

Article

Interacting with Smart Virtual Assistants for Individuals with Dysarthria: A Comparative Study on Usability and User Preferences

Aisha Jaddoh ^{1,2,*} , Fernando Loizides ¹ , Omer Rana ^{1,*}  and Yasir Ahmed Syed ¹ 

¹ School of Computer Science and Informatics, Cardiff University, Cardiff CF24 4AG, UK; loizidesf@cardiff.ac.uk (F.L.); syedy@cardiff.ac.uk (Y.A.S.)

² Computer Science and Engineering Department, Yanbu Industrial College, Yanbu 46411, Saudi Arabia

* Correspondence: jaddoha@cardiff.ac.uk (A.J.); ranaof@cardiff.ac.uk (O.R.)

Abstract: This study explores the effectiveness and user experience of different interaction methods used by individuals with dysarthria when engaging with Smart Virtual Assistants (SVAs). It focuses on three primary modalities: direct speech commands through Alexa, non-verbal voice cues via the Daria system, and eye gaze control. The objective is to assess the usability, workload, and user preferences associated with each method, catering to the varying communication capabilities of individuals with dysarthria. While Alexa and Daria facilitate voice-based interactions, eye gaze control offers an alternative for those unable to use voice commands, including users with severe dysarthria. This comparative approach aims to determine how the usability of each interaction method varies, conducted with eight participants with dysarthria. The results indicated that non-verbal voice interactions, particularly with the Daria system, were favored because of their lower workload and ease of use. The eye gaze technology, while viable, presented challenges in terms of the higher workload and usability. These findings highlight the necessity of diversifying interaction methods with SVAs to accommodate the unique needs of individuals with dysarthria.

Keywords: usability; dysarthria; smart virtual assistants; accessibility; assistive technology



Citation: Jaddoh, A.; Loizides, F.; Rana, O.; Syed, Y.A. Interacting with Smart Virtual Assistants for Individuals with Dysarthria: A Comparative Study on Usability and User Preferences. *Appl. Sci.* **2024**, *14*, 1409. <https://doi.org/10.3390/app14041409>

Academic Editor: Tamás Grósz

Received: 31 December 2023

Revised: 25 January 2024

Accepted: 6 February 2024

Published: 8 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Smart virtual assistants (SVAs) are transforming how we interact with Internet of Things devices and electronic services hosted on or accessed through such devices. These devices utilize natural language processing as a means to support interaction. They can be integrated into cell phones or operating systems (e.g., Siri on iOS devices or Cortana on Windows) or exist as standalone devices, such as Google Home and Amazon Alexa [1,2].

Users can employ these devices to perform various tasks [3–5], such as controlling smart homes by adjusting lights or turning on and off appliances or retrieving information by inquiring about the news, traffic, or weather. Interacting with SVAs through natural language can be beneficial for a variety of groups of people. For instance, those who have low technical literacy may find it simpler to interact with such devices [6]. In addition, people who have a disability can also benefit; for example, individuals who have limited mobility can control their homes using their voices [7], and those who have visual impairments or limited dexterity can interact with an SVA without the need for an intermediary device [8].

However, not all people who have a disability can use SVAs [9]. For example, people who have speech impairments may face difficulty using these devices [10–12] given that they are designed and trained to understand non-impaired speech. Indeed, the difficulty in using such devices increases with the severity of the speech impairment [13]. Providing access to SVAs can greatly improve the well-being of affected individuals by offering them independence [14], a means of communication [15], social inclusion [16], and safety [17]. Multiple modalities (beyond speech and voice) are needed to interact with and use SVAs.

One such approach involves the use of augmentative and alternative communication (AAC) systems [18]. These systems assist people who have speech difficulties to convey their messages using alternative methods, such as a keyboard, mouse, joystick, or other suitable interfaces on a tablet [19]. AAC systems encompass a range of technologies designed to aid communication, especially for individuals who have speech or language difficulties. These technologies include symbol- or text-based systems, speech-generating devices, and computer-based applications [20,21]. In addition, eye gaze technology enables users to control an AAC device connected to a voice user interface simply by directing their gaze [22].

Prior studies have significantly contributed to the understanding of assistive technologies' usability and workload. For example, Pasqualotto et al. [23] compared access technologies, highlighting the need for user-friendly and low-workload solutions for individuals who have severe motor impairments. Similarly, other studies [22] have demonstrated their potential in enhancing smartphone authentication and smart home control for users who have disabilities using eye gaze interaction. These insights emphasize the necessity of designing human–computer interaction (HCI) technologies that are not only accessible but also align with the users' cognitive and physical capabilities.

In this study, we focus on dysarthria, a neurological motor speech impairment that hinders proper speech production, causing slow and weak speech muscle movements that lead to poor articulation, difficulty coordinating breathing and speaking, and low speech intelligibility. Furthermore, dysarthria is often accompanied by physical disabilities, leading to limitations in interactions with assistant devices [15,24].

This article aims to compare the usability of various interaction methods:

- (i) Direct speech commands through Alexa (version 2.2). Alexa is a widely known hands-free smart voice assistant device developed by Amazon [25]. Choosing Alexa was informed by its status as the most widely used device globally for natural language processing and voice-activated assistance [26]. It operates primarily through speech recognition and natural language processing to understand and respond to voice commands (voice commands as input and voice replies or actions performed as output). This interaction method involves using speech, in which users directly send commands to the SVA by uttering a sentence command, for example, "Alexa, what is the weather today?"
- (ii) Nonverbal voice cues through the Daria system [15,27]. We refer to this system as "Daria", an easy to pronounce name. Further, all the letters from Daria are in "DysARthrIA" and in the same order. This choice was informed by emerging research indicating the potential of nonverbal vocalizations in enhancing interaction for individuals who have speech impairments [15,27]. Daria is a custom-developed system that allows for interaction with SVAs by using nonverbal voice cues, offering a more straightforward, shorter, and less fatiguing alternative to traditional speech commands. For example, users can simply make the sound /a/ ("aaa") to turn on lights, which is significantly simpler than uttering complex sentences such as "Alexa, turn on the lights". Daria is programmed using five distinct nonverbal voice cues, each mapped to a specific action. This mapping includes /a/ for lights, /i/ for news, /ɪ/ to initiate a call, humming for music, and /u/ for weather updates. This design ensures ease of control and enhanced accessibility, particularly for users who have severe dysarthria, allowing them to perform a variety of tasks using minimal effort. Prior studies have been conducted on the design of Daria [15,27], underscoring its primary goal of empowering individuals who have dysarthria. The system's design involved collaboration with individuals diagnosed with dysarthria, ensuring that Daria is sensitively and effectively attuned to their specific communication challenges and preferences.
- (iii) Eye gaze control. This method employs eye gaze control by which users control a tablet connected to the SVA using only their eyes.

These methods were chosen for their potential to accommodate the varying communication capabilities of individuals who have dysarthria, ensuring a broad and inclusive approach to interaction with SVAs. Understanding the strengths and limitations of each method enables us to enhance and optimize these technologies. Usability measures for each modality were assessed according to the International Organization for Standardization (ISO), which defines usability as the “extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [28]. Furthermore, the workload required for each interaction was also evaluated.

Finally, this study aimed to evaluate and compare the usability of voice commands, nonverbal voice cues, and eye gaze interactions for individuals who have dysarthria. By doing so, it sought to offer comprehensive insights into the most effective and user-friendly methods of interaction for this population, thus contributing to the development of more accessible and efficient communication technologies.

2. Background

Individuals who have dysarthria interact with voice technologies differently than those who do not have speech disorders. The primary issue in this interaction is the performance of these systems, which tends to deteriorate for people who have dysarthria, worsening as the severity of the condition increases. Prior studies have indicated that these commercial devices require further improvements to better accommodate users who have dysarthria [13,19,29,30]. One primary challenge lies in their accuracy, particularly in understanding commands issued with dysarthric speech. Variability in volume and pitch adds another layer of complexity [19,31]. For example, fluctuations in volume and pitch within a single word or sentence can confuse these systems, making it difficult to accurately capture the intended command [32]. Moreover, these devices may time out before the user finishes speaking [19,31]. Another issue arises owing to the unique characteristics of dysarthric speech, which often includes breaths between syllables. Current devices struggle to handle this as a form of input [30].

As an alternative to verbal interactions, another technique that aligns with the capabilities of users who have dysarthria has been suggested. This alternative involves using nonverbal voice cues for interaction [15,27]. In interviews conducted with 19 participants who had dysarthria [15], it was found that this method of interaction was well accepted by the individuals. This acceptance reflected their willingness and interest in using any potential mode of interaction that may streamline their experience. A pilot study testing nonverbal voice cues focused on system design and the memorability aspects of usability. The findings indicated that the participants could effectively recall the nonverbal cues after a period and such cues could be a viable alternative to traditional verbal commands.

Although SVAs are controlled mainly through voice, it has been suggested that they could also be controlled through other modalities to overcome the challenges faced by people who have dysarthria [12]. This study evaluates interactions with SVAs using eye gaze interaction, because people who have dysarthria usually do not suffer from issues with eye movement or visual impairment, so using eye gaze technology is an appropriate practical option [33]. Eye gaze technology works by tracking eye movement to determine the eye gaze position, enabling users to perform a variety of actions. For example, such a device could allow a user to turn on a light by looking at a “light” button. This could occur through detecting dwell time or blinking, depending on the eye tracking device setting.

To date, there is a relatively small body of literature concerned with the usability of eye gaze systems when used by people who have speech impairments in general and dysarthria specifically [34–36]. Donegan [37] investigated some aspects of usability by exploring user satisfaction and how eye gaze can affect the quality of life for people who have disabilities. The study found that users were satisfied with the use of eye gaze and it had a positive impact on their daily life. Work by Najafi [38] also evaluated the use of eye gaze devices. In addition to obtaining user feedback about the use of the device,

it evaluated the issues that may arise and the adjustments that were required to allow for efficient use. Hemmingeoon and Borgestig [34] used surveys to assess the usability of eye gaze technology for people who have physical and communication impairments. The findings showed that most of the participants were satisfied with using eye gaze to control computers and it was efficient to use.

3. Methods

Dysarthria's unique communication challenges necessitate the exploration of alternative methods for interacting with SVAs. Given the variable severity of dysarthria, it was imperative to evaluate various interaction methods to determine their effectiveness across this spectrum. Each method represented a distinct mode of interaction, providing valuable insights into which modes are most accessible and user friendly for people who have varying degrees of speech impairment.

In this study, we measured several key attributes to assess the interaction methods, including the usability, effectiveness, satisfaction, and workload:

- The usability attribute measured the user's ease and efficiency in interacting with the system. This was measured by the system usability scale (SUS), a widely used tool for testing usability [39] that has been used across various domains, including the usability of SVAs [40–44]. This survey comprises 10 questions rated on a 5-point Likert scale that ranges from strongly disagree to strongly agree.
- The effectiveness attribute measured the user's ability to complete a task (task success rate) [28]. The task was considered successful if the SVA successfully replied to or performed the command requested. The success was recorded during the study and confirmed by video recordings. Although the SUS provides subjective user feedback, the effectiveness attribute offers objective concrete data on how well a system performs in achieving its intended tasks [45] and many studies have used it in combination with the SUS [45–48].
- The preference for each system was evaluated through direct feedback from a post-study interview, which focused specifically on system preference as an indicator of satisfaction [49]. This approach complemented the other measures and provided a deeper understanding of the user feedback.
- The workload identified the effort required to perform a task. This was measured using the NASA Task Load Index (NASA-TLX) questionnaire [50], which contained six questions focusing on mental demand, physical demand, temporal demand, performance, effort, and frustration level. Using this measure was particularly crucial for individuals who have dysarthria and often experience rapid fatigue. By employing the NASA-TLX, we aimed to gain a deeper understanding of the workload implications for this specific user group [41,44,51].

3.1. Participants

Eight participants who had dysarthria participated in this study. All the participants were patients at Sultan Bin Abdulaziz Humanitarian City. Participants were adults whose ages ranged from 18 to 65 years. None of the participants had cognitive issues, ensuring that their responses and interactions with the systems were solely influenced by their dysarthria condition. The severity of dysarthria for these individuals was provided by their speech and language therapists. To maintain consistency and reliability across assessments, all therapists employed the same standardized assessment known as the "Motor Speech Assessment". Although all participants had experience using voice technologies, none had prior exposure to the Daria system or eye tracking systems. Table 1 provides details about the participants.

Table 1. Participants details.

Participant	Gender	Severity	Age Range	Diagnosis
P1	Male	Mild	25–44	Traumatic brain injury
P2	Male	Mild	45–65	Stroke
P3	Female	Mild	25–44	Cerebral palsy
P4	Male	Moderate	45–65	Spinal cord injury
P5	Male	Moderate	25–44	Traumatic brain injury
P6	Male	Severe	25–44	Stroke
P7	Male	Severe	25–44	Traumatic brain injury
P8	Male	Severe	18–24	Traumatic Brain Injury

3.2. Setup and Equipment

The study occurred in a clinic at the medical city. To test verbal interactions, we used Alexa on a cell phone, primarily because it supported the Arabic language and all participants were Arabic speakers. To test nonverbal voice interactions, we employed the Daria system. Finally, to test eye gaze interactions, we used Tobii Eye Tracker 4C, an off-the-shelf eye tracker from a leading eye tracking company [52]. Tobii Eye Tracker 4C is compatible with Windows PCs and easy to use. The tracker was magnetically affixed to the bottom of the laptop screen. For the eye tracking interactions, the key components were the eye tracker device and a HTML web page containing buttons, each of which represented a command. This page was connected to Raspberry Pi through the RabbitMQ message broker. The user interface of this HTML page was designed in accordance with guidelines that recommended large buttons to ensure ease of selection and interaction [53].

During the study, participants were asked to instruct the devices to perform five tasks: turn on the lights, play music, play the news, call someone, and ask about the weather. These tasks were selected because they were the most commonly used by individuals who had dysarthria, as indicated in [15]. First, participants verbally asked Alexa to perform these five tasks. The order of the tasks was randomized and varied for each participant.

Next, participants switched to Daria and gave commands using nonverbal voice cues. Given that participants were less familiar with this system, we provided a brief description of Daria and introduced the commands for the expected actions (the same as those for Alexa).

Finally, participants tested the eye gaze system. When they used the eye tracker, a bar containing buttons appeared at the top of the screen, as shown in Figure 1. A red circle, functioning as a cursor (see Figure 1), also appeared, positioned on the second item from the left on the bar, which participants controlled using their eyes. Participants were instructed on how to use the tracker, including identifying which buttons represented the left-click mouse function and confirm button. The users were required to select the mouse click button through dwell time. Once this bar disappeared, they needed to use the cursor (see Figure 2) to choose one of the boxes, each of which represented a command, that appeared on the screen. Once the user pointed to one of the boxes, it was highlighted, and the action was performed.

After each of the three parts, participants completed the SUS and the NASA-TLX questionnaire. Following completion, post-study interviews were conducted to ask about their preference among the three systems.

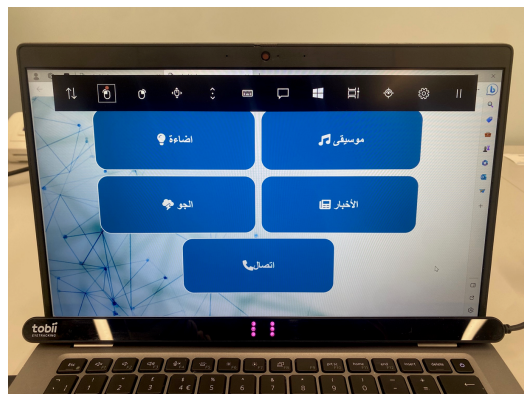


Figure 1. Eye tracker control bar. The text on the buttons appearing on the screen are as follows, starting from the first line, first button on the left: Light, Music, Weather, News, and Call.

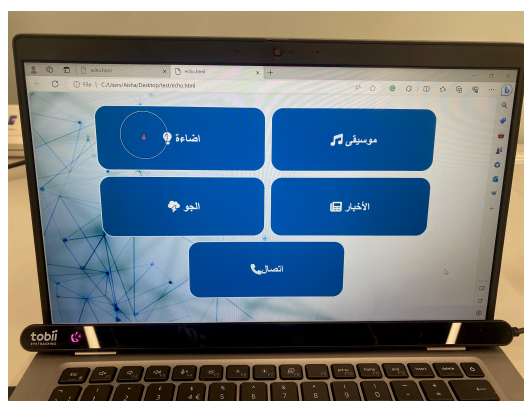


Figure 2. Selecting a command. The text on the buttons appearing on the screen are as follows, starting from the first line, first button on the left: Light, Music, Weather, News, and Call.

4. Results

4.1. SUS

The SUS result for Alexa was 79.06, which, according to the SUS rating scale, is equivalent to “Good”. For the Daria system, the score was 84.68, which is also equivalent to “Good”. Finally, for the eye gaze system, the score was 52.81, which the evaluation only rates as “OK”.

To compare the three approaches, a statistical analysis was conducted using the Friedman test, which is suitable for small sample sizes and comparing the same subjects. The results indicated that there was a significant overall difference between the options. To better understand these differences, we conducted a pairwise statistical analysis using the Wilcoxon test. This test revealed significant differences when comparing the eye gaze approach with the Daria system ($p = 0.011$), favoring the Daria system. Moreover, there was a significant difference between the eye gaze approach and Alexa ($p = 0.011$), in which Alexa was favored. There was no significant difference between the Daria system and Alexa. To further understand how these usability scores relate to the workload experienced by participants during the interactions, we turn to the NASA-TLX assessment.

4.2. Workload

Given that NASA-TLX is commonly used to compare results between tasks, we used this survey to evaluate the differences between the three systems. In this survey, a lower score indicated less workload and effort, which translated to better results. The averages are presented in Figure 3. Starting with the mental demand, Daria had the least demand, followed closely by Alexa and then the eye gaze interaction, which showed a significant difference to the other two. In terms of the physical demand, Alexa scored the lowest,

followed by Daria, which had only a slight difference. However, the eye gaze interaction imposed a notably higher level of demand. For the temporal demand, the pattern was similar to that of the mental demand: Daria scored the lowest, followed by Alexa and then the eye gaze interaction. In terms of the performance, which is how successful the user was in accomplishing what they were asked to do, Alexa had the lowest score, then Daria, showing a marginal difference, and the eye gaze interaction, which had a significantly higher level of demand. Finally, considering the effort and frustration, Daria had the lowest score, followed by Alexa and then the eye gaze interaction.

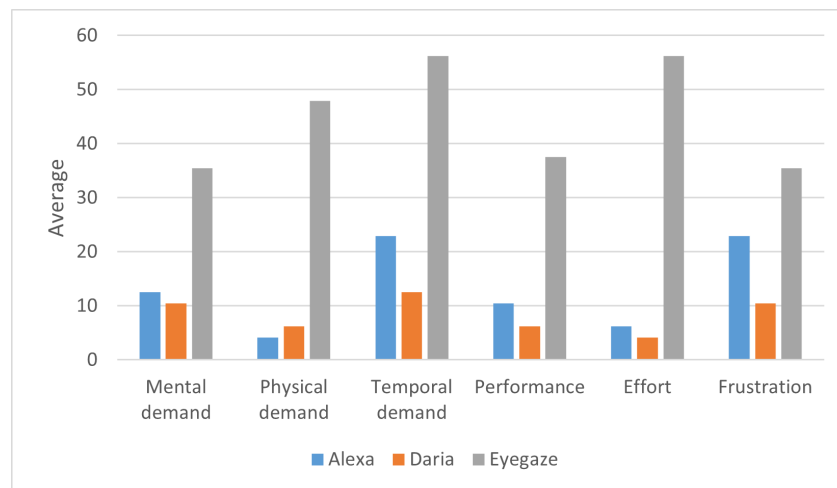


Figure 3. NASA-TLX workload.

Similar to the statistical analysis for the SUS, the Friedman test was used to determine whether there were statistically significant differences in the workload scores across the three interaction methods (see Table 2). The test revealed statistically significant differences in the physical demand ($p = 0.003$), performance ($p = 0.040$), and effort ($p = 0.011$) whereas no significant differences were found for the mental demand ($p = 0.174$), temporal demand ($p = 0.054$), and frustration ($p = 0.244$).

Table 2. Workload—significance between the three interaction systems. Friedman’s ANOVA was used for overall comparison. The significance level is 0.05.

Category	System	Mean Rank	Chi-Square	p -Value	p -Value Assessment
Mental demand	Daria	2.25	3.50	0.174	Not Significant
	Eyegaze	1.63			
	Alexa	2.13			
Physical demand	Daria	2.31	11.47	0.003	Significant
	Eyegaze	1.25			
	Alexa	2.44			
Temporal demand	Daria	2.44	5.85	0.054	Not Significant
	Eyegaze	1.38			
	Alexa	2.19			
Performance	Daria	2.31	6.42	0.040	Significant
	Eyegaze	1.44			
	Alexa	2.25			
Effort	Daria	2.50	8.96	0.011	Significant
	Eyegaze	1.25			
	Alexa	2.25			
Frustration	Daria	2.38	2.82	0.244	Not Significant
	Eyegaze	1.69			
	Alexa	1.94			

To understand the specific differences between each pair of methods, we conducted a Wilcoxon test for the categories that had significant Friedman results (see Table 3). For the physical demand, the pairwise comparison showed significant differences in favor of Alexa over the eye gaze interaction ($p = 0.024$) and Daria over the eye gaze interaction ($p = 0.025$), indicating that Daria and Alexa required less physical effort than the eye gaze method. Similarly, for the temporal demand, significant differences were noted between Daria and the eye gaze interaction ($p = 0.020$), in which Daria required less time to perform tasks. For the performance, there was a significant difference between Daria and the eye gaze interaction ($p = 0.041$). The participants found themselves to be more successful in performing the tasks using Daria. For the effort, a significant difference was found again in favor of Daria over the eye gaze interaction ($p = 0.018$), suggesting that Daria interactions demand less effort from users. Similarly, Alexa required less effort than the eye gaze method ($p = 0.042$).

These findings suggest that for individuals who have varying degrees of speech impairment, nonverbal and verbal voice command methods (Daria and Alexa) may impose a lower workload and be more accessible than eye gaze interaction methods. However, it is noteworthy that no significant differences were found between Daria and Alexa, indicating that the two voice-based interaction methods performed similarly in terms of the workload.

4.3. Task Success Rate

A successful interaction occurs when a command is executed by the SVA as requested by the user. If the device detected the command, regardless of the number of attempts, it was counted as a success. Given that there were five commands and eight users, the total number of successful attempts would have been 40 if all the commands were successful for each user. When interacting with Daria, 38 interactions were successful; for Alexa, 33 were successful; and for the eye gaze method, 13 were successful.

4.4. Preference

The participants were asked to share their preferences across the three systems. Five participants preferred the Daria system, citing its ability to accurately understand their utterances and the ease of use. One participant specifically appreciated that it did not require pronouncing challenging letters, such as "R".

Two participants favored Alexa, including one who had moderate and one who had mild dysarthria, explaining that they were comfortable with it and capable of articulating words and sentences using this system. However, for some participants, using Alexa was not preferred because continuous speech was found to be tiring, particularly for those who had more severe forms of dysarthria in which prolonged speaking can be physically demanding.

Meanwhile, one participant (moderate dysarthria) preferred the eye gaze interaction, valuing the option to interact without the need to use their voice. However, two participants (P5 and P1) reported that the eye gaze system was not their preferred choice because of the discomfort caused by the laser from the tracker, which was uncomfortable or even painful for their eyes. In addition, the effort required to accurately control the eye gaze system was mentioned.

This feedback sheds light on the diverse experiences and preferences of the participants with each interaction modality. A breakdown of the participants' preferences and their respective diagnoses is provided in Table 4.

Table 3. Pairwise workload statistical analysis and significance. The significance level is 0.05.

Category	System	Mean	p-Value	p-Value Assessment
Mental demand	Daria Eye gaze	93.75 43.75	0.223	Not Significant
	Daria Alexa	93.75 77.08	0.317	Not Significant
	Eye gaze Alexa	43.75 77.08	0.223	Not Significant
Physical demand	Daria Eye gaze	93.75 52.08	0.025	Significant
	Daria Alexa	93.75 95.83	0.317	Not Significant
	Eye gaze Alexa	52.08 95.83	0.024	Significant
Temporal demand	Daria Eye gaze	89.58 64.58	0.020	Significant
	Daria Alexa	89.58 87.50	0.285	Not Significant
	Eye gaze Alexa	64.58 87.50	0.205	Not Significant
Performance	Daria Eye gaze	93.75 62.50	0.041	Significant
	Daria Alexa	93.75 89.58	0.655	Not Significant
	Eye gaze Alexa	62.50 89.58	0.242	Not Significant
Effort	Daria Eye gaze	95.83 43.75	0.018	Significant
	Daria Alexa	95.83 93.74	0.564	Not Significant
	Eye gaze Alexa	43.75 93.74	0.042	Significant
Frustration	Daria Eye gaze	89.58 64.58	0.068	Not Significant
	Daria Alexa	89.58 77.08	0.461	Not Significant
	Eye gaze Alexa	64.58 77.08	0.498	Not Significant

Table 4. Participant preferences for assistive communication systems.

Preference	Participant	Diagnosis
Alexa	P1	Mild
	P5	Moderate
Daria	P2	Mild
	P3	Mild
	P6	Severe
	P7	Severe
Eye gaze	P8	Severe
	P4	Moderate

5. Discussion

This study contributes significantly to the field of HCI and the accessibility of HCI technologies. By examining three distinct HCI methods—direct speech commands, non-verbal voice cues via the Daria system, and eye gaze interactions—our study not only reveals their effectiveness, usability, and participant preferences but also provides a com-

prehensive comparison of these methods. These findings are highly valuable for future researchers in this field and contribute to the development of more inclusive and accessible communication technologies.

The diverse preferences expressed by the participants in our study revealed a nuanced picture of interactions with SVAs. Our findings indicate a preference for the Daria system among most of the participants, which is attributed to its ease of use and adeptness at understanding commands. This preference was particularly notable among the participants who had severe dysarthria, suggesting that Daria's design is well suited to users who have significant speech impairments. This finding shows how using nonverbal voice cues, which is within users' capabilities, aligns with Wobbrock's principles, specifically, ability-based design principles [54], for creating systems in accordance with the strengths and capabilities of users, thereby enhancing accessibility. However, Alexa was preferred by the participants who had milder forms of dysarthria, indicating its effectiveness for users who can articulate clearer speech patterns. The eye gaze interaction was uniquely valued by a participant who had moderate dysarthria, highlighting its potential as an alternative communication method for those who find voice-based interaction challenging.

These preferences correlate with our findings on usability, which indicate that SVAs that use verbal or nonverbal commands are more usable than those using eye gaze interactions. This increased usability arises from the relative ease of speaking and the ability of the device to understand speech. In addition, prior studies, such as that of [55,56], have indicated that voice interactions are closely aligned with natural human communication patterns. Further, users who have dysarthria prefer to use their voice to the maximum extent. However, this finding contradicts that of [22], who found that participants rated the usability of eye gaze interactions with SVAs as exceptional, providing an average SUS score of 92.5. However, the limited scope of this study, which focused on a single user who had a disability, raises questions about the generalizability of the findings. Another study, [23], found that users who had a motor disability (but provided no information on their speech ability) gave eye gaze interactions an average SUS score of 78.54, which is higher than our result but lower than that of [22]. A broader participant base in future studies could offer more comprehensive insights into the usability of eye gaze systems.

The alignment of the user preferences with usability scores in our study resonates with the technology acceptance model [57] and the unified theory of the acceptance and usage of technology [58]. These models emphasize the ease of use and effort expectancy as critical factors in technology adoption. This is confirmed by our findings, in which participants gravitated toward systems that offered a greater ease of use and less effort, reflecting a natural inclination toward technologies that align with their individual abilities and communication preferences.

In addition, interacting through voice is likely to be a more intuitive and natural method [59], even for individuals who have impaired speech capabilities [56]. However, it is important to consider the influence of the participants' lack of prior experience with the eye tracking device and the Daria system. None of the participants had previously used these systems, introducing significant factors that may have affected the system usability, including the intuitiveness required to use the system and the learning curve associated with unfamiliar technologies [60]. Although the Daria system, which uses nonverbal voice commands, relies on the inherent familiarity most individuals have with vocal communication, making it more intuitive, eye tracking systems may require a steeper learning curve because of their unconventional interaction mode. Therefore, these factors may influence the overall usability of each system for first-time users [61], underscoring the importance of considering the novelty and intuitiveness of HCI technologies in their evaluation.

These findings are further supported by our workload results, which offer important insights into the experience of users who have dysarthria when interacting with various technologies. The data show that eye gaze interactions involve considerably more effort across several dimensions. This higher level of workload suggests that although eye gaze interactions remain a viable option for individuals who have dysarthria, especially those

who have severe cases, the extensive demands may affect this technology's practicality and acceptance for long-term use. This has been noted in prior studies [62], suggesting that the burdensome nature of eye gaze interactions may extend to other populations who have similar challenges.

Further, the perceived effectiveness and reduced effort associated with voice-based interactions suggest a higher likelihood of long-term acceptance and use. Their ease of use and lower physical and mental demands position these methods as more sustainable and practical for individuals who have dysarthria. This aligns with the broader goal of assistive technologies, which is to enhance the quality of life through user-friendly and efficient solutions [63]. Therefore, our findings underscore the critical need to consider the workload and user effort as key factors in the design and implementation of HCI methods, specifically, in terms of assistive technologies for individuals who have dysarthria.

These findings are also reflected in the success rate, which was higher for the interactions that were more usable and required less effort. The participants who had varying levels of dysarthria severity preferred using their voices to interact with the SVAs. This finding aligns with that of [55,56], who found that users prefer using their voices as much as possible.

Our study offers valuable insights into the interaction preferences of individuals with dysarthria. It serves as a foundation for further research in this area. To build on these initial findings, future studies could benefit from exploring a wider range of participant experiences, enhancing the generalizability and depth of the research. Additionally, investigating the learning curve associated with different interaction systems, especially for users new to eye gaze technology, would provide a more comprehensive understanding of how the user familiarity impacts effectiveness.

6. Conclusions

This study provides important contributions to the field of HCI for individuals who have dysarthria by elucidating the effectiveness, usability, and user preferences of three distinct interaction methods. Our findings distinctly highlight the Daria system as the preferred interaction method for the majority of the participants, especially those who had severe dysarthria, owing to its ease of use and effective command recognition. This underlines the potential of nonverbal voice cues in enhancing accessibility for users who have significant speech impairments.

However, the study revealed that the participants who had milder forms of dysarthria favored the voice-activated systems, such as Alexa, indicating their suitability for those who can articulate clearer speech patterns. This preference emphasizes the need for HCI technologies to cater to varying levels of speech ability.

Furthermore, the eye gaze interaction method, although identified as more effort intensive than the other two methods, emerged as a vital alternative for the users who have severe dysarthria and are unable to use voice-based systems. This finding is crucial because it highlights the importance of including diverse interaction methods in assistive technologies to accommodate the broad spectrum of user needs.

The study's insights into how individuals who have dysarthria interact with various types of SVAs contribute significantly to understanding the factors influencing usability and user experience in this domain. These insights are invaluable for guiding the development of future technologies to better meet the diverse communication needs of individuals who have dysarthria across the spectrum of impairment severity.

Author Contributions: Conceptualization, A.J. and F.L.; methodology, A.J. and F.L.; software, A.J. and F.L.; validation, A.J., F.L., O.R. and Y.A.S.; formal analysis, A.J., F.L., O.R. and Y.A.S.; investigation, A.J.; resources, A.J.; data curation, A.J.; writing—original draft preparation, A.J. and F.L.; writing—review and editing, A.J., F.L., O.R. and Y.A.S.; visualization, A.J.; supervision, F.L., O.R. and Y.A.S.; project administration, A.J. and F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Cardiff University, and approved by the Institutional Review Board (or Ethics Committee) of Cardiff University, School of Computer Science and Informatics.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is contained within the article.

Acknowledgments: We would like to express our acknowledgment and appreciation for the support provided by the Sultan Bin Abdulaziz Humanitarian City, which helped in conducting this study with their patients. Moreover, special thanks go to the research center and to the SLP team for their support. We would also like to thank all the participants who took part in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kepuska, V.; Bohouta, G. Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 99–103.
2. Hoy, M.B. Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Med. Ref. Serv. Q.* **2018**, *37*, 81–88. [CrossRef]
3. Bentley, F.; Luvogt, C.; Silverman, M.; Wirasinghe, R.; White, B.; Lottridge, D. Understanding the long-term use of smart speaker assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 91. [CrossRef]
4. Moore, C. OK, Google: What Can Home do? The Speaker's Most Useful Skills. Available online: <https://www.digitaltrends.com/home/google-home-most-useful-skills/> (accessed on 15 December 2023).
5. Ammari, T.; Kaye, J.; Tsai, J.Y.; Bentley, F. Music, search, and IoT: How people (really) use voice assistants. *ACM Trans. Comput.-Hum. Interact.* **2019**, *26*, 17. [CrossRef]
6. Sciarretta, E.; Alimenti, L. Smart speakers for inclusion: How can intelligent virtual assistants really assist everybody? In *Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, 24–29 July 2021, Proceedings, Part I 23*; Springer: Cham, Switzerland, 2021; pp. 77–93.
7. Masina, F.; Orso, V.; Pluchino, P.; Dainese, G.; Volpato, S.; Nelini, C.; Mapelli, D.; Spagnoli, A.; Gamberini, L. Investigating the accessibility of voice assistants with impaired users: Mixed methods study. *J. Med Internet Res.* **2020**, *22*, e18431. [CrossRef]
8. Corbett, E.; Weber, A. What can I say? addressing user experience challenges of a mobile voice user interface for accessibility. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, Florence, Italy, 6–9 September 2016; pp. 72–82.
9. Morris, J.T.; Thompson, N.A.; Center, S. User personas: Smart speakers, home automation and people with disabilities. *J. Technol. Pers. Disabil.* **2020**, *8*, 237–256.
10. Takashima, Y.; Takiguchi, T.; Arika, Y. End-to-end dysarthric speech recognition using multiple databases. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 6395–6399.
11. Pradhan, A.; Mehta, K.; Findlater, L. “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, 21–26 April 2018; pp. 1–13.
12. Masina, F.; Pluchino, P.; Orso, V.; Ruggiero, R.; Dainese, G.; Mameli, I.; Volpato, S.; Mapelli, D.; Gamberini, L. VOICE Actuated Control Systems (VACS) for accessible and assistive smart homes. A preliminary investigation on accessibility and user experience with disabled users. In *Proceedings of the Ambient Assisted Living: Italian Forum 2019*; Springer: Cham, Switzerland, 2021; pp. 153–160.
13. De Russis, L.; Corno, F. On the impact of dysarthric speech on contemporary ASR cloud platforms. *J. Reliab. Intell. Environ.* **2019**, *5*, 163–172. [CrossRef]
14. Teixeira, A.; Braga, D.; Coelho, L.; Fonseca, J.; Alvarelhão, J.; Martín, I.; Queirós, A.; Rocha, N.; Calado, A.; Dias, M. Speech as the basic interface for assistive technology. In Proceedings of the DSAI 2009, 2th International Conference on Software Development for Enhancing Accessibility and Fighting Info-Exclusion, Lisboa, Portugal, 3–5 June 2009.
15. Jaddoh, A.; Loizides, F.; Lee, J.; Rana, O. An interaction framework for designing systems for virtual home assistants and people with dysarthria. *Univers. Access Inf. Soc.* **2023**, 1–13. [CrossRef]
16. Fried-Oken, M.; Fox, L.; Rau, M.T.; Tullman, J.; Baker, G.; Hindal, M.; Wile, N.; Lou, J.S. Purposes of AAC device use for persons with ALS as reported by caregivers. *Augment. Altern. Commun.* **2006**, *22*, 209–221. [CrossRef]
17. Beukelman, D.R.; Mirenda, P. *Augmentative and Alternative Communication*; Paul H. Brookes: Baltimore, MD, USA, 1998.
18. Bryen, D.N.; Chung, Y. What adults who use AAC say about their use of mainstream mobile technologies. *Assist. Technol. Outcomes Benefits* **2018**, *12*, 73–106.
19. Ballati, F.; Corno, F.; De Russis, L. Assessing virtual assistant capabilities with italian dysarthric speech. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 93–101.

20. Curtis, H.; Neate, T.; Vazquez Gonzalez, C. State of the Art in AAC: A Systematic Review and Taxonomy. In Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility, Athens, Greece, 23–26 October 2022; pp. 1–22.
21. Matters, C. TYPES OF AAC. Available online: <https://www.communicationmatters.org.uk/what-is-aac/types-of-aac/> (accessed on 11 December 2023).
22. Bissoli, A.; Lavino-Junior, D.; Sime, M.; Encarnaç o, L.; Bastos-Filho, T. A human–machine interface based on eye tracking for controlling and monitoring a smart home using the internet of things. *Sensors* **2019**, *19*, 859. [CrossRef]
23. Pasqualotto, E.; Matuz, T.; Federici, S.; Ruf, C.A.; Bartl, M.; Olivetti Belardinelli, M.; Birbaumer, N.; Halder, S. Usability and workload of access technology for people with severe motor impairment: A comparison of brain-computer interfacing and eye tracking. *Neurorehabilit. Neural Repair* **2015**, *29*, 950–957. [CrossRef]
24. Ansel, B.M.; Kent, R.D. Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *J. Speech Lang. Hear. Res.* **1992**, *35*, 296–308. [CrossRef] [PubMed]
25. Alexa. Available online: <https://www.amazon.com/b?node=21576558011> (accessed on 15 December 2023).
26. Statista.com. Number of Households with Smart Home Products and Services in Use Worldwide from 2017 to 2025. Available online: <https://www.statista.com/statistics/1252975/smart-home-households-worldwide/> (accessed on 12 December 2023).
27. Jaddoh, A.; Loizides, F.; Rana, O. Non-verbal interaction with virtual home assistants for people with dysarthria. *J. Technol. Pers. Disabil.* **2021**, *9*, 71–84.
28. ISO 9241-11:2018; Ergonomic Requirements for Office Work with Visual Display Terminals (VDT)s- Part 11 Guidance on Usability. ISO: Geneva, Switzerland, 2018.
29. Ballati, F.; Corno, F.; De Russis, L. “Hey Siri, Do You Understand Me?”: Virtual Assistants and Dysarthria. In Proceedings of the International Workshop on the Reliability of Intelligent Environments (Workshops), Rome, Italy, 25–28 June 2018; pp. 557–566.
30. Moore, M.; Venkateswara, H.; Panchanathan, S. Whistle-blowing asrs: Evaluating the need for more inclusive automatic speech recognition systems. In Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Hyderabad, India, 2–6 September 2018; Volume 2018, pp. 466–470.
31. Derboven, J.; Huyghe, J.; De Grooff, D. Designing voice interaction for people with physical and speech impairments. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, Helsinki, Finland, 26–30 October 2014; pp. 217–226.
32. Moore, M. Speech Recognition for Individuals with Voice Disorders. In *Multimedia for Accessible Human Computer Interfaces*; Springer: Cham, Switzerland, 2021; pp. 115–144.
33. Corno, F.; Farinetti, L.; Signorile, I. A cost-effective solution for eye-gaze assistive technology. In Proceedings of the IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland, 26–29 August 2002; Volume 2; pp. 433–436.
34. Hemmingsson, H.; Borgestig, M. Usability of eye-gaze controlled computers in Sweden: A total population survey. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1639. [CrossRef]
35. Caligari, M.; Godi, M.; Guglielmetti, S.; Franchignoni, F.; Nardone, A. Eye tracking communication devices in amyotrophic lateral sclerosis: impact on disability and quality of life. *Amyotroph. Lateral Scler. Front. Degener.* **2013**, *14*, 546–552. [CrossRef]
36. Karlsson, P.; Allsop, A.; Dee-Price, B.J.; Wallen, M. Eye-gaze control technology for children, adolescents and adults with cerebral palsy with significant physical disability: Findings from a systematic review. *Dev. Neurorehabilit.* **2018**, *21*, 497–505. [CrossRef]
37. Donegan, M.; Morris, J.D.; Corno, F.; Signorile, I.; Chi o, A.; Pasian, V.; Vignola, A.; Buchholz, M.; Holmqvist, E. Understanding users and their needs. *Univers. Access Inf. Soc.* **2009**, *8*, 259–275. [CrossRef]
38. Najafi, L.; Friday, M.; Robertson, Z. Two case studies describing assessment and provision of eye gaze technology for people with severe physical disabilities. *J. Assist. Technol.* **2008**, *2*, 6–12. [CrossRef]
39. Gil-G omez, J.A.; Manzano-Hern andez, P.; Albiol-P erez, S.; Aula-Valero, C.; Gil-G omez, H.; Lozano-Quilis, J.A. USEQ: A short questionnaire for satisfaction evaluation of virtual rehabilitation systems. *Sensors* **2017**, *17*, 1589. [CrossRef]
40. Kocabalil, A.B.; Laranjo, L.; Coiera, E. Measuring user experience in conversational interfaces: A comparison of six questionnaires. In Proceedings of the 32nd International BCS Human Computer Interaction Conference, Belfast, UK, 4–6 July 2018.
41. Vtyurina, A.; Fournay, A. Exploring the role of conversational cues in guided task support with virtual assistants. In Proceedings of the 2018 CHI conference on human factors in computing systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–7.
42. Pyae, A.; Joelsson, T.N. Investigating the usability and user experiences of voice user interface: A case of Google home smart speaker. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, Barcelona, Spain, 3–6 September 2018; pp. 127–131.
43. Bogers, T.; Al-Basri, A.A.A.; Ostermann Rytlig, C.; Bak M oller, M.E.; Juhl Rasmussen, M.; Bates Michelsen, N.K.; Gerling J rgensen, S. A study of usage and usability of intelligent personal assistants in Denmark. In *Information in Contemporary Society, Proceedings of the 14th International Conference, iConference 2019, Washington, DC, USA, 31 March–3 April 2019*; Proceedings 14; Springer: Cham, Switzerland, 2019; pp. 79–90.
44. Anbarasan; Lee, J.S. Speech and gestures for smart-home control and interaction for older adults. In Proceedings of the 3rd International Workshop on Multimedia for Personal Health and Health Care, Seoul, Republic of Korea, 22 October 2018; pp. 49–57.
45. Kortum, P.; Peres, S.C. The relationship between system effectiveness and subjective usability scores using the System Usability Scale. *Int. J. Hum.-Comput. Interact.* **2014**, *30*, 575–584. [CrossRef]
46. Demir, F.; Kim, D.; Jung, E. Hey Google, Help Doing My Homework: Surveying Voice Interactive Systems. *J. Usability Stud.* **2022**, *18*, 41–61.

47. Iannizzotto, G.; Bello, L.L.; Nucita, A.; Grasso, G.M. A vision and speech enabled, customizable, virtual assistant for smart environments. In Proceedings of the 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 4–6 July 2018; pp. 50–56.
48. Barricelli, B.R.; Fogli, D.; Iemmolo, L.; Locoro, A. A multi-modal approach to creating routines for smart speakers. In Proceedings of the 2022 International Conference on Advanced Visual Interfaces, Frascati, Italy, 6–10 June 2022; pp. 1–5.
49. Frøkjær, E.; Hertzum, M.; Hornbæk, K. Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, The Hague, The Netherlands, 1–6 April 2000; pp. 345–352.
50. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.
51. Kim, S.; Ko, I.Y. A Conversational Approach for Modifying Service Mashups in IoT Environments. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April–5 May 2022; pp. 1–16.
52. Tobii. Tobii Eye Tracker. Available online: <https://www.tobii.com/> (accessed on 11 December 2023).
53. Feit, A.M.; Williams, S.; Toledo, A.; Paradiso, A.; Kulkarni, H.; Kane, S.; Morris, M.R. Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 1118–1130.
54. Wobbrock, J.O.; Kane, S.K.; Gajos, K.Z.; Harada, S.; Froehlich, J. Ability-based design: Concept, principles and examples. *ACM Trans. Access. Comput.* **2011**, *3*, 9. [[CrossRef](#)]
55. Patel, R.; Dromey, C.; Kunov, H. *Control of Prosodic Parameters by an Individual with Severe Dysarthria*; Technical Report; University of Toronto: Toronto, ON, Canada, 1998.
56. Ferrier, L.; Shane, H.; Ballard, H.; Carpenter, T.; Benoit, A. Dysarthric speakers’ intelligibility and speech characteristics in relation to computer speech recognition. *Augment. Altern. Commun.* **1995**, *11*, 165–175. [[CrossRef](#)]
57. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–340. [[CrossRef](#)]
58. Venkatesh, V.; Thong, J.Y.; Xu, X. Unified theory of acceptance and use of technology: A synthesis and the road ahead. *J. Assoc. Inf. Syst.* **2016**, *17*, 328–376. [[CrossRef](#)]
59. Munteanu, C.; Jones, M.; Oviatt, S.; Brewster, S.; Penn, G.; Whittaker, S.; Rajput, N.; Nanavati, A. We need to talk: HCI and the delicate topic of spoken language interaction. In *CHI’13 Extended Abstracts on Human Factors in Computing Systems*; Association for Computing Machinery: New York, NY, USA, 2013; pp. 2459–2464.
60. Wang, Y. Gaps between continuous measurement methods: A longitudinal study of perceived usability. *Interact. Comput.* **2021**, *33*, 223–237. [[CrossRef](#)]
61. Kabacińska, K.; Vu, K.; Tam, M.; Edwards, O.; Miller, W.C.; Robillard, J.M. “Functioning better is doing better”: Older adults’ priorities for the evaluation of assistive technology. *Assist. Technol.* **2023**, *35*, 367–373. [[CrossRef](#)]
62. Wang, K.; Wang, S.; Ji, Q. Deep eye fixation map learning for calibration-free eye gaze tracking. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, Charleston, SC, USA, 14–17 March 2016; pp. 47–55.
63. Arthanat, S.; Bauer, S.M.; Lenker, J.A.; Nochajski, S.M.; Wu, Y.W.B. Conceptualization and measurement of assistive technology usability. *Disabil. Rehabil. Assist. Technol.* **2007**, *2*, 235–248. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.