering (POSHA-S) and bipolar adjective scale (BAS): aspects of validity. *Journal of Communication Disorders* 50 (2014), 36–50.
- [57] Michal Luria, Guy Hoffman, and Oren Zuckerman. 2017. Comparing social robot, screen and voice interfaces for smart-home control. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 580–628.
- [58] Juan Martínez-Miranda. 2017. Embodied conversational agents for the detection and prevention of suicidal behaviour: current applications and open challenges. *Journal of medical systems* 41, 9 (2017), 135.
- [59] Romain Maure and Barbara Bruno. 2023. Participatory design of a social robot and robot-mediated storytelling activity to raise awareness of gender inequality among children. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 974–981.
- [60] Conor McGinn and Ilaria Torre. 2019. Can you tell the robot by the voice? An exploratory study on the role of voice in the perception of robots. In *2019 14th ACM/IEEE international Conference on human-robot interaction (HRI)*. IEEE, 211–221.
- [61] Christina Moro, Shayne Lin, Goldie Nejat, and Alex Mihailidis. 2019. Social robots and seniors: A comparative study on the influence of dynamic social features on human–robot interaction. *International Journal of Social Robotics* 11 (2019), 5–24.
- [62] Hailey Nelson. 2020. Changing college student attitudes toward people who stutter. (2020).
- [63] Sue O’Brian, Mark Jones, Ann Packman, Ross Menzies, and Mark Onslow. 2011. Stuttering severity and educational attainment. *Journal of fluency disorders* 36, 2 (2011), 86–92.
- [64] Jane E Prasse and George E Kikano. 2008. Stuttering: an overview. *American family physician* 77, 9 (2008), 1271–1276.
- [65] Leon Reicherts, Yvonne Rogers, Licia Capra, Ethan Wood, Tu Dinh Duong, and Neil Sebire. 2022. It’s good to talk: A comparison of using voice versus screen-based interactions for agent-assisted tasks. *ACM Transactions on Computer-Human Interaction* 29, 3 (2022), 1–41.
- [66] Andra Rice, Katarzyna Kłęczek, and Maryam Alimardani. 2023. The Effectiveness of Social Robots in Stress Management Interventions for University Students. In *International Conference on Social Robotics*. Springer, 181–190.
- [67] Vladimir Robles-Bykbaev, Mario Ochoa-Guaraca, Marco Carpio-Moreta, Daniel Pulla-Sánchez, Luis Serpa-Andrade, Martín López-Nores, and Jorge García-Duque. 2016. Robotic assistant for support in speech therapy for children with cerebral palsy. In *2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*. IEEE, 1–6.
- [68] Rinat B Rosenberg-Kima, Yaacov Koren, and Goren Gordon. 2020. Robot-supported collaborative learning (RSCL): Social robots as teaching assistants for higher education small group facilitation. *Frontiers in Robotics and AI* 6 (2020), 148.
- [69] Stela H Seo, Denise Geiskkovitch, Masayuki Nakane, Corey King, and James E Young. 2015. Poor thing! Would you feel sorry for a simulated robot? A comparison of empathy toward a physical and a simulated robot. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. 125–132.
- [70] Micol Spitale, Silvia Silleresi, Franca Garzotto, and Maja J Matarić. 2023. Using socially assistive robots in speech-language therapy for children with language impairments. *International Journal of Social Robotics* 15, 9 (2023), 1525–1542.
- [71] R Nathan Spreng*, Margaret C McKinnon*, Raymond A Mar, and Brian Levine. 2009. The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment* 91, 1 (2009), 62–71.
- [72] Kenneth O St. Louis and Timothy W Flynn. 2018. Maintenance of improved attitudes toward stuttering. *American Journal of Speech-Language Pathology* 27, 2 (2018), 721–736.
- [73] Kenneth O St. Louis, Katarzyna Węsierska, and Irena Polewczyk. 2018. Improving Polish stuttering attitudes: An experimental study of teachers and university students. *American journal of speech-language pathology* 27, 3S (2018), 1195–1210.
- [74] Marc Steen. 2013. Co-design as a process of joint inquiry and imagination. *Design issues* 29, 2 (2013), 16–28.
- [75] Glen Tellis and Kenneth O St Louis. 2015. *Stuttering meets stereotype, stigma, and discrimination: an overview of attitude research*. West Virginia University Press.
- [76] John Van Borsel. 2014. Acquired stuttering: A note on terminology. *Journal of Neurolinguistics* 27, 1 (2014), 41–49.
- [77] Rianne Van den Berghe, Josje Verhagen, Ora Oudgenoeg-Paz, Sanne Van der Ven, and Paul Leseman. 2019. Social robots for language learning: A review. *Review of Educational Research* 89, 2 (2019), 259–295.
- [78] Anne Zappi Williams, John Tetnowski, and Kenneth O St. Louis. 2023. Addressing Attitudes About Stuttering in Preservice Teachers. *Perspectives of the ASHA Special Interest Groups* 8, 2 (2023), 372–379.
- [79] Melonee Wise, Michael Ferguson, Derek King, Eric Diehr, and David Dymesich. 2016. Fetch and freight: Standard platforms for service robot applications. In *Workshop on autonomous mobile service robots*. 1–6.
- [80] Ehud Yairi and Noline Ambrose. 2013. Epidemiology of stuttering: 21st century advances. *Journal of fluency disorders* 38, 2 (2013), 66–87.

- [81] Yuichiro Yoshikawa, Hiroaki Kobayashi, Naomi Sakai, Hiroshi Ishiguro, and Hirokazu Kumazaki. 2024. Therapeutic potential of robots for people who stutter: a preliminary study. *Frontiers in Psychiatry* 15 (2024).
- [82] Aida Zhanatkyzy, Aizada Turarova, Zhansaule Telisheva, Galiya Abylkasymova, and Anara Sandygulova. 2019. Robot-assisted therapy for children with delayed speech development: a pilot study. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–5.
- [83] Zhanel Zhexenova, Aida Amirova, Manshuk Abdikarimova, Kuanysh Kudaibergenov, Nurakhmet Baimakhan, Bolat Tleubayev, Thibault Asselborn, Wafa Johal, Pierre Dillenbourg, Anna CohenMiller, et al. 2020. A comparison of social robot to tablet and teacher in a new script learning context. *Frontiers in Robotics and AI* 7 (2020), 99.