I Can Do It: Exploring Voice Assistants for Adults with Intellectual Disabilities

Madhuka Nadeeshani
Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
madhuka.koralalagedona@monash.edu

Jacqueline Johnstone
Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
jacqueline.johnstone@monash.edu

Kirsten Ellis
Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
kirsten.ellis@monash.edu

Swamy Ananthanarayan
Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
swamy.ananthanarayan@monash.edu

Abstract

Voice Assistants (VAs) are increasingly integrated into smartphones and smart home devices, offering potential support for diverse groups of users. However, limited research has examined how individuals with Intellectual Disabilities (ID) engage with these technologies. We conducted an eight week study with 17 adults with ID and four support workers at a disability support organization, integrating screen based VAs within a STEAM (Science, Technology, Engineering, Arts, Mathematics) program. Data were collected through interviews with participants and support workers before, during, and after deployment, complemented by analysis of VA interaction logs and researcher observations. Participants initiated 260 interactions with the VAs, using them for information retrieval, entertainment, and learning, with peer support playing a critical role in sustaining engagement. Some participants experienced difficulties with pronunciation and cognitive challenges, while several formed emotional connections with the devices. Based on these findings, we propose six design considerations to guide the development of more inclusive VAs.

CCS Concepts

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

Keywords

voice assistants, Google Nest Hub, Amazon Echo Show, intellectual disability, design considerations

ACM Reference Format:
Madhuka Nadeeshani, Jacqueline Johnstone, Kirsten Ellis, and Swamy Ananthanarayan. 2026. I Can Do It: Exploring Voice Assistants for Adults with Intellectual Disabilities. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3772318.3790573

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3790573

1 Introduction

Voice Assistants (VAs) have become ubiquitous in smartphones and smart home devices, allowing users to access information, play media, and control technology through voice commands. They are available in multiple formats, such as standalone smart speakers (e.g., Google Home [34], Amazon Echo [1]), smart displays (e.g., Google Nest Hub [30], Amazon Echo Show [4]), and built-in smartphone assistants (e.g., Siri [69], Google Assistant [8]). These technologies have been widely adopted and extensively studied in mainstream contexts [13, 28, 74]. More recently, VAs have been explored in relation to specific need populations, including older adults [24, 39, 76], individuals who are blind or have low vision [19, 64, 67], and people with physical disabilities [27, 57]. In these contexts, VAs support daily activities, education, and employment [39, 81], and they can also act as social companions, potentially reducing isolation [39]. These opportunities highlight the need to investigate how VAs can effectively support populations with distinct needs, such as individuals with Intellectual Disabilities (ID).

Individuals with ID have difficulties retaining and manipulating information simultaneously, which affects their engagement with digital systems. These working memory demands make it harder to track spellings, follow multi-step instructions, or rephrase misunderstood questions [59]. Limited literacy and reading comprehension can further affect how easily text based or visually dense interfaces can be used, reducing the overall usability of many mainstream technologies [12, 44]. VAs offer an intuitive, voice based interface that can reduce cognitive load and support greater independence [12, 81]. Prior research shows that VAs can assist with accessing information, managing routines, and operating smart devices [12, 50], enhancing autonomy and reducing reliance on caregivers [16]. Many individuals with ID also experience co-occurring physical disabilities [21], making hands free interaction particularly beneficial [57].

Beyond functional support, VAs may serve as non-judgmental communication partners that foster social interaction and reduce isolation [70, 71]. Despite these benefits, research on VAs for individuals with ID remains limited, particularly regarding real world adoption and long term interaction patterns. To our knowledge, no prior work has analysed voice interaction logs from extended VA use, which is essential for understanding authentic usage patterns, persistent challenges, and user adaptation strategies over time. Existing studies often focus on structured tasks [12, 50] (e.g., retrieving information or playing videos via voice commands), whereas our work provided initial training and then allowed participants to freely choose how and when to engage with VAs, supporting open ended and naturalistic use.

The study was conducted within the context of the STEAM (Science, Technology, Engineering, Arts, Mathematics) program at a local disability support organisation, where participants engage in creative, social, and technology focused activities. The program routinely introduces new technologies and encourages hands-on exploration as part of its broader independence building goals, providing a familiar and supportive environment for technology use. In this work, the STEAM program serves as the setting in which VA use was explored; our focus is on how individuals with ID learn to use VAs in this context, rather than on their engagement with or evaluation of the STEAM program itself.

We deployed Google Nest Hub and Amazon Echo Show devices across five sites within a disability support organisation for two months that supports adults with ID, autism spectrum disorder, cerebral palsy, Rett syndrome, and partial trisomy 9p. Each site participated in six hour weekly session, and data collection followed a three phase methodology involving preliminary interviews, deployment interviews, post reflection interviews, researcher observations, and analysis of voice interaction logs. Incorporating perspectives from both individuals with ID and their support workers allowed us to capture interaction patterns, challenges, and forms of assistance that emerged during everyday use of VAs in real world support settings. This study addresses the following research questions:

- How do individuals with ID engage with VAs during creative, social, and technology focused activities?
- What challenges, strategies, and learning patterns emerge as individuals with ID use VAs over an eight weeks period?
- How can VAs be designed to better support the communication and interaction needs of individuals with ID?

This paper contributes to human computer interaction and accessibility research by offering empirical insights into how adults with ID engaged with VAs across a wide range of activities, drawing on 240 hours of multimodal interaction data. It further examines the communication and memory challenges that shaped these interactions, as well as the learning trajectories and repair strategies that supported continued engagement over time, and it contributes a set of design considerations for developing more accessible VA systems tailored to the needs of adults with ID.

Building on these contributions, our study shows that individuals with ID experienced both meaningful engagement and ongoing difficulties when interacting with VAs. Participants gradually became more confident and explored a wide range of activities, while still encountering issues such as unclear pronunciation, ambiguous phrasing, delays in formulating prompts, and challenges recalling wake actions. These observations informed six design considerations focused on mispronunciation tolerance, adaptable listening windows, simplified system responses, mechanisms for sustained engagement and routine building, and multimodal interaction cues. Together, these insights outline how future VAs could better support communication, participation, and autonomy for adults with ID.

2 Related Work

While off-the-shelf VAs are increasingly adopted to support diverse user groups, research on their use among individuals with ID remains limited, highlighting the need for a deeper understanding of their experiences, challenges, and opportunities. The section begins with work on VA use across specific-need groups, followed by research centred on individuals with ID.

2.1 Exploratory use of VA for Specific-Need Groups

VAs have demonstrated significant potential in supporting specific-need groups, including older adults [24, 25, 39, 63, 76], individuals who are blind or have low vision [19, 64], and people with physical disabilities [57, 78]. Previous research has shown that older adults used VAs for various tasks such as accessing information, managing daily activities, controlling smart home devices, and supporting health related routines [36]. Studies [76, 81] highlighted that older adults who used Amazon Alexa and Google Home for daily activities, such as setting alarms, managing schedules, and retrieving information, not only experienced practical benefits but also reported improvements in mood and increased social connectedness. However, Pradhan et al. [63] observed that although older adults initially used VAs for a broad range of tasks, their usage declined over time, with increasing reliance on specific functions, particularly those related to health. A year long study [75] with older adults in a long term care community found that, although participants used non-screen based Alexa to access information, they often preferred screen based devices, such as smartphones, for tasks like information searching, music playback, and checking the weather, due to their more interactive features.

Another study [38] highlighted the use of Alexa for medication reminders among older adults, emphasizing its importance as a valuable tool for supporting health management and assisting individuals in adhering to their medication schedules. Several studies [39, 72] reported that many older adults considered VAs not only as task oriented tools but also as digital companions that provided both functional support and emotional comfort. Despite challenges related to speech intelligibility, older adults largely adopted VAs because they helped overcome common barriers, such as reading small text, navigating complex digital systems, and interacting with traditional graphical user interfaces [76, 81].

For blind individuals, VAs offer essential voice driven solutions that enhance accessibility and support independence. Research [64] indicates that these users frequently rely on VAs for tasks such as listening to music, checking the time, and managing assistive devices. In addition, Herskovitz et al. [33] demonstrated that VAs allow users to control smart devices that were previously inaccessible, further promoting autonomy. VAs have also been shown to improve the quality of life for individuals with physical disabilities [27, 57]. Research on the use of Google Home for home automation finds that users manage their environment more effectively through VAs, reducing reliance on caregivers and supporting greater independence [27, 57]. Despite these benefits, recent work shows that underlying speech recognition is not equally reliable for all users where recognition performance varies across genders, accents, and speech styles, with atypical speech more likely to be misrecognized [48, 49, 64]. Studies with specific-need groups further describe interaction difficulties related to pronunciation [64, 76, 81] and question formulation (e.g., ambiguous queries [64]), as well as challenges recalling wake words [33, 64, 81] and taking longer to formulate queries [64, 76]. Together, these findings suggest that while VAs can reduce access barriers, they may also shift where interaction effort and breakdowns occur, highlighting the need to understand and mitigate these patterns for underserved user groups.

2.2 Clarifying ID Terminology

According to the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) [9], ID is defined as a neurodevelopmental condition involving significant limitations in intellectual functioning and in the adaptive skills needed for communication, social participation, and everyday activities. In our study, the participant group included individuals with ID, autism spectrum disorder, cerebral palsy, Rett syndrome, and partial trisomy 9p. Although autism spectrum disorder and ID are distinct diagnoses, participants in our study who had autism spectrum disorder showed functional support needs comparable to those typically associated with ID. Moreover, our partner organisation uses ID as an umbrella term for this broader group, and we adopt the same terminology in this paper.

At the same time, we acknowledge that terminology varies internationally [20], and in some countries the term "learning disabilities" is used to describe the same population that diagnostic frameworks identify as having ID. This variation in usage can create ambiguity, particularly because learning disabilities is also used to refer to specific learning disabilities such as dyslexia, dysgraphia, and dyscalculia which relate to difficulties in particular academic domains and are conceptually distinct from ID [26]. These specific learning disabilities do not imply an ID; in contrast ID involves broader cognitive and adaptive support needs that extend beyond difficulties in a single academic area.

2.3 Exploratory use of VA for People with ID

VAs are emerging as accessible tools that support individuals with ID in managing daily tasks [12], improving communication [46, 71], and enhancing overall well-being [12, 44]. Although research has highlighted their potential benefits, VAs remain underexplored, with significant gaps in understanding that hinder widespread adoption and effective use.

Two studies [12, 46] have investigated the use of VAs in daily activities for individuals with ID. Balasuriya et al. [12] specifically examined mobile phone based VAs, such as Siri and Google Assistant, for tasks including searching for images and videos, as well as managing calendars. The study found that 72% of participants with ID preferred voice input, citing its ease of use and ability to mitigate spelling issues when typing. However, the study's short duration (three weeks) and focus on mobile phone based assistants may limit its applicability, as standalone devices offer a more hands-free experience that can enhance user engagement and adoption. Lewis et al. [46] examined how individuals with cognitive disabilities and their caregivers used the Amazon Echo Dot and Echo Show over a four week period. This study found that participants were able to establish routines and complete various tasks with caregiver assistance. Moreover, the Echo Dot supported tasks such as retrieving weather reports and answering health-related questions.

Masina et al. [50] investigated whether participants' cognitive and linguistic abilities could predict their performance with Google Home. Participants completed structured tasks such as controlling fan and lights, adjusting light colour, and operating the TV, and their pronunciation skills were subsequently evaluated. While the study provides insight into how these abilities relate to short term VA performance, its single session design prevents understanding of how performance may evolve with repeated or long-term use.

Two studies [71, 77] evaluated the impact of VAs on the well-being of individuals with ID over a two-month period, using well-being assessment scales to evaluate how VA use influenced daily activities, including smart home control. Both studies reported positive outcomes, demonstrating that VAs can enhance well-being by supporting activities like entertainment and promoting independence, thereby highlighting the importance of long-term evaluations [77]. Another study by Smith et al. [70] found that pronunciation clarity among individuals with ID improved after 12 weeks of interacting with a smart speaker, highlighting the benefits of prolonged use. However, they also noted that without detailed tracking of usage frequency and interaction patterns, it is difficult to fully understand how sustained engagement with the devices contributes to communication improvements. While existing research highlights the potential of VAs to support independence, communication, and well-being among individuals with ID, key gaps remain particularly in understanding how to address challenges related to atypical speech such as mispronunciations [12, 50, 71, 77], taking longer to formulate questions [12, 77], ambiguous questions [12] and diverse cognitive difficulties, such as recalling commands and wake words [46, 77]. Building on this foundation, our study aims to deepen understanding of these issues and explore solutions to better support individuals with ID.

3 Methods

This study was conducted in collaboration with a local disability support organization that operates across 20 sites within a single geographic region, supporting over 500 adults with disabilities that affect cognition. The researchers maintain an ongoing engagement with the disability support organization through their involvement in delivering the STEAM program and providing materials for the activities. The support workers deliver programs covering a wide range of interests, skills, and employment areas to help people with ID achieve their goals and participate in meaningful activities.

Our work was conducted in the context of the STEAM program over an eight week period, with weekly six hour sessions held at each site. Early sessions emphasised creative building, including LEGO constructions and animated paper characters, followed by engineering and technology challenges such as designing a hydraulic lift and exploring the Makey Makey [1] (a small electronic board that supports the connection of everyday conductive materials such as foil, fruits, or playdough to a computer using alligator clips, enabling these materials to act as inputs for interactive creations), where participants created projects such as a digital musical guitar. Later sessions incorporated festive themed crafts, interactive system designs (e.g., fish catching devices), science experiments involving pressure and balloons, and Makey Makey projects where participants triggered animal sounds by touching conductive objects.

Ethics approval for the study was obtained from both the university's research ethics committee and the disability support organization to ensure that the research met all necessary ethical standards. To support informed decision making, we provided consent information in three accessible formats: plain language sheets with images, step by step verbal explanations, and a short video offering an easy to understand overview of the study. We also shared project information with guardians and carers to keep them informed about the study. All participants were over 18, and even those who were minimally verbal had sufficient understanding to provide their own informed consent independently. In some cases, to confirm that the information was clearly understood, support coaches asked brief clarification questions, which participants were able to answer successfully.

Participants were informed of the study's goals, procedures, potential benefits, and risks, and were reminded that they could pause activities, withdraw at any time, or request deletion of their data. Participation was supported through ongoing assent, indicated by participants' verbal responses or non-verbal cues, interpreted together with support workers who were familiar with their communication styles. Throughout the study, researchers and support workers monitored these cues to ensure participants remained comfortable, and regularly reiterated that participation was voluntary and that they could stop or withdraw at any point.

3.0.1 Participants : The term *"participants"* refers to adults with ID who are enrolled in the disability support organization's STEAM program (Table 1). The study initially included 17 participants (9 males, 8 females) with a mean age of 25.0 years (SD = 3.7), all of whom had experience using mobile phones. Six participants (P1, P3, P4, P6, P8, P16) had prior experience with VAs (Siri, Google Assistant, Google Nest).

The disability support organisation provided information about whether participants were described as having lower or higher support needs. These terms are part of the organisation's routine practice and refer to the amount of assistance a participant typically requires for example, the level of communication support, prompting, or guidance needed to stay engaged in activities. Lower support needs indicate that a participant can engage with minimal assistance, including following simple instructions, completing short tasks independently, and staying focused with occasional reminders. Individuals with lower support needs typically join group activities supported by one coach for four participants (1:4). In contrast, higher support needs indicate that an individual benefits from more frequent prompting and closer guidance, such as having instructions broken into smaller steps, or needing one-to-one assistance during tasks or group activities. Within our participant group, P8 was the only individual the organisation classified as having higher support needs; all other participants were described as having lower support needs.

All participants except P7 communicated verbally, with abilities ranging from short phrases to full sentences. To capture both verbal and non-verbal behaviours, data collection combined participant interviews with researcher observations. The non-verbal participant (P7) used gestures for communication and was assisted by a peer throughout the study. During the study, participant numbers fluctuated across the three phases, as some transitioned to other programs, secured employment, or discontinued their engagement with the disability support organization.

3.0.2 Coaches : The term *"coaches"* refers to the employed professional support workers from the disability support organization who facilitate the STEAM program across five different sites (Table 2). Two lead coaches, (C1, C2), alternate between sites throughout the week to oversee the program, while the other coaches are stationed at a single site and assist the lead coaches in running the sessions. Three of the four coaches (C1, C3, C4) had prior experience using VAs (Siri, Google Assistant, Google Nest Hub) in their daily activities, including tasks such as entertainment, accessing information, cooking, and connecting to other smart devices, while C2 had no prior experience with VAs.

3.1 Researcher Positionality

The involvement of participants and coaches was central to this study, and it is also important to acknowledge how the researchers' positionality shaped the study design and the interpretation of participants' experiences. All authors maintain an ongoing relationship with the organisation where this study took place, and this continued engagement informed both the design and conduct of the study. The first and second authors regularly volunteered at the organisation prior to the research, attending sessions and supporting coaches and participants during activities. Together, they have spent more than 700 hours working within the organisation's programs.

The last two authors have a long standing collaboration with the disability support organisation, contributing resources, materials, and guidance for inclusive technology initiatives. Their sustained engagement provided valuable contextual understanding of participant routines, support needs, and organisational practices, while also requiring reflexivity regarding how familiarity might shape participant behaviour. Collectively, the research team's expertise in HCI, accessibility, and qualitative fieldwork informed the study's use of observations, and semi-structured interviews. As none of the authors identify as having an ID, we approached data collection and analysis with awareness of this outsider perspective, triangulating observations with insights from coaches and centring participants' own accounts of their interactions.

3.2 Devices Used in the Study

The study used two screen-based VAs: the Google Nest Hub (2nd generation) and the Amazon Echo Show 8 (2nd generation). Both supported multimodal interaction through voice input, visual output, and touch. Participants activated the assistants using wake

[1] https://makeymakey.com/

Table 1: Participant Demographic Information

Participant ID	Site	Age	Gender	Self-Identified Disability	Research Phase Participation
P1	Site 1	29	M	Autism Spectrum	1
P2	Site 1	27	M	Intellectual Disability	1,2
P3	Site 1	21	M	Intellectual Disability & Cerebral Palsy	1,2
P4	Site 1	22	M	Autism Spectrum	1,2,3
P5	Site 1	21	F	Intellectual Disability & Rett syndrome	1,2,3
P6	Site 1	30	F	Intellectual Disability & Autism Spectrum	1,2,3
P7	Site 1	22	M	Autism Spectrum	1,2,3
P8	Site 2	26	F	Intellectual Disability	1,2,3
P9	Site 3	27	M	Intellectual Disability	1
P10	Site 3	24	F	Autism Spectrum & Partial Trisomy 9p	1,2,3
P11	Site 4	23	F	Intellectual Disability	1
P12	Site 4	25	M	Intellectual Disability	1,2,3
P13	Site 5	34	F	Intellectual Disability & Autism Spectrum	1
P14	Site 5	20	M	Autism Spectrum	1
P15	Site 5	22	M	Autism Spectrum	1
P16	Site 5	27	F	Autism Spectrum	1,2,3
P17	Site 5	25	F	Intellectual Disability	1,2,3

Table 2: Coach Demographic Information

Coach ID	Site	Gender	Disability Experience	Research Phase Participation
C1	Site 1–5	F	5+ years	1,2,3
C2	Site 1–5	F	2 years	1,2,3
C3	Site 1	M	5+ years	1,2
C4	Site 1	F	3 years	1

words ("Hey Google" and "Alexa"), and the devices provided spoken responses accompanied by on screen visual feedback (Figure 1). Touch input further enabled participants to navigate menus, follow step-by-step instructions, open links, adjust settings, and control media playback.

Beyond these shared capabilities, the devices differed slightly in their physical design and external features where the Echo Show includes a front facing camera, whereas the Nest Hub does not. However, camera based features such as gesture input or facial recognition were not introduced as part of the study activities, and participant interactions focused on voice and touch. The devices also differed slightly in their physical layout for example, in screen size (seven inches for the Nest Hub and eight inches for the Echo Show) and in the placement of volume buttons (on the side of the Nest Hub and on the top of the Echo Show) but these variations did not appear to shape how participants interacted with them. Across sessions, participants used both devices in comparable ways, relying primarily on the core affordances of voice and touch rather than device specific hardware features or physical controls.

3.3 Study Procedure

We employed a three-phase approach (preliminary interviews, deployment interviews, and post-reflection interviews) to explore the integration of VAs into the program for individuals with ID. In preliminary interviews, we explored participants' prior experiences with VAs, their perceived benefits and challenges, and their interest in using these technologies. Building on participants' interest and motivation, deployment phase commenced with training sessions using the Google Nest Hub at each site. We also incorporated the Amazon Echo Show three weeks into the program. We utilized these screen-based VAs based on feedback from coaches and participants, who noted that visual elements help participants better understand instructions and engage more effectively with the system. Additionally, prior research on specific-need groups also highlights the benefits of screen-based technologies [18, 46], demonstrating improvements in motivation and user engagement. Together, these insights informed our exploration of screen-based VAs within the program.

Both devices (Google Nest Hub, Amazon Echo Show) were placed centrally on tables to facilitate access for small groups of 3-4 participants (Figure 2). On one occasion at site 1, when five participants were present, we suggested introducing a backup Google Home to ease sharing demands; however, participants preferred to continue engaging collectively with a single device. Some participants consistently favoured one type of assistant, whereas others used both

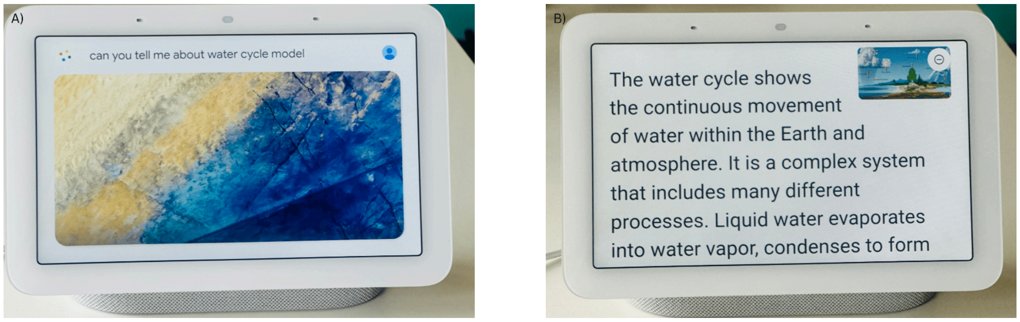

Figure 1: Screenshots of the Google Nest Hub interface: (A) the participant's spoken question is transcribed on the screen; (B) the device's response is displayed as text and spoken aloud.

devices depending on their preference. Coaches also encouraged participants to use the VA freely throughout the program.

Deployment interviews began in the third week and were completed by the end of the fourth week to gather participants' and coaches' insights about the VAs. Additionally, the Google Nest Hub and Amazon Echo Show devices were delivered to each site by the coaches as part of their routine transport of program materials. Post-reflection interviews were conducted after two months of initial use to capture participants' and coaches' feedback and analyse changes in their perceptions over time.

All interviews were semi-structured, with questions designed and refined several times in collaboration with all authors and coaches. Interview questions were phrased in simple language without technical terms, ensuring that participants could understand all questions. All the interviews were conducted by the first author to ensure consistency and foster connections with the participants, with support from coaches. The consistent presence of the first author throughout the study helped to build trust and comfort among participants. Interviews with participants and coaches were conducted separately across all three phases to capture their perspectives independently. However, in some participant interviews (P4, P9), a coach was present to provide communication support when needed, particularly for participants who were less verbally communicative and required assistance to express or clarify their ideas. Field observation notes were recorded, capturing both verbal and non-verbal behaviours, including gestures, facial expressions, vocalizations, and interaction tempo. We also recorded and analysed voice interaction logs from the Google Nest Hub and Amazon Echo Show to understand usage patterns and identify recurring issues or preferences.

3.3.1 Phase 1 - Preliminary Interviews : One-on-one semi-structured interviews were conducted with participants and their coaches separately at their respective sites. The interview questions focused on participants and coaches prior experience with VAs, perceived or anticipated advantages and disadvantages, and their interest in using voice-based devices within the program, including the specific activities they would like to perform. For participants with no prior experience using VAs, we provided verbal explanations and images of the devices, and when further clarification was needed, a demonstration using Siri or Google Assistant was given. For participants, we began with non technical questions such as *"Have you ever communicated through voice commands?"* and *"Do you like using voice commands?"*, before asking whether they had prior experience with VAs. When participants had difficulty recalling device names, visual aids displaying each VA alongside its name were provided to support memory and recall. For coaches, questions directly addressed the use of VAs and their perspectives on which participants they felt would most benefit from integrating VAs. Some participant interviews were brief due to limited verbal abilities, which made it difficult for participants to fully express their ideas. As a result, they often preferred responding with simple *"yes"* or *"no"* answers. Interview durations with participants ranged from 2 to 8 minutes, while those with coaches ranged from 9 to 18 minutes.

3.3.2 Phase 2 - Deployment Interviews : Initially, we conducted one-on-one training sessions with participants over a two week period, totalling ten days across all sites (approximately two days per site). During each session, researchers introduced the wake words along with a small set of voice commands (Table 3) through brief demonstrations, guided practice, and printed cheat sheets. We began with short, familiar queries (e.g., asking about the weather) and gradually introduced more complex commands as participants became more comfortable. These commands served as flexible practice opportunities rather than a fixed sequence: some participants experimented with the full set, while others focused on those that matched their interests. After the training activities, we conducted interviews with both participants and coaches to reflect on their experiences with the VAs. During the second week, P3 assisted the first author by helping demonstrate Google Nest Hub features to other participants. In the following weeks, support was provided as needed to ensure participants felt comfortable using the device and could engage with its features.

After participants became familiar with the Google Nest Hub, training on the Amazon Echo Show was introduced in the third week, with each site receiving one day of instruction. Participants were also encouraged to use the Amazon Alexa device to ask questions related to the activities in the following weeks. Semi-structured interviews were conducted during weeks three and four to gather feedback from participants and coaches on their overall use of the devices, the functions they engaged with, and the

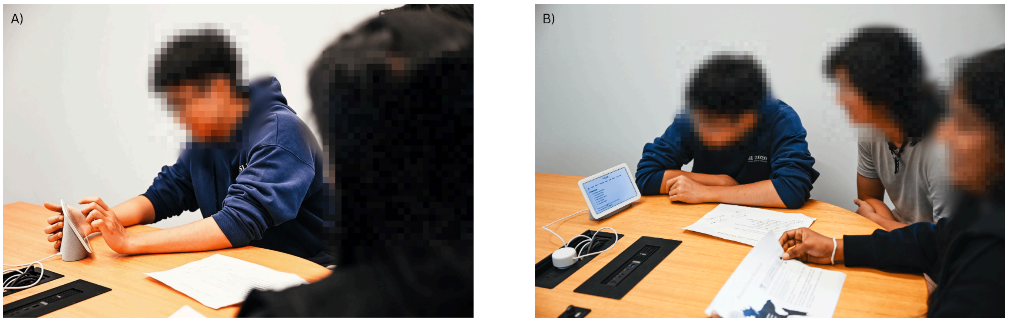

Figure 2: Scenarios of participant engagement with the Google Nest Hub: (A) a participant interacting with the screen while browsing images; (B) a group of participants engaging with the device in a shared session.

challenges they encountered. In addition to general reflections, participants were asked about specific experiences, such as their reactions when the assistant failed to understand them (e.g., *"How did you feel when the assistant didn't understand you?"*). Interview durations ranged from 6 to 38 minutes for participants and 10 to 15 minutes for coaches.

3.3.3 Phase 3 - Post-reflection Interviews : In the post-reflection phase, semi-structured interviews were conducted with both participants and coaches to gain a deeper understanding of their experiences after two months of VA use. These interviews aimed to explore participants and coaches interactions with the devices, the challenges they faced, the strategies they used to overcome those challenges, and any preferences for future VA implementations. Participants were also prompted to reflect on changes in their comfort and confidence with the VAs, for example: *"Do you feel more comfortable using the assistant now than when you first started?"*. Coaches were asked about their experiences with participants' use of the VAs, the perceived benefits, and suggestions for improving VA integration in the program. The interviews ranged from 8 to 12 minutes for participants and 10 to 15 minutes for coaches.

3.4 Data Analysis

3.4.1 Device Usage Data : VA interactions during the study were retrieved from Google's My Activity [29] and Alexa's Voice History [5], which log participants' voice commands, system responses, and recorded audio. Over the eight-week study, participants initiated a total of 234 interactions with the Google Nest Hub and 26 interactions with the Amazon Echo Show across all sites. All authors reviewed the full set of interaction logs to determine an appropriate level of categorisation. Given the specificity of many queries, we adopted a straightforward topic based scheme that captured the general intent of each request (Table 4). For example, requests such as *"Show me an image of Stitch"* was classified within the Cartoon category, *"Tell me about photosynthesis?"* was classified within Science, and *"Show me a video of a Toyota Prius car"* was classified within Vehicles. The first author developed an initial set of topic categories after examining all logs, and the second author cross checked a subset to confirm alignment. Any uncertainties or mismatches were discussed with the research team to ensure consistency. This process resulted in 17 categories that reflected the distinct types of questions present in the dataset.

3.4.2 Observational Data : Field notes were recorded during each session to capture verbal and non-verbal behaviours, such as gestures, facial expressions, and vocalizations, as well as contextual details like peer collaboration and coach support. The notes were reviewed to identify behaviours indicating excitement or frustration, and used descriptively to complement interview and VA usage data.

3.4.3 Interview Data : All interview data (preliminary, deployment, post-reflection) from participants and coaches were audio recorded, transcribed, and analysed to identify key insights and themes related to their experiences with the VA in each phases.

3.4.4 Thematic Analysis : We employed an inductive thematic analysis approach to identify and interpret patterns in the data [15]. To ensure coding consistency, two transcripts (one from a participant and one from a coach) were independently coded by the first and second authors, and the coding was then discussed with the research team. Following this, the first author conducted open coding on all transcripts, including preliminary, deployment, and post reflection interviews. Initial low level descriptive codes were generated and then refined into high-level thematic codes through eight weekly collaborative meetings with the research team. During these meetings, the team discussed excerpts, resolved disagreements, and iteratively updated the document by adding, renaming, or removing codes. When new codes emerged, the first author re-examined all transcripts to locate and annotate additional relevant excerpts. For example low-level codes such as *"Information overload"* and *"Frustrations due to incorrect responses"* were grouped under the high-level code *"user challenges with VA"*.

Themes from the preliminary interviews with participants and coaches were independently generated by the first and second authors to ensure accuracy, and were discussed collaboratively with the research team. Building on this, the first author derived themes from the deployment and post-reflection interviews, which were iteratively reviewed and refined through weekly discussions. We collaboratively explored connections between high-level codes and identified broader themes that captured participants and coaches experiences with the VA. From the thematic analysis, we generated

Table 3: Sample voice commands introduced during training sessions.

Query Type	Example Commands
Foundational Queries	"What is the date today?" "Can you tell me about weather today?"
Utility Queries	"Set a timer for five minutes." "Set an alarm for 10am."
Image Retrieval Queries	"Show me a picture of a car." "Find pictures of a lion."
Media Playback Queries	"Play Peppa Pig music." "Play lion videos on YouTube."
Activity Aligned Queries	"Tell me something about ocean." "Find simple science experiments."
Spelling Support Queries	"How do you spell water cycle?" "Spell the word environment."

48 codes in the preliminary phase, 92 in the deployment phase, and 95 in the post-reflection phase. These were then organized into five themes for the preliminary interviews and four shared themes across the other two phases (Figure 3).

4 Findings

Based on our analysis, Figure 3 presents a summary of the themes derived from interviews across all phases, including preliminary, deployment, and post-reflection sessions. The themes identified during the preliminary interviews with participants and coaches are detailed in the section *"Initial Perceptions of the VAs"* (subsection 4.1). Deployment and post-reflection interviews revealed the same themes, suggesting consistency in the experiences reported by participants and coaches. These recurring themes are further elaborated in the section *"Engagement and Reflections on the VA Use"* (subsection 4.2).

4.1 Initial Perceptions of the VA

In this section, we examine initial perceptions of VAs, which varied according to participants' and coaches' familiarity and comfort with the technology. Several participants (P1, P2, P3, P6, P9, P14, P16, P17) expressed enthusiasm, often linked to previous experience using similar technologies. In contrast, others (P4, P5, P8, P10, P11, P12, P13) showed hesitation, primarily due to a lack of prior exposure. One participant (P15) refused to use the VA due to a habitual preference for typing, which he associated with his identity as a writer. Coaches offered diverse perspectives; three (C1, C3, C4) expressed optimism about the role of VAs in the program, highlighting their simplicity and potential to support learning, while also noting usability challenges. C2, who also preferred typing, raised concerns about the privacy implications of VAs but acknowledged their potential benefits for participants.

Prior Knowledge and Adoption of VAs : In the preliminary interviews, coaches (C1-C3) noted that participants with prior experience using technology, such as smartphones or tablets, tended to be more open to adopting VAs. However, C3 observed that familiarity with technology did not guarantee smooth adoption, highlighting that participants needed time and repeated exposure to feel comfortable using the VAs. As C3 remarked, *"It is not very easy for them to get it, but once they get it, it's gonna become an addiction."* C1 expressed that she had personally introduced some participants to VAs on their mobile phones before our study, highlighting the benefits of voice technology. She described an instance where P2 observed P4 interacting with a VA, which enabled P2 to independently search and access images. Although a few participants had prior experience with VAs and other devices, most expressed a desire for structured training before using the VA. For example, P16 said, *"Training will help us to learn all concepts."* In contrast, P3 demonstrated confidence in their ability to learn independently: *"I learn quickly with technology like the iPad, phone, computer, so I can learn Nest Hub on my own."* While some participants required support to build confidence with VAs, others were more independent and readily able to engage with new technologies.

Mitigating Disabilities and Social Isolation : All coaches highlighted that VAs could help address various communication and accessibility barriers faced by participants. Coaches C1, C2, and C4, along with participant P17, suggested that VAs could be particularly beneficial for individuals with physical disabilities. For example, Coach C1 noted that participant P3, who has use of only one hand, could benefit from voice user interfaces as an alternative to typing. Coach C2 also noted that VAs could assist with challenges related to spelling, mentioning, *"P13 cannot even search something on YouTube because of incorrect spelling."* Beyond their functional benefits, VAs were perceived as tools to reduce social isolation, with Coach C1 emphasizing that participants with limited social exposure (P13), could use the VA as a conversational partner to foster social interaction and emotional engagement.

Supplemental and Simplified Approach to Activities : Coaches (C1-C4) and participants (P1, P3, P6, P16) identified several ways in which VAs could support both daily and educational activities. Coach C2 highlighted the usefulness of screen-based VAs in providing *"step-by-step instructions paired with visuals"* which could help participants complete tasks independently. Coach C3 viewed

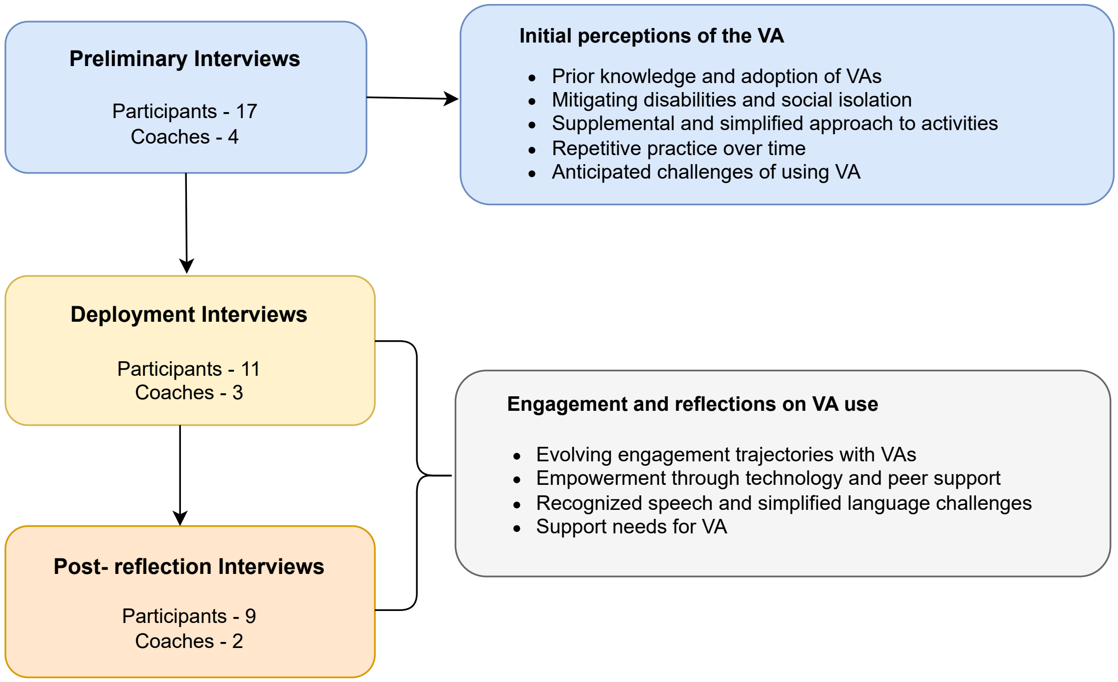

Figure 3: Summary of interview phases and resulting themes. Preliminary interviews (17 participants, 4 coaches) surfaced themes grouped under Initial Perceptions of VAs. Deployment (11 participants, 3 coaches) and post-reflection interviews (9 participants, 2 coaches) highlighted recurring themes grouped under Engagement and Reflections on VA Use, showing continuity across study phases.

VAs as a supplementary tool to assist coaches in visually demonstrating concepts, particularly for non-verbal clients, making it easier for them to understand content. C4 also pointed out the challenges participants face in obtaining specialized communication devices, which typically involve a lengthy and complex selection process. In contrast, VAs provide a more immediate and accessible alternative. Participants recognized the potential benefits of VAs in supporting various activities, with several noting that VAs could assist with tasks like searching for information (P3, P11, P14), checking the weather (P1, P3, P16), spellings (P3, P6) and watch videos (P1, P3, P14). Some participants (P4), also viewed VAs as a source of entertainment, using them to listen to music or watch videos. Additionally, many appreciated the convenience of voice commands, especially when feeling tired or overwhelmed, with P4 sharing,*"When I'm tired of typing, I can use the VA"* highlighting the situational value of voice interaction in reducing cognitive or physical effort.

Repetitive Practice over Time : Coaches (C1, C3) emphasized that learning to use VAs is not a *"one-time process"* but requires sustained engagement and repetitive practice. Coach C1 highlighted the importance of ongoing practice, stating, *"This will take a long time, with repeated practice, until they really know."* Coach C3 similarly noted that VA use must be *"embedded into their routines"* explaining that without regular interaction, participants may lose their familiarity with the technology. They further remarked that consistent practice not only helps participants retain their skills but also enables them to gradually build confidence, making it *"harder for them to back up"* (C3) once they become accustomed to the technology.

Anticipated Challenges of using VA : Despite the advantages, coaches (C1-C4) and participants (P3, P16) both shared several challenges related to the use of VA. A frequently raised concern was speech recognition, particularly difficulties with unclear speech (C1-C4), slow speech patterns (C1, C2), and atypical pronunciations (C1). Coach C1 emphasized the lack of bilingual support in VAs, referring to the technology as *"it's not bilingual… it's more English"* stating that in a communal environment where language settings favour the dominant language, such limitations fail to meet the diverse linguistic needs of users. Coaches also highlighted the challenge VAs face in multi-speaker environments. Coach C1 pointed out that background noise and overlapping voices in group settings can reduce the VA's accuracy and responsiveness. Similarly, Coach C4 highlighted the variability in cognitive abilities among participants, stating, *"the clients here are very different... their intellectual ability is lower than the clients in other sites, it's going to be a bit tricky here, but it's trial and error."* Additionally, Coach C3 expressed concerns about participant frustrations if the VA fails to respond as expected, explaining, *"If they want it to work and it doesn't, they* *smash it... they are not patient."* This suggests that participants with lower frustration tolerance may discontinue using the technology if their experiences are consistently negative. Even participants with prior VA experience at home also reported challenges, particularly with pronunciation (P3, P16) and slow response times, with P16 stating, *"It takes forever to load answers."*

4.2 Engagement and Reflections on VA Use

In this section, we explore participants' engagement with VAs, focusing on the facilitators and barriers that influenced use, as well as the support needs identified by coaches and participants. During the two month deployment, participants used either both the Google Nest Hub and Amazon Echo Show devices, or just one based on their preference, across a variety of activities. These activities included searching for variety of information such as checking weather, accessing entertainment, and asking learning related questions (Table 4). As the deployment progressed, some participants became more confident in using the VAs (P2, P3, P4, P6, P11, P12, P16, P17), while some continued to rely on assistance (P5, P8, P10), primarily due to challenges they encountered when using the VAs. These findings suggest that although participants' familiarity with the VAs increased over time, communication challenges continued to affect sustained engagement.

Evolving Engagement Trajectories with VAs : At the beginning of the deployment, participants tended to use the VAs for simple tasks such as checking the weather, searching for information, or spelling. Over time, their engagement broadened: several participants explored additional functions, including playing games (P16, P17), asking science related questions (P1, P3, P6, P16, P17), and using the assistants in more socially oriented ways (P13, P17). Alongside this expansion in activities, participants' queries also became more specific and structured over time. Early interactions typically involved short requests with a single intent (e.g., *"Tell me the weather"*, *"Show me an image of a dog"*). As participants became more familiar with the VA, some (P3, P16, P17) began asking longer, more detailed questions that added contextual information, such as *"Tell me the weather in Seaford tomorrow"* and *"Explain photosynthesis step by step"*.

Several participants showed hesitation when interacting with the VA during the early sessions. Some paused after the wake word (P5, P8), whispered their queries (P10),or abandoned the questions when the device did not respond (P8, P9). Participants P8 and P10 at times declined to use the device and expressed that they felt afraid of using the VA. Others (P12, P13) commented, *"I don't know what to ask"*, often looking to coaches or researchers for reassurance before speaking. As the deployment progressed, these behaviours shifted noticeably, with participants approaching the device with greater ease, laughing at unexpected responses (P16, P17), repeating queries for amusement (P16, P17), and experimenting with features that had been unused in earlier sessions. For example, many adjusted the volume (P4, P16, P17), switched screen layouts (P16, P17), scrolled or tapped through results (P2, P3, P4, P6, P16, P17). P17 commented, *"I will try that again until I get it right,"* while refining their phrasing across repeated attempts. Coaches (C1, C3) similarly observed that participants' (P2-P5, P16, P17) questions became more varied and detailed over time.

The deployment revealed varied learning trajectories, with participants advancing at different paces and engaging with the VA in distinct ways. Some showed substantial growth and confidence (P3, P6, P16, P17), others maintained consistent usage patterns (P4), a few preferred using the VA only on certain days (P6), and some continued to rely on coaches or peers to engage meaningfully (P5, P7, P10). For example, P12 who had never used a VA before gradually used it for creative purposes such as gathering design suggestions for multiple projects. P16 expanded from basic queries (weather, time, image searches) to interactive Alexa games and science questions such as how heart functions. In contrast, P4 primarily used the VA for listening to music and only occasionally for search. Although participants and coaches (C1–C3, P3, P6, P16) valued spelling functions in pre-interviews, spelling was used less frequently than searching for information or checking the weather.

Beyond these learning trajectories, participants' interaction methods also changed over time. Although both VAs are primarily designed for spoken commands, participants frequently combined voice with touch, scrolling through results, selecting images or videos, and reopening links. Across the deployment, 59 display based interactions (view) were documented in the log files. Log data and session observations together suggest that participants increasingly incorporated touch from around the fourth week onwards, particularly when they wanted finer control, preferred to browse options visually, or were engaging with video content. Their consistent use of both speech and touch across the two devices indicates that they were drawing on a shared set of multimodal affordances voice output, on-screen feedback, and touch based navigation rather than relying on device specific features. In addition, P13 and P17 began forming a more personal relationship with the assistant, greeting (e.g., *"How are you today?" "Good morning Google"*), expressing their feelings, and commenting on its responses. These moments suggest that participants' engagement extended beyond task oriented use and, for some, offered a sense of companionship.

Empowerment through Technology and Peer Support : The use of VAs had varied impacts across participants, revealing both empowering outcomes and ongoing challenges. Coaches (C1, C2) reported that VAs were particularly helpful for participants who struggled with specific tasks like spelling (P3, P6, P16), searching for information (P2, P5, P8, P10). They also observed that some participants struggled during interactions, highlighting the need for VAs to support diverse speech patterns and communication styles (P5, P10).

Despite these limitations, many participants gained confidence and demonstrated increased autonomy over time, suggesting that repeated exposure and familiarization with the technology can support a sense of empowerment. For instance, P3 emphasized the VA's role in helping with tasks, stating, *"It helps to get knowledge and skills."* Similarly, P17 reported fewer typing-related errors, stating, *"I don't get the lines for incorrect words in a word document"* highlighting how voice commands could reduce cognitive demands by simplifying tasks. P6 articulated the advantage of having access to real-time information through the VA, saying, *"Books are not updated; VAs are updated"* reflecting the importance of current and responsive content for supporting learning and curiosity. In addition to these functional benefits, several other participants (P4, P6,

Table 4: Weekly and categorical totals of user queries with VAs across all sites over eight weeks.

Category/Week	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Total (Category)
Vehicles	-	3	5	-	-	1	-	4	13
Sports	2	-	1	2	-	-	-	-	5
Animals	-	-	6	15	2	6	8	3	40
Foods	-	-	-	2	-	-	-	-	2
Seasonal Info	-	-	-	-	16	1	4	-	21
Arts	1	10	2	1	-	6	-	5	25
Science	-	-	-	4	-	-	4	-	8
Spelling	3	-	1	-	-	1	2	1	8
Cartoon	-	-	-	7	-	4	4	-	15
Jokes	-	2	5	2	-	-	-	-	9
Music	1	1	-	10	-	2	-	4	18
Movies	1	-	1	1	-	-	1	4	8
Games	-	-	-	-	-	-	-	2	2
Weather	2	3	5	6	2	-	7	5	30
Alarms	2	-	10	4	4	3	-	3	26
Date and Time	-	1	1	4	-	-	3	-	9
Other	3	-	3	5	3	2	5	-	21
Total (Weekly)	15	20	40	63	27	26	38	31	260

Note. "Other" includes queries that did not fit into the main categories (e.g., what is your name, tell me a random thing, turn down the volume).

P12, P16, P17) expressed that VAs helped them take initiative and engage and search information on their own. P15, who initially refused to use the VA, began interacting with it in the third week to check spellings. However, we were unable to observe long-term usage because the participant left the organization shortly afterward. Coach C3 occasionally used the VA to demonstrate concepts to clients, highlighting its potential for staff-led facilitation.

The impact of VAs went beyond individual empowerment, promoting peer support and collaboration. Our observations revealed that when one participant faced challenges, others stepped to help, promoting collective problem solving. For instance, P16 assisted P17 with voice queries (e.g., wake actions) early in the deployment, and by the end of the phase, P17 became independent in resolving similar issues. Furthermore, P3 expressed a strong desire to support peers with more severe disabilities, stating, *"I want to teach other people like those who have worse disabilities than me to use VAs… show them that disabilities can be overcome, and then follow my footsteps."* Proxy use was common where participants with stronger verbal skills meaningfully assisted those with limited or non-verbal in interacting with the VA. For example, P3 translated P7's gestures into voice commands, enabling access to videos, images, and interactive content. Interestingly, P7 used his hand on the VA screen to scroll through and select videos and images from the generated results. This mutual support highlights how VA interactions can foster inclusive technology use.

The quantitative results show that engagement in Weeks 1 and 2 was low, as participants were hesitant to approach the unfamiliar digital systems. By Week 4, usage increased sharply, indicating that with time and support, participants gained confidence and engaged more meaningfully with the VAs. Participants most frequently interacted with topics such as searching for animals, weather, arts, seasonal information, and alarms, suggesting that familiar and emotionally engaging content supports comprehension and encourages continued use of VAs. Participants' interactions were classified based on topics, as either general questions or STEAM related questions. Among the 260 recorded interactions, 161 (62%) involved general use, while 99 (38%) were STEAM related. Several participants (P3, P6, P12, P16, P17) expressed that errors became less frequent over time, particularly compared to their initial experiences with the VA. Our interaction logs also showed a gradual reduction in errors for some participants. For example, P3 initially had a 39% error rate during the first four weeks, largely due to taking a long time to formulate questions. By preparing and speaking

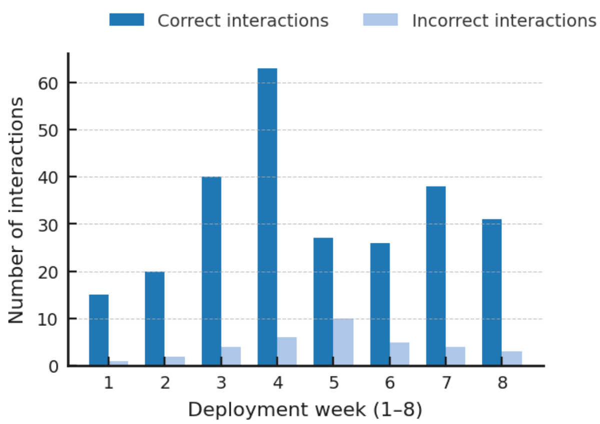

Figure 4: Summary of weekly correct and incorrect interactions over the eight week deployment.

out questions in advance, her error rate dropped significantly to 12% in the final four weeks.

These findings suggest that extended usage is crucial for individuals with ID, as continued use leads to improvements in voice user interface interaction. By preparing and speaking out questions in advance, her error rate dropped significantly to 12% in the final four weeks. Beyond this individual trajectory, Figure 4 presents a summary of correct and incorrect interactions across the eight week deployment. Correct interactions consistently outnumbered errors (260 correct vs. 35 incorrect overall), indicating that participants were able to complete their intended requests. In the early weeks, both correct and incorrect interactions were relatively low, reflecting an initial learning phase with shorter, simpler queries. Usage peaked in the middle of the deployment, with week 4 showing the highest volume of correct interactions (63 correct, 6 errors), followed by a smaller number of correct interactions but a noticeable spike in errors in week 5 (27 correct, 10 errors), suggesting that participants were beginning to experiment with more complex queries. In the later weeks (weeks 6–8), both correct and incorrect interactions decreased and stabilised, pointing to the emergence of more settled usage patterns in which participants continued to use the VA successfully while still occasionally encountering errors.

Recognized Speech and Simplified Language Challenges : Coaches and participants identified several challenges that affected the usability of VAs, including pronunciation difficulties (C1–C3, P4, P5, P8, P10–P13, P16, P17); cognitive challenges such as complex responses (C1–C3, P16, P17) and difficulty remembering wake words (P5, P6, P9, P10); and other challenges including slow response generation (C1, C2, P16) and unintentional activations (P1, P4, P6, P11, P16).

Analysis of the pronunciation related errors in the logs revealed four prominent issue types, including mispronounced words by participants (P5, P8, P11, P13, P17), ambiguous or grammatically incorrect sentences (P8, P11, P12, P13), soft-spoken speech (P4, P5) and timeout errors (P10, P11). C1 described, these complications often made partcipants feel confused, while P17 humorously expressed frustration, remarking with a smile, *"Google is not working at all. I feel like putting Google out the window."* Quantitative logs revealed around 15% error rate with Google Assistant, with 234 successful interactions and 35 errors. Errors were primarily due to mispronunciation (16 errors), ambiguous questions (10 errors) and timeout errors (nine errors) often followed by the response, *"I don't know, but I found these results on search."* By contrast, Alexa showed no recognition errors across 26 logged interactions. This likely reflects its later adoption in the study and its use primarily by participants with stronger verbal skills. Further analysis of mispronounced words (P5, P8, P11, P13, P17) revealed three dominant patterns [53, 55].

- Minimal pair confusion: where words differing by a single phoneme were misrecognized (e.g., *"leader"* as "ladder").
- Phonemic substitution: where one or more phonemes were misrecognized (e.g., *"sandwich"* as *"Swedish"*).
- Named entity distortion: where proper nouns were misrecognized (e.g., *"Makey Makey"* as *"Mickey Mickey"*).

When participants mispronounced a word or a sentence, the researcher asked them what they actually intended to say. Their intended interaction was then compared with the VA logs to determine whether the error came from the VAs or from the participant's speech. Another common challenge related to pronunciation was the formulation of queries using ambiguous or grammatically incorrect sentences. Participants often asked ambiguous questions such as *"How was the cricket match?"* which the VA struggled to interpret accurately. Timeout errors occurred when users required extended time to formulate their questions or when they stammered.

Soft-spoken speech created problems with voice recognition, which were identified through researcher and coach observations since the device only logged errors after successful activation. P5 explained, *"VA don't understand whether I speak far or close, so I don't talk,"* reflecting a reluctance to engage due to inconsistent voice recognition. While some participants hesitated following initial errors, several persisted, rephrasing or adjusting their speech until achieving success, reflecting growing confidence in their task performance. For example, P16 stated, *"I'm going to ask that question louder again"*, and P17 noted, *"I will try that until I get it correct"*, illustrating a proactive approach to self-correction and persistence, reflecting an *"I can do it"* attitude.

In addition to pronunciation challenges, various cognitive challenges persisted throughout the study duration, including issues with comprehending complex responses (C1-C3, P16, P17) and difficulties in remembering wake words (P5, P6, P9, P10). Participants frequently had difficulty comprehending VA responses when the answers were long or complex, particularly when simple questions received overly detailed or verbose replies. Several participants (P2, P12, P17) asked the assistant to stop mid-response, indicating difficulty in processing or retaining lengthy information. Coaches (C1, C2) observed that participants often felt cognitively overloaded by the assistant's lengthy and complex responses. As C1 reflected: *"It comes out with 100 million things… they are already stuck; they don't need another question."*. Wake word issues, often tied to memory and recall challenges, were not reflected in the interaction logs since unsuccessful wake attempts go unrecorded. However, observational data revealed that these failures disrupted interactions and negatively affected participants' overall experience with the VAs. Even by the end of Week 8, some participants (P5, P10) continued to struggle with recalling the wake word. Notably, one participant (P5) who had previously used the device but rejoined the program after a three-week break, questioned *"What is this device?"* when encountering the VA.

Other challenges included slow response times, unstable internet connections, and background noise, all of which contributed to disengagement. Unintentional activations also posed a barrier, as casual or incidental use of wake phrases during conversation frequently triggered the assistant unexpectedly, disrupting the flow of interaction. Coach C2 explained, *"There's not a lot of patience,"* while Coach C1 pointed out, *"people will get frustrated, especially when you are in a rush."* P16 expressed frustration with the speed of loading answers, saying, *"It takes ages to load."* Additionally, environmental noise made it difficult for some participants to concentrate, as P16 noted, *"It's difficult to concentrate".*

Support Needs for VA : Participants and coaches identified several areas where the VAs required additional support or adjustment to better suit the needs of the group. A commonly observed challenge was that the VA often returned responses that were more detailed than the activity required. For example, Coach C1 explained, *"For the water cycle project, I told them to search for it, but it came up with complex projects… not what they wanted."* This reflected a broader pattern across sessions, where lengthy or highly detailed responses made it difficult for participants to follow the information. Coach C1 suggested that allowing participants to request simpler explanations could reduce confusion and support clearer understanding of the responses. Several participants (P2, P3, P5, P16, P17) described VA responses as *"too long"* and difficult to follow. When replies became overly complex, some used the *"stop"* command to interrupt the assistant sometimes laughing (P2, P16, P17), and at other times showing visible frustration (P1, P8). These moments were often accompanied by signs of confusion such as looking away from the screen (P3), or turning to a coach or researcher for help (P2, P5).

Across the sessions, several participants found it difficult to activate the VA, particularly when the wake phrase was not recognised or when they forgot the exact wake word. Some repeated the wake word multiple times (P4, P9, P17), spoke more loudly or slowly (P11, P12, P17), or leaned closer to the device (P4, P10); a few succeeded using these strategies, while others (P5, P8) eventually sought help from coaches or researchers or asked them to *"can you tell that for me".* To address this, C1 suggested adding a press-to-talk button or adopting shorter activation cues such as *"max."* P17 also proposed using cartoon character names for example, Peppa or Dora to make activation easier to remember, noting that a cartoon style voice would feel more engaging and easier to remember.

Issues related to slow responses and unstable internet connectivity also affected how smoothly participants could use the VA. Coaches emphasised the importance of fast VA responses, noting that participants should not have to wait long for an answer. Participants also expressed similar frustrations; for example, P6 described the device as *"take years to load"* signalling dissatisfaction with delayed responses. During these pauses, several participants disengaged; some stopped mid-task (P4), and others shifted their attention to other activities (P10, P13). The VA also froze or stopped responding mid-query at several points, often due to connectivity issues, leaving participants unsure about how to proceed. In these moments, some participants waited silently before repeating their question (P4, P16, P17), while others (P8, P10) turned to coaches or researchers to ask whether the device was broken. Coaches noted that such interruptions made it difficult for participants to maintain continuity in their activities, and suggested that an offline mode could help reduce the impact of these disruptions.

To improve engagement and reduce monotony, coaches recommended incorporating interactive elements such as games and quizzes, with one coach emphasising, *"Games and quizzes attract clients,"* expressing how gamified experiences can capture attention, encourage participation, and prevent the VA from becoming tedious over time. When we asked participants whether they would like game or quiz based interactions, P16 and P17 responded enthusiastically, asking whether they could earn *"points"* or *"stars"* for correct answers, highlighting their interest in reward based, gamified interactions. Another challenge noted by coaches was the difficulty of managing multiple devices during sessions, particularly when participants were already using computers and tablets for their activities. One coach commented, *"There's too many devices"* referring to the cognitive effort involved in switching between systems. This challenge was also visible during observations, where several participants (P4, P5, P12, P13) attempted to use the VA while using the computer and became confused about which device to address. Some (P4, P5) tapped the computer screen instead of the VA screen, indicating that physical and attentional switching created friction.

For non-verbal participants, coaches proposed adding multimodal interaction options, such as a simple touch interface where users could tap on symbols or images to receive responses. As Coach C1 explained, *"Having a simple touch interface, where they could tap on symbols or images to get a response, would work well."* Coaches further suggested incorporating visual cues text, icons, or small animations to help users follow along with the VAs actions. Some also noted the importance of interpreting facial expressions to better understand non-verbal participants' needs. Assistive typing methods were also identified as important for non-verbal users, with C1 noting, *"If you can have some assistive typing method also, it would be beneficial".* A few verbal participants (P4, P17) noted that offering both text and voice input options would make the VA easier to use, as the ability to switch between modalities could better support their comfort and communication needs.

5 Discussion

Our study provides empirical insights into how adults with ID engaged with VAs during an eight week deployment in a group settings. While many participants experienced the devices as enabling and enjoyable, others encountered challenges such as pronunciation related misrecognition and cognitive difficulties related to memory and recall. We begin by situating our findings within research on accessibility and human computer interaction, focusing on VA adoption, user agency, and social engagement, and then present the design considerations that emerged from our study.

5.1 VA Adoption, User Agency and Social Interaction

Our findings on VA adoption align with prior studies showing that VA use can support people with ID by enabling access to information (e.g., checking the weather, looking up facts, or asking questions) [12, 46, 70], entertainment (e.g., playing music, videos, or simple games) [12, 77], and everyday routines (e.g., setting timers and alarms, organising daily activities) [77], as well as by contributing to a sense of increased independence [46, 70, 77]. At the same time, our study extends this prior work by examining how deployment duration and the social context of use shape VA adoption. Unlike prior studies that reported short deployments (typically less than one month) [46, 50], our study enabled us to observe how VA adoption evolved over time. As participants became more familiar with the VAs, recognition errors decreased and interactions shifted from simple requests to more complex questions. In addition, most studies in this space examine one-to-one VA use, often within controlled or semi-structured activities [12, 46], our deployment took place in a group based environment in which 3–4 participants shared a single VA device and used it for unstructured activities. This configuration illustrates that VA adoption can be a collaborative process, with VAs becoming embedded in shared activities rather than solely individual interactions. In this setting, adoption was strongly scaffolded by both peers and coaches: peers repeated commands for one another, modelled successful phrasing, and suggested new queries, while coaches encouraged retry attempts and cued familiar routines. Although previous work has primarily emphasised the role of caregivers or coaches in sustaining VA use [46, 70], our observations foreground peer scaffolding as an additional mechanism that supports continued adoption and helps participants refine how they use VAs across different activities.

Prior work suggests that people with ID experience increased agency as they become more comfortable using VAs when performing different tasks. Smith et al. [71] found that most participants felt VAs made them *"better able to do things for themselves"* even though quantitative measures indicated only modest gains in agency. In contrast to Smith et al. [71], our observations and interaction logs suggest expanded forms of agency as participants grew more familiar with the VAs. Participants increasingly initiated their own queries [12], configured the devices independently [77], and shifted from simple, single intent requests to more complex questions [70] consistent with prior work. Beyond these changes, we also observed further shifts in agency over the course of the deployment, as participants began experimenting with device settings (e.g., changing screen layouts) and became less reliant on coaches and researchers when using the VA. Rather than interpreting these patterns as a straightforward shift from dependence to independence, we observed agency as situated and collaboratively developed. Participants routinely welcomed support from coaches and peers asking for help to rephrase a query, checking the meaning of a response, or jointly deciding what to ask next when they encountered difficulties while still treating the interaction as something they were doing themselves. This contrasts with studies suggesting that some people with ID may be reluctant to seek help when performing activities [40]. In our group based setting, collaborative activities instead appeared to normalise help seeking and frame it as a shared effort toward accomplishing a task.

In our deployment, participants engaged with the VA in ways that often resembled companion like interactions where they routinely greeted the device (e.g., saying "good morning") [46], engaged in small talk [17], and requested jokes or playful content [12, 70]. Beyond this, prior work has reported that sustained VA use can reduce feelings of social isolation and strengthen perceived social connectedness not only among individuals with ID [77] but also among older adults [22, 39]. However, we did not observe VAs taking on a substitutive role for human contact or social isolation in our study, likely because interactions were embedded within group sessions where participants were co-present and actively engaged with peers and coaches.

Although VAs can provide added benefits for social interaction, their use also needs to be interpreted within a broader sociotechnical context. Research on disability and accessibility has cautioned against positioning VAs as replacements for human social interaction, noting that socially expressive technologies can inadvertently encourage over anthropomorphism, reduce opportunities for face-to-face communication, or shift relational expectations from humans to machines [41, 80]. Similar concerns have been raised in human robot interaction research, where socially engaging behaviours have been shown to shape users' trust, privacy expectations, and assumptions about emotional support [6, 68]. Recognising this boundary is essential for designing systems that support social participation without displacing the human relationships that structure everyday life for individuals with ID. Future work should therefore incorporate safeguards such as clear signalling of system limitations, strong privacy protections, and interaction designs that foreground collaboration, so that VAs augment rather than replace existing social supports.

5.2 Design Consideration 1: Auto-Correction for Mispronunciation and Grammar

Our findings revealed that VAs often misinterpret atypical speech from participants, leading to errors and reduced engagement. The primary challenges observed were mispronunciations (P5, P8, P11, P13) and ambiguous or grammatically incorrect utterances (P8, P11, P12, P13), both of which challenged VA recognition and interpretation. Similar challenges have been reported in prior studies involving individuals with ID [12, 46, 50], people with speech impairments [37, 66] and older adults [60, 75]. To address this, VAs could integrate an auto correction mechanism, similar to spelling correction tools in word processors [52], that detects atypical pronunciations and offers real-time correction or prompts. Machine learning models, such as random forest classifiers, have been effective in identifying lisped sounds, like mispronunciations of *"s"* or *"z"* [35], and could also extend to handle confusion between minimal pairs, such as *"leader"* and *"ladder"*, which were observed in our study.

Proenca et al. [65] also propose a two stage mispronunciation detector for children's sentence reading. The approach first segments each spoken word attempt into syllables and then compares the expected and produced sounds to decide whether the word was pronounced correctly, demonstrating that mispronunciations can be explicitly modelled. Training VAs on atypical speech data has been shown to improve recognition accuracy for individuals with autism spectrum disorder [10], suggesting similar benefits for users with ID. Off-the-shelf devices could further enhance accuracy by allowing user-specific adaptation, fine tuning speech recognition models to individual users' pronunciation patterns. This combination of detection, correction, and personalization can reduce communication breakdowns, improve overall interaction flow, and make VAs more usable, accessible, and empowering for individuals facing pronunciation challenges.

Another approach is to adapt predictive text features [79], which have improved typing performance in specific need groups [3], so that VAs can suggest or confirm words during interactions with atypical speech. Such support would not only improve communication accuracy but also scaffold learning by guiding participants toward correct word selection. A common challenge we observed involves phonetic substitutions, where one sound is replaced with another, leading to misinterpretations by VAs [53, 55]. For example, a participant might mispronounce *"sandwich,"* leading the system to interpret it as *"Swedish."* To address such errors, VAs could integrate real-time correction features. When the recognition model produces an unlikely word given the context, the VA could prompt the user for clarification either visually on the screen or verbally. For instance, the misrecognized word could be highlighted, allowing the user to confirm or choose the intended option. This reduces the need to repeat full sentences, enhances recognition accuracy, and supports smoother, more accessible interactions for individuals with ID [56].

The other challenge involved ambiguous or grammatically incomplete queries (P8, P11, P12, P13), which frequently lacked the contextual information needed for the system to infer the user's intent. For example, P11 asked, *"How was the cricket match?"*, a vague query that lacked the contextual details needed for the VA to identify which specific match was being referenced. Similar patterns have been reported in studies with individuals with ID [12], older adults [39, 63], and general users [43] where underspecified queries often lead to system misunderstandings. To address such issues, we recommend integrating repair mechanisms into VA systems, such as clarification prompts [23]. When a vague query is detected, the VA could respond with a follow-up like, *"Did you mean the [x] cricket match?"*, and display a few alternative options on the screen. For users with ID, these prompts scaffold query reformulation, reduce the likelihood of communication breakdowns, and help them stay oriented in the interaction. More broadly, such error tolerant designs operationalize universal design principles [73] particularly tolerance for error and simple and intuitive use by shifting part of the communicative burden from the participant to the system. The same mechanisms could also benefit other groups, such as older adults who experience memory or recall difficulties [39, 64].

5.3 Design Consideration 2: Customize the Listening Window of VAs

Several participants (P5, P6, P9, P10, P13) experienced difficulties interacting with VAs due to delayed speech, including stammering or long pauses while formulating questions. For instance, when P2 attempted to say *"Can you show a video of an earthquake?"*, the assistant only registered *"Can you show me a video?"*, missing the crucial part of the request because it timed out too early. Similar challenges have been observed in individuals with ID [12, 77], and older adults [39, 42, 64] highlighting the need for VAs to accommodate varied speech patterns. These delays often caused the VA to misinterpret pauses as the end of input, leading to premature cut-offs or timeouts. The problem arises from the fixed endpoint detection in off-the-shelf VAs, which typically rely on short silence thresholds (500 ms–1s) that cannot be manually adjusted [32]. While commercial platforms like Alexa Skills do not allow direct modification of these thresholds, developers can partially mitigate the issue using multi-turn sessions, extended re-prompt intervals, and session attributes to handle partial or interrupted input.

To improve usability, we recommend extending the listening window to provide participants with slower speech additional time to complete their utterances [36, 51]. Listening windows could be dynamically adjusted by monitoring speech patterns, such as pause durations and speech rate, and adapting endpoint detection thresholds in real time. Providing a clear visual cue, such as a *"Listening"* prompt on screen-based devices, signals when the VA is actively processing input, reducing uncertainty and increasing confidence. Minor adjustments to backend speech processing (1–2%) can also improve recognition accuracy by minimizing extended pauses between syllables [42]. Consistent with universal design principles, these capabilities should be exposed as configurable options in VA settings rather than being restricted to accessibility modes. Adjustable listening windows can extend usability beyond adults with ID to a wider range of users who speak slowly or require more time, including older adults [39] and non-native speakers [54].

5.4 Design Consideration 3: Simplified and Easy to Understand Responses

Several participants, including more independent participants (P6, P8, P16, P17), struggled to comprehend the responses provided by VAs. When P16 asked about the water cycle, the VA responded: *"The water cycle is the continuous movement of water within the Earth and atmosphere... processes like evaporation, condensation, precipitation, and collection... "*. This response was too complex for the participant, who became confused by unfamiliar terms and struggled to follow the explanation. Similar challenges have been reported in studies with individuals with ID [12], and older adults [39, 64] where they struggled to understand information that was overly complex or not tailored to their cognitive needs. A simplified version, as suggested by C1, might be phrased as: *"The water cycle is how water moves around the Earth. Water heats up and turns into vapour. It forms clouds and then falls as rain or snow. The water goes back to rivers and oceans, and the cycle starts again."*

These findings highlight the need for responses that align with participants' cognitive processing abilities, including those of non-native speakers and people with cognitive disabilities. For example, Ospina-Henao et al. [61] show that, with prompt engineering, large language models can rewrite complex texts into Easy-to-Read versions with shorter sentences, simpler vocabulary, and added explanations, improving accessibility. Building on this, we suggest using reading measures such as Lexile scores [45] to dynamically adjust VA responses. Lexile scores provide a standardised measure of text complexity across 12 levels, with lower levels associated with simpler texts and higher levels with more advanced terminologies.

Level 1 corresponds to scores between 0-300L on the Lexile scale, while Level 12 corresponds to scores between 940-1210L. Lower levels use shorter sentences and more familiar vocabulary, which can make explanations easier to follow for individuals with ID, who in our study often struggled with multi-clause sentences and abstract terminology. For example, when users asked *"What is an apple?"*, VAs responded with explanation: *"An apple is the round, edible fruit of an apple tree (Malus spp.). Fruit trees of the orchard or domestic apple (Malus domestica), the most widely grown in the genus, are cultivated worldwide. The tree originated in Central Asia, where its wild ancestor, Malus sieversii, is still found."* Using a standard readability formula, this response corresponds to approximately a Grade 7–8 reading level, and it introduces specialist botanical terms that are unlikely to be familiar to our participants. Several participants asked coaches to explain these terms, indicating that the default explanation was too complex for their needs. A simplified alternative, of Lexile level 1 is, *"An apple is a fruit that grows on trees. It can be red, green, or yellow."* expresses one idea per sentence using everyday language and avoids abstract terms, aligning with the explanation style participants found more comprehensible. To further support comprehension, VAs could also offer interactive simplification features. A screen based *"Simplify further"* button or a spoken prompt such as *"Would you like a simpler version?"* would allow participants to request clarification when needed, ensuring that responses remain cognitively accessible.

This could potentially be accomplished through language models that produce multiple versions of a response at different Lexile levels, but further research is needed to explore this area. Although our study focuses on individuals with ID, similar simplification mechanisms could also support children with learning disabilities by simplifying text and making it easier to read, thereby extending the accessibility of VA explanations.

5.5 Design Consideration 4: Supporting the Establishment of VA Routines

Participants often relied on familiar patterns when using VAs, and once a device became part of their routine, they were reluctant to switch to alternatives (P2-P8, P10, P11). For example, even after the Echo Show was introduced, most participants continued using the Google Nest Hub they had first learned, while only two independent participants (P16, P17) explored both devices and eventually preferred the Echo Show. Peer influence also shaped adoption, when P16 began using Alexa, P17 quickly followed, showing how social dynamics in group settings could encourage or discourage experimentation. Some participants (P2, P11) developed routines for interacting with the VA, such as consistently asking about the weather before starting other activities.

This behaviour suggests that supporting habitual interaction patterns could help participants engage more effectively with VAs. These observations align with prior work showing that individuals with ID can benefit from structured or repeated interaction patterns to scaffold task completion and increase confidence [81]. Moreover, participants and coaches favoured screen based VAs, as visual prompts supported memory, comprehension, and navigation, reinforcing consistent use over time. Studies with older adults have noted that established routines strongly shape technology use and are often difficult to change [14, 18]. Research with individuals with ID similarly shows that VAs can become embedded in daily activities and provide consistent support for routines [46, 77].

Designing for routine formation requires VAs to recognise and support patterns in participants' repeated behaviours. For example, if a participant typically wakes around 6:00 a.m. and checks the weather at 6:15 a.m., the assistant could proactively suggest *"Would you like to check the weather?"* rather than waiting for an explicit request, in line with prior work [81]. Routine formation can also be reinforced through interface design by placing frequently used actions on the home screen, thus allowing participants to begin familiar tasks without needing to recall specific phrasing. These strategies may additionally benefit older adults who rely on VAs for regular medication reminders or daily routines.

5.6 Design Consideration 5: Supporting Attention and Engagement in VUI

Prior work has reported difficulties sustaining attention and engagement in children with ID, autistic teenagers, and other populations with specific needs [2, 7, 31], our findings also show that these challenges are equally salient for adults with ID. Across our deployment, participants (P1, P8 - P11, P16, P17) engaged more actively with the VA when coaches provided motivation or proactive guidance. Prior research has similarly shown that motivational prompts (e.g., positive feedback after completing tasks) [62] and proactive cues (e.g., routine linked reminders to resume tasks) [11] can support engagement for individuals with ID and related support needs. In our study coaches also often used short affirmative prompts such as *"Great job!"* or *"Are you ready to search for videos?"* and we observed that these cues helped participants stay oriented, follow instructions, and progress through tasks with greater confidence. Coaches also suggested incorporating interactive elements such as step-by-step guidance or short quizzes, which can help maintain attention by providing predictable cues and reinforcing task progression.

Although these strategies were effective in our study, proactive prompts are rarely used outside task focused activities in current VA designs [47, 58]. This raises a broader design concern that prompts can enhance engagement, but they must be introduced carefully to avoid disrupting ongoing activities or overwhelming participants. In the context of individuals with ID, this kind of prompting may be particularly beneficial in educational settings, where participants can struggle to maintain attention or follow multi-step instructions independently [11]. Timely prompts can remind participants of the next step and reinforce key ideas at the moments when they are needed most. To support this, VAs could offer configurable engagement modes or allow support workers to adjust prompt frequency, providing structure when needed while still allowing users to interact independently when appropriate.

5.7 Design Consideration 6: Incorporating Multi-modal Input and Feedback in VAs

VAs are primarily designed for spoken interactions, which can limit accessibility for individuals with speech difficulties or those who are non-verbal [21]. Consistent with these limitations, our study found that voice only interaction created persistent barriers for several participants with ID, particularly non-verbal participants (P7) and minimally verbal participants (P5, P8), who were unable to reliably produce spoken commands. Some verbal participants (P4, P16) also reported that speaking to the device was effortful at times and expressed a preference for alternative input methods, such as typing. These findings align with prior work showing that speech recognition difficulties and reliance on caregiver support can constrain VA use for individuals with ID [11, 62], highlighting the importance of multi-modal interaction. Providing alternative input options not only supports non-verbal users who rely on other communication methods but also benefits verbal users with ID who may face challenges with speech clarity or comprehension. In addition to alternative input methods, our findings highlight the need for VAs to provide multimodal feedback rather than relying on voice responses alone. In our deployment, when the VA displayed images or on-screen text alongside spoken responses (e.g., images of the water cycle), these visual cues substantially improved comprehension for several participants, particularly P4, P16, and P17, by turning abstract explanations into concrete, easy-to-follow representations.

Multimodal options including touch, gestures, symbolic interfaces, and haptic feedback, can provide more flexible and accessible ways to interact with VAs. For example, users could tap large, high-contrast buttons or select pictorial icons for common phrases such as *"What time is it?"* or *"Play music"* Although current VAs already support limited gesture recognition (e.g., swiping or simple hand raise detection for media control), future systems could expand these capabilities through advanced cameras and motion tracking. Such improvements could enable more natural interactions, allowing users to wave to pause media, point to select on-screen items, or even clap to activate the device, thereby broadening the interaction space beyond voice. Haptic feedback, such as vibrations, can also reinforce these interactions by confirming recognised commands, signalling errors, guiding multi-step tasks, or alerting users to reminders. Together, these multimodal approaches can reduce reliance on voice alone, making VAs more inclusive and usable for individuals with ID.

6 Limitations and Future Work

Our study revealed that although VAs held significant potential to empower certain participants, they also presented considerable challenges for others, highlighting important limitations in their accessibility and usability. One limitation of our study is that the Google Nest Hub and Amazon Echo Show were introduced at different times, which may have influenced participants' preferences and interactions. Moreover, the study took place within the structured environment of a disability support organisation, which may not fully reflect how VAs are used in everyday or more varied real-world contexts (e.g., homes, workplaces, schools). However, the clubhouse like setting provided a supportive and predictable environment that enabled many participants to engage comfortably with the devices.

A promising direction is to involve people with ID more directly in the design of VA tools, using participatory methods that accommodate diverse communication styles and interaction needs so that key design decisions are grounded in their lived experience. Approaches such as co-design workshops, scaffolded choice making, and guided prototyping can enable participants to select key system elements, including wake words, response complexity, and interaction modalities. Recognising the wide diversity of abilities and experiences within the ID population (e.g., vocal articulation, receptive language, working memory, and attention), adopting a competency based perspective that focuses on behaviours demonstrated in real interactions rather than assumed ability levels can help ensure that design decisions reflect actual user capabilities and preferences. Future systems should support configurable or adaptive features that can be tuned to different support needs, rather than assuming that a single VA configuration will suit all participants with ID. Following these participatory design efforts, it will be important to evaluate how the resulting VAs perform in longer deployments, and how participants' familiarity and confidence with the systems evolve over time. Evaluating such systems across a wider range of real-world settings, and examining practical interaction factors such as listening window duration, preferred reading and comprehension levels, and favoured interaction modalities, will help ensure that VAs are tailored to participants' everyday needs and communication preferences.

7 Conclusion

In this paper, we explored how adults with ID interacted with off-the-shelf VAs over a two-month period as part of a STEAM program in a disability support organization. Our three-phase study: comprising of preliminary interviews, deployment interviews, and post-reflection interviews, alongside voice interaction logs and observational data revealed how VAs can support a range of activities including daily routines, entertainment, and learning. Our findings revealed the importance of collaborative learning and peer support in group settings, highlighting how social dynamics influence technology adoption. While VAs offered clear benefits, participants encountered speech and cognition related challenges such as mispronunciations, ambiguous phrasing, soft-spoken input, timeout errors, difficulty processing complex responses, and problems recalling wake words. Addressing these limitations requires VA designs that are more accessible and adaptable to diverse communication needs. Our study emphasizes the value of adaptive features, including auto-correction for atypical speech, extended listening windows, simplified responses, routine support, and multimodal interaction capabilities. We also highlight the importance of proactive and engaging interfaces, including visual prompts, and supportive cues, to sustain attention and promote long-term use. This work advances the design of accessible VA technologies that support communication and engagement for individuals with ID, providing actionable guidance for researchers and practitioners across diverse contexts. More broadly, our findings show that with thoughtful adaptation, VAs can better support users with varied communication profiles by applying universal design principle, helping to reduce barriers to everyday technology use.

Acknowledgments

We thank our participants and coaches from the disability support organization for their time and insights. This research was funded by the Australian Government through the Australian Research

Council (Project No. LP210300465). This work was also supported by the Monash University Faculty of Information Technology International Postgraduate Research Scholarship.

References

[1] Amazon Alexa. 2025. https://alexa.amazon.com/. Accessed: 2025-02-20.

[2] Mohammad Rafayet Ali, Seyedeh Zahra Razavi, Raina Langevin, Abdullah Al Mamun, Benjamin Kane, Reza Rawassizadeh, Lenhart K. Schubert, and Ehsan Hoque. 2020. A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20)*. Association for Computing Machinery, New York, NY, USA, Article 2, 8 pages. doi:10.1145/3383652.3423900

[3] Mrim M Alnfiai and Muhammad Ashad Kabir. 2024. Exploring the impact of word prediction assistive features on smartphone keyboards for blind users. *Heliyon* 10, 17 (2024). doi:10.1016/j.heliyon.2024.e36653

[4] Amazon. 2021. *Amazon Echo Show 8 User Manual*. https://www.bhphotovideo.com/lit_files/782196.pdf Accessed: 2025-04-14.

[5] Amazon. 2025. Alexa Voice History (Personal Account). Data extracted from personal Alexa device logs. Accessed: 2025-04-14.

[6] Aida Amirova, Nazerke Rakhymbayeva, Elmira Yadollahi, Anara Sandygulova, and Wafa Johal. 2021. 10 years of human-nao interaction research: A scoping review. *Frontiers in Robotics and AI* 8 (2021), 744526. doi:10.3389/frobt.2021.744526

[7] Valentina Andries and Judy Robertson. 2023. Alexa doesn't have that many feelings: Children's understanding of AI through interactions with smart speakers in their homes. *Computers and Education: Artificial Intelligence* 5 (2023), 100176. doi:10.1016/j.caeai.2023.100176

[8] Google Assistant. 2025. https://assistant.google.com/. Accessed: 2025-02-20.

[9] American Psychiatric Association. 2013. *Neurodevelopmental Disorders* (5th ed.). American Psychiatric Association. doi:10.1176/appi.books.9780890425787.x01_Neurodevelopmental_Disorders

[10] Sujirat Attawibulkul, Boonserm Kaewkamnerdpong, and Yoshikazu Miyanaga. 2017. Noisy speech training in MFCC-based speech recognition with noise suppression toward robot assisted autism therapy. In *2017 10th Biomedical Engineering International Conference (BMEiCON)*. 1–5. doi:10.1109/BMEiCON.2017.8229135

[11] Ami Bakhai, Aurora Constantin, and Cristina Adriana Alexandru. 2020. Motivateme!: An Alexa skill to support higher education students with autism. In *14th International Conference on Interfaces and Human Computer Interaction*. Internal Association for Development of the Information Society, 187–191. https://www.ihci-conf.org/

[12] Saminda Sundeepa Balasuriya, Laurianne Sitbon, Andrew A. Bayor, Maria Hoogstrate, and Margot Brereton. 2018. Use of voice activated interfaces by people with intellectual disability. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (Melbourne, Australia) *(OzCHI '18)*. Association for Computing Machinery, New York, NY, USA, 102–112. doi:10.1145/3292147.3292161

[13] Ana Berdasco, Gustavo López, Ignacio Diaz, Luis Quesada, and Luis A Guerrero. 2019. User experience comparison of intelligent personal assistants: Alexa, Google Assistant, Siri and Cortana. In *Proceedings*, Vol. 31. MDPI, 51. https://api.semanticscholar.org/CorpusID:214119479

[14] Valérie Bergua, Colette Fabrigoule, Pascale Barberger-Gateau, Jean-François Dartigues, Joel Swendsen, and Jean Bouisson. 2006. Preferences for routines in older people: associations with cognitive and psychological vulnerability. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences* 21, 10 (2006), 990–998. doi:10.1002/gps.1597

[15] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. American Psychological Association. *APA handbook of research methods in psychology* 2 (2012), 71. doi:10.1037/13620-004

[16] Jared M. Bruce, Michael McQuiggan, Vanessa Williams, Holly Westervelt, and Geoffrey Tremont. 2008. Burden among spousal and child caregivers of patients with mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders* 25, 4 (4 2008), 385–390. doi:10.1159/000122587

[17] Inha Cha, Sung-In Kim, Hwajung Hong, Heejeong Yoo, and Youn-kyung Lim. 2021. Exploring the Use of a Voice-based Conversational Agent to Empower Adolescents with Autism Spectrum Disorder. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 42, 15 pages. doi:10.1145/3411764.3445116

[18] Chen Chen, Ella T Lifset, Yichen Han, Arkajyoti Roy, Michael Hogarth, Alison A Moore, Emilia Farcas, and Nadir Weibel. 2023. Screen or No Screen? Lessons Learnt from a Real-World Deployment Study of Using Voice Assistants With and Without Touchscreen for Older Adults. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 52, 21 pages. doi:10.1145/3597638.3608378

[19] Dasom Choi, Daehyun Kwak, Minji Cho, and Sangsu Lee. 2020. "Nobody Speaks that Fast!" An Empirical Study of Speech Rate in Conversational Agents for People with Vision Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376569

[20] Victoria Cluley. 2018. From "Learning disability to intellectual disability"—Perceptions of the increasing use of the term "intellectual disability" in learning disability policy, research and practice. *British journal of learning disabilities.* 46, 1 (2018). doi:10.1111/bld.12209

[21] Sally-Ann Cooper, Gary McLean, Bruce Guthrie, Alex McConnachie, Stewart Mercer, Frank Sullivan, and Jill Morrison. 2015. Multiple physical and mental health comorbidity in adults with intellectual disabilities: population-based cross-sectional analysis. *BMC family practice* 16, 1 (2015), 110. doi:doi:10.1186/s12875-015-0329-3

[22] Andrea Cuadra, Jessica Bethune, Rony Krell, Alexa Lempel, Katrin Hänsel, Armin Shahrokni, Deborah Estrin, and Nicola Dell. 2023. Designing Voice-First Ambient Interfaces to Support Aging in Place. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2189–2205. doi:10.1145/3563657.3596104

[23] Andrea Cuadra, Shuran Li, Hansol Lee, Jason Cho, and Wendy Ju. 2021. My Bad! Repairing Intelligent Voice Assistant Errors Improves Interaction. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 27 (April 2021), 24 pages. doi:10.1145/3449101

[24] Melisa Duque, Sarah Pink, Yolande Strengers, Rex Martin, and Larissa Nicholls. 2021. Automation, wellbeing and digital voice assistants: Older people and Google devices. *Convergence* 27, 5 (2021), 1189–1206. doi:10.1177/13548565211038537

[25] Christiane Even, Torsten Hammann, Vera Heyl, Christian Rietz, Hans-Werner Wahl, Peter Zentel, and Anna Schlomann. 2022. Benefits and challenges of conversational agents in older adults: a scoping review. *Zeitschrift für Gerontologie und Geriatrie* 55, 5 (2022), 381–387. doi:10.1007/s00391-022-02085-9

[26] Jack M. Fletcher and Jeremy Miciak. 2024. Assessment of Specific Learning Disabilities and Intellectual Disabilities. *Assessment* 31, 1 (2024), 53–74. doi:10.1177/10731911231194992

[27] Anne Fota and Hanna Schramm-Klein. 2024. The Impact of Digital Voice Assistants on the Everyday Life of Physically Disabled Users: Barriers, Drivers and Potential for Improvement. (2024). https://scholarspace.manoa.hawaii.edu/items/1d981edf-b981-40fa-8dce-1f0680e205f4

[28] Marta Perez Garcia, Sarita Saffon Lopez, and Hector Donis. 2018. Everybody is talking about virtual assistants, but how are people really using them?. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference* (Belfast, United Kingdom) *(HCI '18)*. BCS Learning & Development Ltd., Swindon, GBR, Article 96, 5 pages. doi:10.14236/ewic/HCI2018.96

[29] Google. 2025. Google My Activity. https://myactivity.google.com/ Accessed: 2025-04-14.

[30] Google. 2025. Introducing Google Nest Hub. https://support.google.com/googlenest/answer/9136909?hl=en Accessed: 2025-04-14.

[31] Rebecca Grist, Joanna Porter, and Paul Stallard. 2017. Mental health mobile apps for preadolescents and adolescents: a systematic review. *Journal of medical internet research* 19, 5 (2017), e176. doi:10.2196/jmir.7332

[32] Ramalingam Hariharan, Jula Hakkinen, and Kari Laurila. 2001. Robust end-of-utterance detection for real-time speech recognition applications. In *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, Vol. 1. IEEE, 249–252. doi:10.1109/ICASSP.2001.940814

[33] Jaylin Herskovitz, Andi Xu, Rahaf Alharbi, and Anhong Guo. 2023. Hacking, Switching, Combining: Understanding and Supporting DIY Assistive Technology Design by Blind People. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 57, 17 pages. doi:10.1145/3544548.3581249

[34] Google Home. 2025. https://assistant.google.com/. Accessed: 2025-02-20.

[35] Anirudh Itagi, Cyril Joe Baby, Subham Rout, KP Bharath, R Karthik, and M Rajesh Kumar. 2019. Lisp detection and correction based on feature extraction and random forest classifier. In *Microelectronics, Electromagnetics and Telecommunications: Proceedings of the Fourth ICMEET 2018*. Springer, 55–64. doi:10.1007/978-981-13-1906-8_6

[36] Dietmar Jakob, Sebastian Wilhelm, Armin Gerl, Diane Ahrens, and Florian Wahl. 2025. Adapting Voice Assistant Technology for Older Adults: A Comprehensive Study on Usability, Learning Patterns, and Acceptance. *Digital* 5, 1 (2025), 4. doi:10.3390/digital5010004

[37] Madeline Jefferson. 2019. Usability of automatic speech recognition systems for individuals with speech disorders: Past, present, future, and a proposed model. (2019). https://api.semanticscholar.org/CorpusID:196178405

[38] Manuel Jesús-Azabal, José Agustín Medina-Rodríguez, Javier Durán-García, and Daniel García-Pérez. 2020. Remembranza pills: Using alexa to remind the daily medicine doses to elderly. In *Gerontechnology: Second International Workshop, IWoG 2019, Cáceres, Spain, September 4–5, 2019, Revised Selected Papers 2*. Springer, 151–159. doi:10.1007/978-3-030-41494-8_15

[39] Sunyoung Kim. 2021. Exploring How Older Adults Use a Smart Speaker–Based Voice Assistant in Their First Interactions: Qualitative Study. *JMIR Mhealth Uhealth* 9, 1 (13 Jan 2021), e20427. doi:10.2196/20427

[40] Robert M. Klassen. 2002. A Question of Calibration: A Review of the Self-Efficacy Beliefs of Students with Learning Disabilities. *Learning Disability Quarterly* 25 (2002), 102 – 88. doi:10.2307/1511276

[41] Andreas M. Klein, Andreas M. Hinderks, Maria Rauschenberger, and Jörg Thomaschewski. 2020. Exploring Voice Assistant Risks and Potential with Technology-based Users. In *Proceedings of the 16th International Conference on Web Information Systems and Technologies - WEBIST*. INSTICC, SciTePress, 147–154. doi:10.5220/0010150101470154

[42] Soonil Kwon, Sung-Jae Kim, and Joon Yeon Choeh. 2016. Preprocessing for elderly speech recognition of smart devices. *Computer Speech & Language* 36 (2016), 110–121. doi:10.1016/j.csl.2015.09.002

[43] Dounia Lahoual and Myriam Frejus. 2019. When Users Assist the Voice Assistants: From Supervision to Failure Resolution. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3290607.3299053

[44] Youngsun Lee, Michael L Wehmeyer, Susan B Palmer, Kendra Williams-Diehm, Daniel K Davies, and Steven E Stock. 2011. The effect of student-directed transition planning with a computer-based reading support program on the self-determination of students with disabilities. *The Journal of Special Education* 45, 2 (2011), 104–117.

[45] Colleen Lennon and Hal Burdick. 2003. *The Lexile Framework for Reading Measurement and Success*. Technical Report. MetaMetrics, Inc., Durham, NC, USA.

[46] Lundy Lewis and Andre Vellino. 2021. Helping Persons with Cognitive Disabilities using Voice-Activated Personal Assistants. In *2021 8th International Conference on ICT & Accessibility (ICTA)*. 1–3. doi:10.1109/ICTA54582.2021.9809777

[47] Lizi Liao, Grace Hui Yang, and Chirag Shah. 2023. Proactive Conversational Agents in the Post-ChatGPT World. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan) *(SIGIR '23)*. Association for Computing Machinery, New York, NY, USA, 3452–3455. doi:10.1145/3539618.3594250

[48] Lanna Lima, Vasco Furtado, Elizabeth Furtado, and Virgilio Almeida. 2019. Empirical Analysis of Bias in Voice-based Personal Assistants. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 533–538. doi:10.1145/3308560.3317597

[49] Amama Mahmood and Chien-Ming Huang. 2024. Gender Biases in Error Mitigation by Voice Assistants. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 60 (April 2024), 27 pages. doi:10.1145/3637337

[50] Fabio Masina, Valeria Orso, Patrik Pluchino, Giulia Dainese, Stefania Volpato, Cristian Nelini, Daniela Mapelli, Anna Spagnolli, and Luciano Gamberini. 2020. Investigating the accessibility of voice assistants with impaired users: Mixed methods study. *Journal of Medical Internet Research* 22, 9 (9 2020). doi:10.2196/18431

[51] Assaf Hurwitz Michaely, Xuedong Zhang, Gabor Simko, Carolina Parada, and Petar Aleksic. 2017. Keyword spotting for Google assistant using contextual speech recognition. In *2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. 272–278. doi:10.1109/ASRU.2017.8268946

[52] Microsoft Corporation. 2024. Microsoft Word. https://www.microsoft.com/en-us/microsoft-365/word. Accessed: 2025-04-14.

[53] N Miller. 1995. Pronunciation errors in acquired speech disorders: The errors of our ways. *International Journal of Language & Communication Disorders* 30, 3 (1995), 346–361. doi:10.3109/13682829509021447

[54] Tuuli H. Morrill, Melissa Michaud Baese-Berk, and Ann Bradlow. 2016. Speaking rate consistency and variability in spontaneous speech by native and non-native speakers of English. *Speech prosody* (2016), 1119–1123. https://api.semanticscholar.org/CorpusID:148168394

[55] William G Moulton. 1962. Toward a classification of pronunciation errors. *The Modern Language Journal* 46, 3 (1962), 101–109. doi:10.1111/j.1540-4781.1962.tb01773.x

[56] Hoang Long Nguyen, Vincent Renkens, Joris Pelemans, Srividya Pranavi Potharaju, Anil Kumar Nalamalapu, and Murat Akbacak. 2021. User-Initiated Repetition-Based Recovery in Multi-Utterance Dialogue Systems. *ArXiv* abs/2108.01208 (2021). https://api.semanticscholar.org/CorpusID:236881134

[57] Kenichiro Noda. 2018. Google Home: smart speaker as environmental control unit. *Disability and rehabilitation: assistive technology* 13, 7 (2018), 674–675. doi:10.1080/17483107.2017.1369589

[58] Jeesun Oh, Wooseok Kim, Sungbae Kim, Hyeonjeong Im, and Sangsu Lee. 2024. Better to Ask Than Assume: Proactive Voice Assistants' Communication Strategies That Respect User Agency in a Smart Home Environment. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 846, 17 pages. doi:10.1145/3613904.3642193

[59] Thierry Olive. 2004. Working Memory in Writing: Empirical Evidence From the Dual-Task Technique. *European Psychologist* 9, 1 (2004), 32–42. doi:10.1027/1016-9040.9.1.32

[60] Sarit Orlofsky and Kathryn Wozniak. 2022. Older adults' experiences using Alexa. *Geriatric Nursing* 48 (2022), 247–257. doi:10.1016/j.gerinurse.2022.09.017

[61] Valentina Ospina-Henao, Sebastián López Flórez, V Juan M Núñez, Óscar Loureda Lamas, and Fernando De la Prieta. 2024. Generative AI: Simplifying Text for Cognitive Impairments and Non-native Speakers. In *International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning*. Springer, 33–44. doi:10.1007/978-3-031-73538-7_4

[62] Parvaneh Parvin, Marco Manca, Caterina Senette, Maria Claudia Buzzi, Marina Buzzi, and Susanna Pelagatti. 2022. Alexism: ALEXa supporting children with autISM in their oral care at home. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI '22)*. Association for Computing Machinery, New York, NY, USA, Article 18, 5 pages. doi:10.1145/3531073.3531157

[63] Alisha Pradhan, Amanda Lazar, and Leah Findlater. 2020. Use of Intelligent Voice Assistants by Older Adults with Low Technology Use. *ACM Trans. Comput.-Hum. Interact.* 27, 4, Article 31 (Sept. 2020), 27 pages. doi:10.1145/3373759

[64] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3174033

[65] Jorge Proenca, Carla Lopes, Michael Tjalve, Andreas Stolcke, Sara Candeias, and Fernando Perdigao. 2018. Mispronunciation Detection in Children's Reading of Sentences. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.* 26, 7 (July 2018), 1203–1215. doi:10.1109/TASLP.2018.2820429

[66] Kathleen Puckett and B O'Bannon. 2012. Technology applications for students with dyslexia. *Essentials of Dyslexia Assessment and Intervention* (2012), 199–222.

[67] Samuel Reinders, Matthew Butler, Ingrid Zukerman, Bongshin Lee, Lizhen Qu, and Kim Marriott. 2025. When Refreshable Tactile Displays Meet Conversational Agents: Investigating Accessible Data Presentation and Analysis with Touch and Speech. *IEEE Transactions on Visualization and Computer Graphics* 31, 1 (2025), 864–874. doi:10.1109/TVCG.2024.3456358

[68] Naba Rizvi, William Wu, Mya Bolds, Raunak Mondal, Andrew Begel, and Imani N. S. Munyaka. 2024. Are Robots Ready to Deliver Autism Inclusion?: A Critical Review. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 69, 18 pages. doi:10.1145/3613904.3642798

[69] Apple Siri. 2025. https://www.apple.com/siri/. Accessed: 2025-02-20.

[70] Elizabeth Smith, Petroc Sumner, Craig Hedge, and Georgina Powell. 2021. Smart speaker devices can improve speech intelligibility in adults with intellectual disability. *International journal of language & communication disorders* 56, 3 (2021), 583–593. doi:10.1111/1460-6984.12615

[71] Elizabeth Smith, Petroc Sumner, Craig Hedge, and Georgina Powell. 2023. Smart-speaker technology and intellectual disabilities: agency and wellbeing. *Disability and Rehabilitation: Assistive Technology* 18, 4 (2023), 432–442. doi:10.1080/17483107.2020.1864670

[72] Ewan Soubutts, Amid Ayobi, Rachel Eardley, Roisin McNaney, Kirsten Cater, and Aisling Ann O'Kane. 2022. Amazon Echo Show as a Multimodal Human-to-Human Care Support Tool within Self-Isolating Older UK Households. 6, CSCW2, Article 302 (Nov. 2022), 31 pages. doi:10.1145/3555193

[73] Molly Follette Story. 2001. Principles of universal design. *Universal design handbook* 2 (2001).

[74] George Terzopoulos and Maya Satratzemi. 2020. Voice assistants and smart speakers in everyday life and in education. *Informatics in Education* 19, 3 (2020), 473–490. https://api.semanticscholar.org/CorpusID:222496598

[75] Milka Trajkova and Aqueasha Martin-Hammond. 2020. "Alexa is a Toy": Exploring Older Adults' Reasons for Using, Limiting, and Abandoning Echo. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376760

[76] Pooja Upadhyay, Sharon Heung, Shiri Azenkot, and Robin N. Brewer. 2023. Studying Exploration & Long-Term Use of Voice Assistants by Older Adults. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 848, 11 pages. doi:10.1145/3544548.3580925

[77] E Van Wingerden, SV Vacaru, L Holstege, and PS Sterkenburg. 2023. Hey Google! Intelligent personal assistants and well-being in the context of disability during COVID-19. *Journal of Intellectual Disability Research* 67, 10 (2023), 973–985. doi:10.1111/jir.13064

[78] Ana Vitória Lachowski Volochtchuk, Higor Leite, and Alessandro Diogo Vieira. 2024. Voice assistant technology applied to populations with developmental and physical disabilities. *Behaviour & Information Technology* 43, 11 (2024), 2300–2322. doi:10.1080/0144929X.2023.2243343

[79] Sholom M Weiss, Nitin Indurkhya, Tong Zhang, et al. 2010. *Fundamentals of predictive text mining*. Vol. 41. Springer. doi:10.1007/978-1-84996-226-1

[80] Carolin Wienrich, Astrid Carolus, André Markus, Yannik Augustin, Jan Pfister, and Andreas Hotho. 2023. Long-Term Effects of Perceived Friendship with Intelligent Voice Assistants on Usage Behavior, User Experience, and Social Perceptions. *Comput.* 12 (2023), 77. https://api.semanticscholar.org/CorpusID:258152645

[81] Tamara Zubatiy, Kayci L Vickers, Niharika Mathur, and Elizabeth D Mynatt. 2021. Empowering Dyads of Older Adults With Mild Cognitive Impairment And Their Care Partners Using Conversational Agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 657, 15 pages. doi:10.1145/3411764.3445124