

Handling sensory disabilities in a smart society

The case of a simulated social robot

Julie Pivin-Bachler^{a,*}, Egon L. van den Broek^a and Randy Gomez^b

^a *Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands*
E-mails: julie@pivin-bachler.eu, vandenbroek@acm.org

^b *Honda Research Institute, Honda, Japan*
E-mail: r.gomez@jp.honda-ri.com

Received 25 September 2023

Accepted 17 February 2024

Abstract. Billions of people live with visual and/or hearing impairments. Regrettably, their access to systems remains delayed, leaving them socially excluded. The need for universal access of next-generation systems and users' inclusion is paramount. We pose that a smart society should respond to this crucial need. Following ability-based design principles, we introduce a simulated social robot that adapts to users' sensory abilities. Its working was assessed via a Rock–Paper–Scissors game in an Intelligent Environment (IE), using three modes: where the user is able to see and hear, only see, or only hear. With this game, two user-studies were conducted using the UMUX-LITE usability score, an expectation rating, and the gap between experience and expectation, complemented with two open questions. A repeated measures Multivariate ANalysis Of VAriance (MANOVA) on the data from study 1 unveiled an overall difference between the three modes, $F(6, 6) = 6.823$, $\eta_p^2 = .872$, $p = .017$. Users expected applications to be harder to use with a disability, especially a visual impairment. All modes were considered accessible, with the experience exceeding expectations for the mode with a hearing impairment. In parallel, substantial variance was observed across participants and the results from the open questions suggested improvements. To reduce this variance and increase system stability, study 2 was run with an enhanced design. A repeated measures MANOVA on the data from study 2 confirmed study 1's findings, $F(6, 6) = 12.801$, $\eta_p^2 = .928$, $p = .003$. Moreover, experiences exceeded expectations in all modes and the variance among participants was substantially decreased. We conclude that IE applications managed by a social robot can be adapted to user's sensory abilities, improving smart society's accessibility, and, hence, reducing social exclusion.

Keywords: Accessibility, sensory disabilities, social exclusion, Intelligent Environment, social robots

1. Introduction

According to the World Health Organization, more than 1.5 billion people currently experience some degree of hearing loss and at least 2.2 billion people live with a vision impairment. As the aging population continues to grow, both numbers are expected to increase, with a projected 2.5 billion people with hearing loss by 2050, and a growing prevalence of various eye conditions [62,63]. Hearing loss and visual impairments frequently result in

* Corresponding author. E-mail: julie@pivin-bachler.eu.

social isolation and loneliness [12,48]. This is due to information deprivation which stems from limited access to essential information and communication that those without impairments typically enjoy [12]. It became even more evident during the COVID pandemic, where individuals with sensory impairments experienced more emotional distress and social isolation in part because of unadapted technologies [32,37]. Such facts reveal the importance to take into account individuals with diverse sensory abilities when developing new technologies [49]. The same application should be made as accessible to someone with an impairment as it is to someone without one; thus, preventing social exclusion [15,24,53].

Digital technology is taking humanity through its longest phase of evolution on a socio-economic level, since the birth of computers to the current omnipresence of connected items such as smartphones, tablets, laptops, or smartwatches [28]. The issue relative to the achievement of universal access in technological advances lies in the persistent delay in providing access to individuals with disabilities. The thought of universal access is rarely incorporated during the initial conception stages of new technological devices; instead it tends to be deferred until after the product is finished [1,9,16]. This delay presents two major problems. Firstly, individuals with disabilities are excluded during this time as they cannot use the system, which creates a digital divide [29,44]. Secondly, introducing accessibility features post-development is considerably more challenging than if it had been taken into consideration from the start. Consequently, the resultant accessibility measures often fall short of achieving equal accessibility for individuals with disabilities and those without [9]. Similarly, older adults, who usually experience a loss in their physical and/or cognitive abilities, do not have the same acceptance of technology as younger adults, highlighting once again the need to think of accessibility for users with diverse capacities [26,38]. On a more optimistic note, technology keeps evolving, and we are now surrounded not only by isolated devices, but by networks of devices communicating together, which offer real opportunities for universal access. This is commonly referred to as ambient intelligence, internet of things, smart, or Intelligent Environments (IE) [5].

IE have been an area of research for around two decades [5,20], and are now tangible with the availability of embedded systems (i.e. smartwatches) and virtual assistants in our own homes, such as Google Home or Alexa from Amazon. IE are defined as physical spaces enriched with sensors, associated with an ambient intelligence that deals with the information gathered from these sensors [5]. The components of an IE are orchestrated in order to interact with the spaces' occupant(s) in a sensible way, and enhance their experience. With the current rapid technological evolution, IE are expected to become omnipresent, making universal access one of the seven grand challenges for living and interacting in such technology-augmented environments [22,42,50]. Therefore, it is crucial to incorporate the needs, requirements, and preferences of all individual users when designing applications in IE [13,40]. If these elements are constantly ignored, dire consequences could follow, such as the social exclusion for certain users [21]. Indeed, the fast development of new technologies leads to a digital divide where vulnerable people, usually older adults and people with disabilities, do not have access to new technologies because of their disabilities. To prevent this gap from getting larger, accessibility and inclusion are paramount when designing new applications in IE [43].

Box 1.1 (Definitions of sensory disabilities, adopted from [2]).

- **Sensory disability** usually refers to the impairment of the senses such as sight, hearing, taste, touch, smell, and/or spatial awareness. It mainly covers conditions of visual impairment, blindness, hearing loss, and deafness.
- **Visual impairment:** Decrease or severe reduction in vision that cannot be corrected with standard glasses or contact lenses and reduces an individual's ability to function at specific or all tasks.
- **Blindness:** Profound inability to distinguish light from dark or the total inability to see.
- **Hearing loss:** Decrease in hearing sensitivity of any level.
- **Deafness:** Profound or total loss of hearing in both the ears.

Note: In this study, we use a binary approach towards sensory abilities: participants either have the ability to see or hear, or they do not. So, to convey information, the system cannot rely on visuals in case of a visual impairment (i.e., from low vision to blindness) or on sounds in case of a hearing impairment (i.e., from hearing loss to deafness).

IE need to provide accessibility for users with varying sensory abilities from the moment they are released. This will avoid the waiting of post-development accessibility features or the release of assistive technologies tailored to the needs of each user to be able to use applications in IE. To tackle this problem, we propose that using a social robot as intelligent agent in the IE to select appropriate modalities of interaction can secure accessibility for users with varying sensory abilities. We present an interactive application with the social robot Haru [25] in its virtual environment. The system adapts to the user's sensory abilities by adjusting the modalities employed (cf. [52]). Previous work with Haru made use of multimodality for the robot's perception and understanding of the surrounding environment [47]. Here, we especially focused on multimodality for a suitable communication between the social robot and the user. Our aim was to perform a user-study to verify the accessibility of the application when confronted with different sensory disabilities, namely hearing loss and vision impairment, as defined in Box 1.1. This paper sheds light on how IE with social robots can be used to promote accessibility, and so, from the conception phase of new systems. Section 2 presents the related work in this field. Section 3 introduces the selected application, including a presentation of the social robot Haru. Subsequently, Sections 4 and 5 delineate the initial user-study and its follow-up counterpart, encompassing a description of the adapted design's implementation, the proceedings of each user-study, and the resultant findings. Lastly, Sections 6 and 7 summarize and discuss the key outcomes of the present work, and explores potential avenues for future advancements in the field.

2. Related work

The design of new technologies should always take into account the diversity of human abilities. Nevertheless, interviews of senior user-experience professionals showed that considerations linked to accessibility mainly focus on vision-related challenges and are often "somewhat superficial". Many also reported that their formal education did not give them the necessary and appropriate knowledge and tools for accessibility considerations [46]. Even navigating unfamiliar websites can still be a challenge nowadays for individuals with visual impairments [7].

When new technological devices are created, the conception and development often include assumptions about the user's abilities. Only individuals aligning with these initial assumptions can effectively use the devices once they are out, therefore excluding a significant portion of the user population [9,60,61]. After a certain delay, accessibility features post-development are introduced, which is more challenging from a development point of view and rarely achieve equitable accessibility for users with disabilities compared to the users without any [9]. Elements from traditional assistive technologies can help the user "compensate" for the abilities the system assumes they have in order for the user to be able to use the system. In this scenario, the user is the one adapting to the system [60,61]. In contrast, ability-based design refrains from making any initial assumptions on the user's abilities by providing a system that can adapt to various ones [60,61]. Given ability-design is the only approach among the three described that addresses accessibility for users with diverse sensory disabilities from the conception phase, we employ this principle in this paper. This involves configuring the system to employ distinct modalities for communication, adapting to the user's specific sensory abilities. Figure 1 provides a summary of these three different approaches.

Concerning IE, application designs that cater to sensory disabilities have been previously considered. Usually, the presented application targets one specific sensory disability: either visual or hearing impairments. For instance, AmIE [30] is an Ambient Intelligent Environment directed to users with visual impairments for indoor positioning and navigation. Another example is Apollo SignSound [17], which was designed to integrate sign language in IE for users suffering from hearing loss. There are some rare cases where several disabilities are addressed at once, mainly through guidelines on how to create an accessible application. Obrenović et al. [41] and Kartakis et al. [31] proposed respectively a methodological framework and tools to assist designers in developing universally accessible user interfaces. However, to the best of our knowledge, applications in IE that considers the requirements of users with various sensory disabilities have not been presented [57]. When considering applications with social robots, the same assessment holds. Most of the existing applications can only be operated by users with no sensory disability or considering a specific one, with only limited research focusing on several at once. Mixed abilities are occasionally considered, but refer to the same sense. An illustration is Metatla et al. [39], where an educational game that fits the requirements of both children with and without visual impairments is proposed. The game includes children with

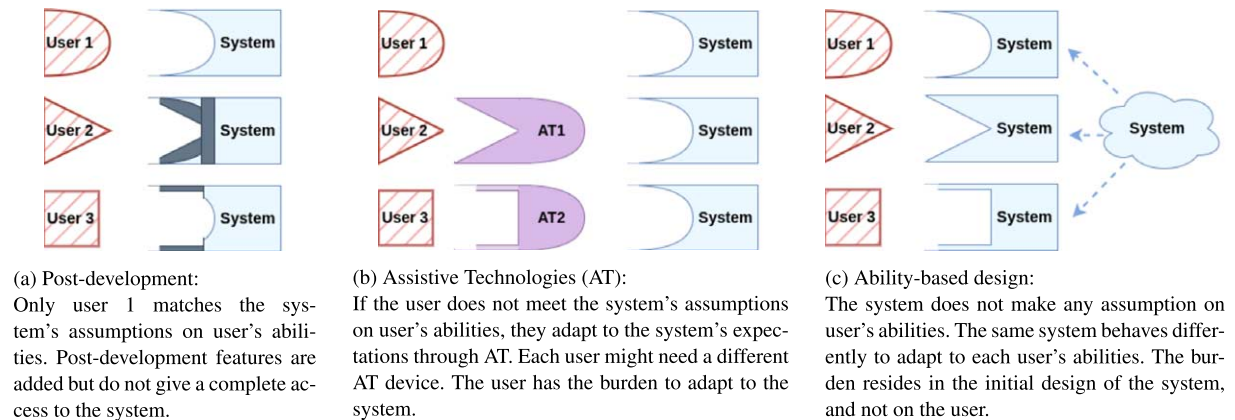


Fig. 1. Three different approaches for accessibility in a system: (a) post-development accessibility features, (b) adaptations from the user to align with the system requirements via Assistive Technologies (AT), and (c) ability-based design. All users can have various sensory abilities. In this figure, we represent different sensory abilities through shapes, with three users.

mixed abilities, but solely linked to the sense of sight. In this paper, we include three various sensory abilities: visual impairment, hearing impairment, or no impairment.

On the one hand, intelligent agents in IE can be seen as butlers, aware of the environment and their owner's (user) attributes, and acting on their behalf [51]. The owner's attributes could include their sensory abilities, and the butler-agent should be able to act accordingly. These agents can be embodied and take a robotic form. Users expect a social robot to act as an assistant or once again, as a butler, rather than a friend [18]. Adaptability to the user has an important role in human acceptance towards social robots, thus giving a social robot the possibility to take into account its owner's capacities and adapt to them is key for long-term use [19]. Intelligent agents and robotic agents are expected to be part of IE [35]: they can be used for intelligence in healthcare [56], for well-being [54], and they could prove useful to promote accessibility [50,52]. On the other hand, IE imply technological richness which allows for multimodal applications, both in terms of inputs from the user and outputs from the system, to create suitable interactions [40,50]. This variety offered by multimodality has the potential to lead to universal access in IE [14]. For instance, older adults, who usually experience some degree of vision loss, could benefit from multimodal devices [64]. Following these two lines of thought, we propose that using an intelligent agent in the form of a social robot as a butler managing the IE, and interacting with the user through different modalities is helpful for accessibility [59].

3. Design

This section introduces the social robot Haru [25] from the Honda Research Institute, with its appearance and capabilities. It also presents the chosen application to implement where multimodal interactions are possible and the general proceedings. The design and general proceedings described in this section are identical for both user-studies presented in Sections 4 and 5.

3.1. The social robot Haru

The main element of the IE is the social robot Haru [25]. Haru is a tabletop robot able to speak, hear, see with the help of cameras, and show different emotional states. The physical robot was not available to the researchers at this time; therefore, Haru is used in its virtual environment. As we are investigating the benefits of an intelligent agent managing an IE for the application's inputs and outputs methods, having the agent in its embodied or virtual form does not alter its management of the IE.

Figure 2 demonstrates that Haru's eyes can help strengthen the expression of affective states. Only happy and sad are shown in Fig. 2, but the possibilities are numerous: surprised, guilty, bored, angry, . . . All tasks were implemented

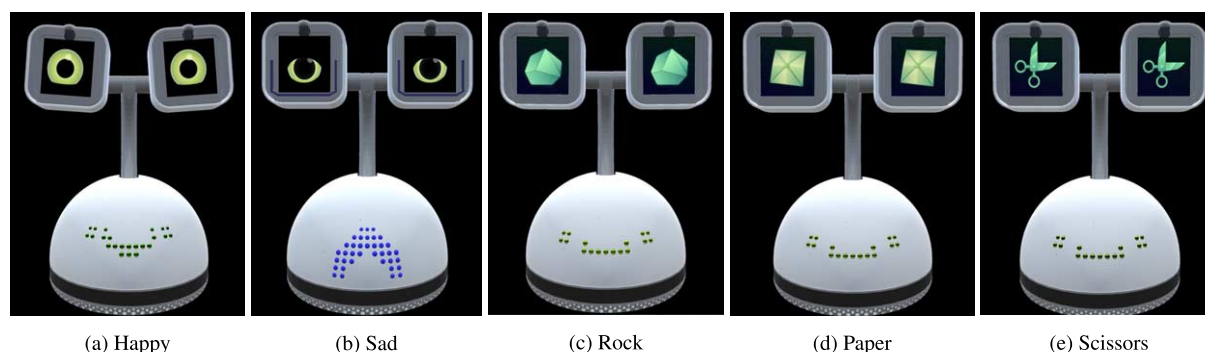


Fig. 2. Eyes' display of the social robot Haru to express emotional states (a, b) and to show its choice in the Rock–Paper–Scissors game (c, d, e).

on Haru with the Robot Operating System (ROS) and Python code. The ROS architecture – in terms of nodes, services, and topics – used for Haru's virtual environment is identical to the one existing for the physical robot. In other words, all the motions, facial expressions, and behaviours demonstrated by the robot in the virtual environment can equally be achieved with the actual robot. In an IE, the robot should be enhanced with cameras to see the user and keep eye contact (cf. [58]), multiple screens could be used as well. For this prototype, the screens are implemented by different windows that appear around Haru in the virtual environment. Instead of the cameras we would use in the IE, the embedded camera from the computer with the virtual environment is used. In this prototype, Haru, acting as manager of the virtual IE, receives information from the camera, the microphone, and the clicks on the computer. To communicate with the user and convey information, Haru is able to interact through visuals – in its eyes or on the various windows on the screen –, sounds, and facial expressions.

3.2. Rock–Paper–Scissors

Following the goal of designing an application with a social robot accessible to users with various sensory abilities, the first choice to make concerns the game to implement. We chose the Rock–Paper–Scissors game for several reasons. First, it is a relatively simple application which suits us as the complexity of the application is of no interest for this research. Second, to play this game, communication between the user and the robot must go both ways as each player must be able to communicate with the other. A lot of information must be conveyed to the user: they must understand what the robot is playing, and they must be able to easily keep track of the score. Last, in previous work, Haru was implemented with the Rock–Paper–Scissors game to be evaluated as a telepresence robot to enhance social interaction between two physically separated users [11]. Therefore, all the necessary designs to play the game with Haru were already available and were kept. Indeed, Haru's eyes can be used as small screens to display elements such as Rock, Paper, or Scissors, as shown in Figs 2c, 2d, and 2e.

Rock–Paper–Scissors has already been implemented on robots [3,27]. Ahn et al. [3] used a four-fingered robotic hand to play the game, associated with a camera to recognize the hand motion of the participant. Hasuda et al. [27] used a social robot, also using a camera to recognize the user's hand motions. The robot responds with facial expressions depending on the outcome of the game. In both examples, vision and hearing are essential to be part of the game and understand its proceedings. Therefore, the game cannot be played by a user living with visual or hearing impairments. The aim of the present research was to demonstrate that a social robot acting as manager of an IE can allow users with various sensory abilities to experience an application with the same ease of use. To that effect, we created a Rock–Paper–Scissors game with an ability-based design where the social robot – in its virtual environment – manages the inputs and outputs of the IE to adapt to the sensory abilities (namely sight and hearing) of its user.

This application begins by a short introduction from Haru where the proceedings of the game are explained, then the game starts. Each round consists of both the user and the robot simultaneously making a choice between Rock, Paper, and Scissors after a countdown from 3 to 1. Once both choices are stated, the winner of the round is determined. The rules are as follows: Rock smashes Scissors, Scissors cuts Paper, and Paper covers Rock. The user

Table 1
Adaptation of the first design for user perception (top part) and interaction (bottom part)

User's abilities	See and hear	Only see	Only hear
Score board	✓	✓	✗
Subtitles	✗	✓	✗
Face tracking	✓	✓	✗
Haru says the score	✗	✗	✓
Haru shows its choice	✓	✓	✗
Haru states its choice	✗	✗	✓
Buttons board	✓	✓	✗
Speech recognition	✗	✗	✓

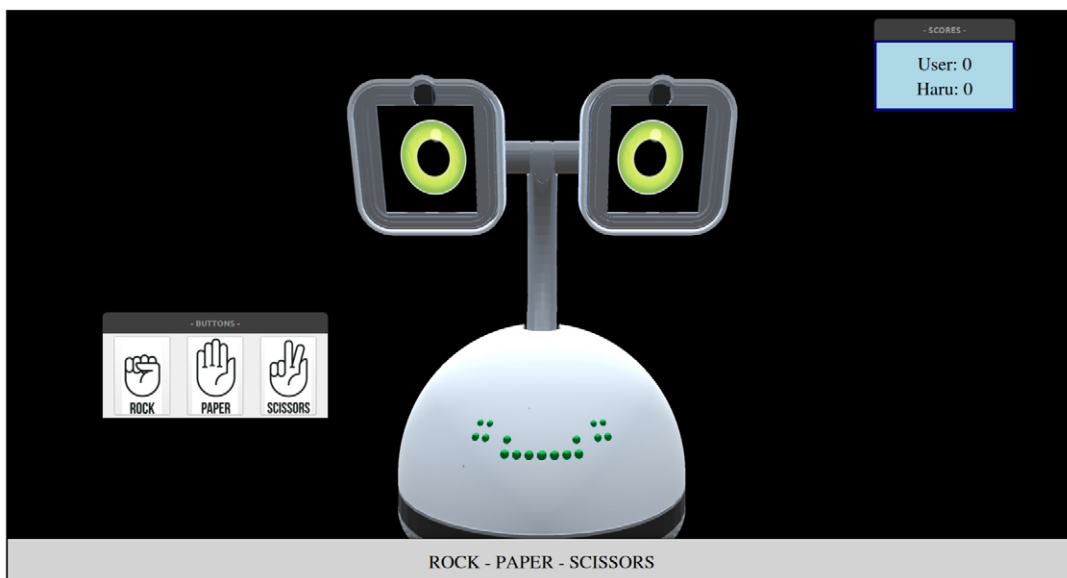


Fig. 3. First design of the Rock–Paper–Scissors application with Haru’s virtual environment for a user only able to see and not hear. For a user able to see and hear, the interface is similar, without subtitles. For a user only able to hear and not see, a black screen is displayed.

can, after each round, either decide to continue or to stop the game. Before finishing, Haru states the final score and the winner of the game.

4. Study 1: Multimodality for perception and interaction

In this section, the first user study is presented. It includes the adapted design of the application, the study design, the obtained results, and a discussion. It corresponds to our work published in [45]. We extend our previous work with a second user-study in Section 5, performed with an improved design of the application.

4.1. Adapted design

Three modes of perception are considered in this study: the user is able to see and to hear, to see but not to hear, or to hear but not to see. The proceedings of the game vary accordingly to these distinct modes, see Table 1.

The application stays very similar when the user is able to see and hear, and when the user is only able to see as most information is conveyed through visuals. The only difference is the subtitles added at the bottom of the screen when the user cannot hear, see Fig. 3. Throughout the game, Haru speaks to the user – through sound or subtitles –,

and expresses different feelings: Haru looks happy – see Fig. 2a – when it wins a round or the game, and sad – see Fig. 2b – when it loses.

For each round, the user makes their choice by clicking on the buttons board on the left side of the screen. At the same time the user clicks on the button, Haru displays its own choice through its eyes as shown in Figs 2c, 2d, and 2e. The interface continuously displays the current score on the top right corner of the screen. Additionally, face tracking is implemented with the camera so that Haru can always be looking in the direction of the participant. To decide whether to continue or stop the game, the user is presented with a pop-up window containing a *Yes* or *No* alternative to pursue the game.

When the user is only able to hear, the interactions through clicks are replaced with speech recognition (with the Google speech-to-text API in Python), with the user stating their choices aloud. Haru states the current score aloud at the beginning of each round. To express its feelings, Haru uses small vocalizations sounding either happy or sad.

4.2. Methods

A within-subjects user-study with 12 participants was conducted. All participants shared the same work profile. All were part of the department of Information and Computing Sciences from Utrecht University, Utrecht, the Netherlands. All interactions in the application were in English in which all participants were fluent.

The three modes – see and hear (SH), see but not hear (S), and hear but not see (H) – were tried by each participant with a randomized order of the modes. There are 6 possible combinations for the modes' order, each done by 2 participants. All participants from the study were able to see and to hear. Therefore the sound from the computer was cut for mode S and the application displayed a black screen for mode H.

Two standardized usability questionnaires and two open questions were used to provide insights on the usability of the application for the different modes, see Table 2. The Usability Metric for User Experience (UMUX-LITE) is a 2-item standardized questionnaire, which rates the usability of a particular application, the higher the score being the better [36]. Both items are rated on a 7-point Likert scale, from “Strongly disagree” (1) to “Strongly agree” (7). To aid interpretation, the scores of these items need to be transformed in order to correspond to the scores from the System Usability Scale (SUS) questionnaire [36]. For that purpose, the following linear regression from [36] is applied:

$$\text{UMUX-LITE} = 0.65 \cdot \left[[(item_1 - 1) + (item_2 - 1)] \cdot \frac{100}{12} \right] + 22.9, \quad (1)$$

where $item_1$ and $item_2$ are the scores on the 7-point Likert scale for each item of the UMUX-LITE. The operations performed inside the square brackets allow to have a result in the interval [0,100]. The rest of the equation is applied to compensate for the small but statistically significant difference that was found in previous work between the UMUX-LITE and the full UMUX version of the questionnaire [36]. The UMUX-LITE adjusted score therefore ranges between 22.9 and 87.9.

The second questionnaire chosen was the Expectation Ratings (ER) [4]. It consists in collecting the expectation of an application's ease of use before using it, and then the actual experienced ease of use of that same application. Similarly to the UMUX-LITE, it can be rated on a 7-point Likert scale which results in being able to use the second

Table 2
Questionnaires and open questions filled in by the participant

Questionnaire	Item
ER [4]	I expect the application to be easy to use knowing that I am able (<i>mode</i>) ^a .
UMUX-LITE [36]	This application capabilities meet my requirements. This application is easy to use.
Open questions	Which mode suited you the best/the least, and why? Can you think of elements that might improve the application's accessibility?

^a(*mode*) = “to see and hear”, “to see but not hear”, or “to hear but not see”.

item of UMUX-LITE as the second item of ER as well, i.e. the one linked to the experienced ease of use. The ER can give two distinct information about accessibility in applications: first, the expectation ratings can show what the participants expect when they are in a situation with a disability (i.e. no hearing or no vision), and when they are not. The ratings directly corresponds to the score given on the 7-point Likert scale, thus ranging from 1 to 7. Secondly, the gap between experiences and expectations informs about whether the proposed adapted design exceeded, matched, or subceeded the participant's expectations considering the mode they were in. It is calculated as follows:

$$Gap = Experience_{[1-7]} - Expectation_{[1-7]}, \quad (2)$$

where *Experience* and *Expectation* corresponds to the two items of the ER questionnaire, thus each ranging from 1 to 7. Therefore, the gap between experience and expectation ranges from -6 to 6 .

The experiment protocol was as follows: the participant was welcomed, invited to sit down in front of the computer, and a brief explanation about the experiment was given. The participant was told about the three modes and could ask about the general rules of the game. No information on how the game is played with the robot was given, as it should be self-explanatory in the application. To enhance the efficiency of face tracking and speech recognition, the participant had a white wall in the background and needed to wait for a "beep" sound before speaking. The participant was warned about this sound. How to rate items on a 7-point Likert scale was explained and the participant was told that they would have three items to rate for each mode. The participant was informed that they had to play from 2 to 4 rounds for each mode to have a complete experience. Before each mode, the participant rated the first ER item. The mode was launched by the researcher, then the two items from UMUX-LITE were rated by the participant. At the end of the experiment, both open questions were asked to the participant. They were then thanked for their participation with chocolates. The whole experiment took around 15 to 20 minutes per participant.

4.3. Results

The collected data met the requirements to run a repeated measures Multivariate ANalysis Of VAriance (MANOVA) that compared the three modes on the i) UMUX-LITE score obtained from Equation (1), ii) expectation rating, and iii) gap between experience and expectation. An overall difference between the three modes was unveiled, $F(6, 6) = 6.823$, $\eta_p^2 = .872$, $p = .017$. Subsequent Bonferroni corrected univariate tests gave more insights into the significance of each dependent variable taken separately as is shown in Fig. 4.

Figure 4a presents the estimated marginal means of the UMUX-LITE scores for the three modes. According to Bangor et al. [6], a score superior or equal to 70 is considered acceptable. The usability scores of mode S and mode H were close, with a value around 70. Even if mode SH exceeded the other two by 10, S and H still presented correct usability scores, meaning the application was usable with visual or hearing impairments. In the open questions,

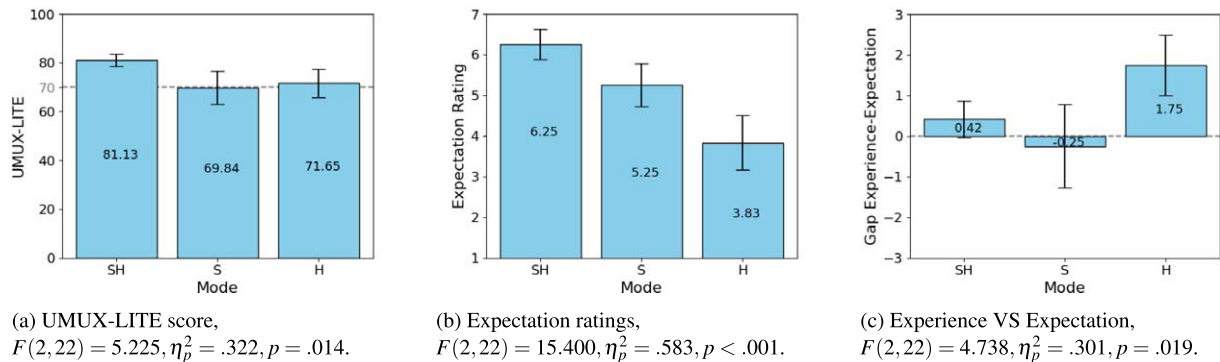


Fig. 4. Estimated marginal means of the three dependent variables – (a) UMUX-LITE score, (b) expectation ratings, and (c) gap between experience and expectation ratings – for the three modes SH, S, and H with error bars for 95% confidence intervals. The initial study shows that all of the dependent variables differed across the three modes.

the majority of participants declared preferring mode SH because it corresponds to what they are used to; thus, explaining the higher ratings for mode SH.

As demonstrated by Fig. 4b, the average expectation ratings for the three modes were substantially different. The expectations for mode H were the lowest with an average rating of 3.83. According to the 7-point Likert scale, it corresponds to the user somewhat disagreeing (3) or having a neutral opinion (4) on the expected application's ease of use before trying it. Mode S also showed lower expectations than mode SH. It can be deduced that while participants expected an application easy to use when able to see and hear, they expected it to be harder when having a disability, especially a visual impairment.

The second element of interest in ER was the gap between the experiences and expectations of the user for the different modes. If the gap is positive or negative, it means the experience respectively exceeded or subceeded the initial expectation. A gap close to 0 indicates that the experience matched the expectations. Figure 4c shows that experiences largely exceeded expectations for mode H, meaning that the application design was substantially easier to use than what the participants expected. The experiences almost matched the expectations for modes SH and S. In average the experience was slightly exceeding expectations in mode SH, and slightly disappointing for mode S.

The results for the three dependent variables reveal that there was a lot of variance among participants. This is notable, especially in the modes with a disability, as illustrated by the error bars for 95% confidence intervals in Fig. 4. These intervals suggest that while, on average, the application was perceived as accessible in the modes with a disability, for a few participants, the experience might have been easier to use or, on the contrary, more challenging.

The answers to the open questions suggest that some participants were overwhelmed by the application in mode S. The visual elements were numerous and possibly hard to keep track of. This would explain the slightly negative gap between experience and expectation in mode S. Most participants preferred mode SH, as it is what they are accustomed to and was more engaging. In contrast, two participants found mode SH to be overwhelming with the use of multimodality – visuals and sound – making it harder to follow than modes S and H. The majority of participants disliked mode H the most, as it really differs from what they are accustomed to, and the pauses in the robot's speech could be confusing, leaving the participant wondering if they needed to say something or not. In mode H, four participants felt the game was unfair because they had to state their choice before Haru, which gave the impression that Haru had an advantage over them.

On the elements to add to make the game more accessible, the choice of interaction with the application was recurrent. Some participants would have appreciated being able to click on physical buttons instead of stating their choice aloud in mode H, and some others would have preferred to always speak instead of clicking in modes S and SH. It was mentioned by several participants that the clicking interaction could be enhanced with a touch screen instead of a mouse or pad. Concerning the display, the only recurring element found concerned mode S, where it was suggested that the visuals should be placed more on the same level to prevent having too many places to look at. Another suggestion was to add more flashy colours or blinking elements to help direct the user's gaze.

4.4. Discussion

Despite the small number of 12 participants, differences between the three modes: i) see and hear, ii) only see, and iii) only hear were unveiled in favor of see and hear. The Rock–Paper–Scissors application's usability was considered good in all modes. Expectations differed among the modes, with see and hear having the highest expectations and only hear the lowest. With both see and hear and only see, the application met the participants' expectations. With only hear, it performed above expectations. The results from this user-study suggest that the social robot Haru can support people with a visual or hearing impairment.

The modes with a disability were expected to be harder to use than the modes without one, especially the mode where the user can only hear. This expectation might have arisen due to participants' awareness that they would need to state their choices out loud.

Two comments were very recurrent in the open questions concerning the modes with a disability. In the mode where the user could only hear, the user had to state their choice out loud, just before the robot. If the participant lost, it felt unfair because Haru could have chosen its play according to their choice. This might have influenced the usability score in this mode. As our goal was to provide the same experience in all modes, another method of interaction should be chosen in order to mitigate this feeling. In the mode where the user could only see, several

participants mentioned that there were visual information all over the screen and it could be difficult to know where to look at, and when. This probably caused the slightly negative gap between experience and expectations in this mode, and surely influenced the usability score. Directing the user's gaze more should allow us to reduce this problem.

The overall preference for the mode with see and hear might be due to the participants' habits. Nevertheless, a large variance was observed among participants in the modes with a disability. This suggests that in average participants have found the application accessible and easy to use with a disability while a few of them found it more challenging. We believe that enhancing the design of the application according to the suggestions in the open questions will help in reducing the variance among participants. These changes could lead to a more stable application in terms of user's satisfaction regarding accessibility. This is what motivated us to continue with a follow-up study.

5. Study 2: Interacting with one modality for all

In this section, the follow-up study conducted after obtaining the results of the initial study (see Section 4.3) is presented. It includes the improvements brought to the application's design, the performed user-study, the new results obtained, a comparison with the first study, and a discussion.

5.1. Improved design

The recurrent answers from the participants to the open questions in the initial study motivated us to improve the design of the application and run the study a second time. Our goals were i) to reduce the overall variance among participants, especially in the modes with a disability, ii) to reduce as much as possible the elements from the first design that might have disturbed the participants' experience (i.e., feeling of unfairness in mode H), and iii) to improve the usability of all modes, especially mode S which had the lowest score in the initial study.

The participants felt that mode H was unfair because of the mode of interaction (speech). Indeed, the participant had to say their choice out loud before the robot to not harm the speech recognition system. Therefore, if the robot won, the user felt it was unfair as they thought Haru chose what to play based on their choice. To cope with this issue, the modes of interaction were changed. We designed a tailored control pad to play the Rock–Paper–Scissors game, as displayed in Fig. 5. Thanks to this new controller, all modes use the control pad to interact with the robot instead of having either a buttons board on screen or a speech recognition system. As the participants in our study did not have any disabilities, they were allowed to see the keys, even in mode H. Indeed, with blind participants,



Fig. 5. Tailored control pad used in the second design to interact with the robot.

Table 3
Adaptation of the second design for user perception (top part) and interaction (bottom part)

User's abilities	See and hear	Only see	Only hear
Score board	✓	✓	✗
Speech balloon	✗	✓	✗
Face tracking	✓	✓	✗
Haru says the score	✗	✗	✓
Haru shows its choice	✓	✓	✗
Haru states its choice	✗	✗	✓
Physical buttons	✓	✓	✓

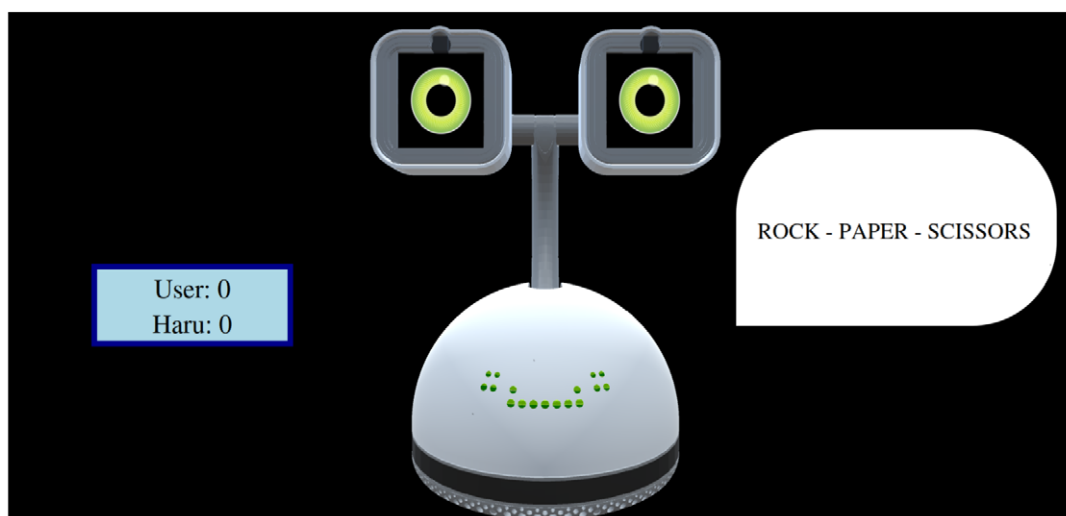


Fig. 6. Second design of the Rock–Paper–Scissors application with Haru's virtual environment for a user only able to see and not hear. For a user able to see and hear, the interface is similar, without the speech balloon for text. For a user only able to hear and not see, a black screen is displayed.

3D printed keys recognizable through touch, for instance with braille, would be used. However, sighted people do not have the habit of recognizing elements through touch alone, and would have to be trained to recognize braille. Even with time, they would probably not reach the level of reading fluency of braille readers. Therefore, using these buttons worked with our prototype. It is however important to keep in mind that when moving from prototype to reality, the keys should be replaced with 3D printed keys that can be recognized through touch alone. The summary of the adaptation to the different modes for the improved design is presented in Table 3.

To improve the usability of mode S, changes were made following the participants' suggestions to avoid being overwhelmed by elements all over the screen, see Fig. 6. First, as the mode of interaction was changed, the buttons board was removed. Second, the score board was placed closer to the robot and made bigger. The subtitles bar was replaced by a speech balloon, which reinforces the impression that Haru addresses the user and avoids having the user feeling like they need to concentrate too much on the bottom part of the screen. Additionally, in order to direct the user's gaze more, the score board blinks three times in different colors depending on who won the previous round: it blinks in green if the user won, in red if Haru won, and in grey if it was a tie.

5.2. Methods

The proceedings for the user-study stayed very similar as the ones for the initial study – cf Section 4.2 – as we wanted to be able to compare both studies. Once again, 12 participants from the Information and Computing Sciences from Utrecht University, Utrecht, the Netherlands were recruited. None of the participants from the initial

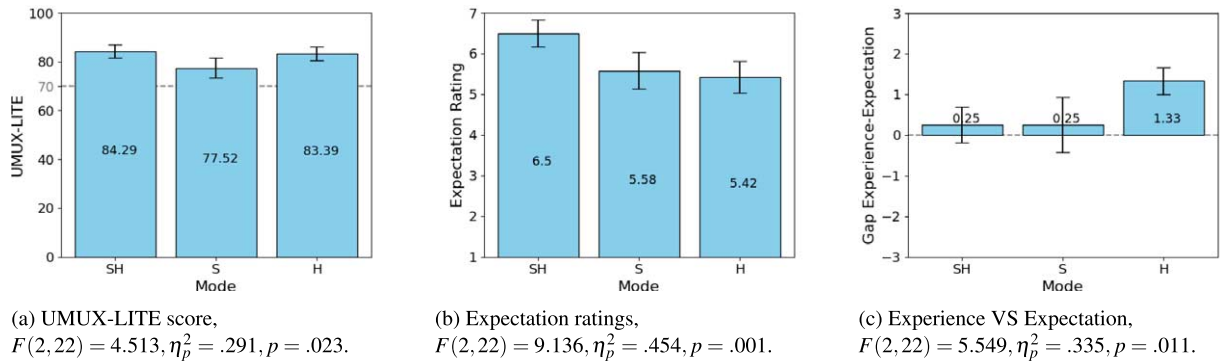


Fig. 7. Estimated marginal means of the three dependent variables – (a) UMUX-LITE score, (b) expectation ratings, and (c) gap between experience and expectation ratings – for the three modes SH, S, and H with error bars for 95% confidence intervals. The follow-up study shows that all of the dependent variables differed across the three modes.

study participated in the follow-up study. The questionnaires used to evaluate the usability and accessibility of the different modes stayed the same as for the initial study.

In mode H, even though the participants were not able to see anything on screen, they were still allowed to look at the buttons to play. Indeed, as mentioned previously, as we were trying to see if all modes appeared as usable and enjoyable for the users, a participant should be able to try the three modes and thus, none of the participants had a disability. Otherwise, for mode H, 3D printed keys recognizable through touch would have been used.

5.3. Results

As for the initial study, a repeated measures MANOVA was run to compare the three modes on the following dependent variables: i) UMUX-LITE score obtained from Equation (1), ii) expectation rating, and iii) gap between experience and expectation. The analysis once again unveiled an overall difference between the three modes, with $F(6, 6) = 12.801, \eta_p^2 = .928, p = .003$. In order to investigate each dependent variable separately – as for the initial study –, subsequent Bonferroni corrected univariate tests followed, see Fig. 7.

As mentioned previously, a UMUX-LITE score superior or equal to 70 is considered acceptable in terms of usability [6]. Figure 7a, which presents the estimated marginal means of the UMUX-LITE scores for the three modes, shows that all modes reached a score largely superior to 70. This proves all modes were considered usable by the participants. Mode SH reached the highest usability score, closely followed by mode H. This was confirmed in the open questions where the majority of participants declared preferring mode SH and added that mode H was equally enjoyable and easy to use. Mode S had the lowest score, but still ended up greater than 70.

Figure 7b presents the average expectation ratings for the three modes. The modes with a disability (mode S with 5.58 and H with 5.42) had a slightly lower expectation ratings than mode SH (with 6.5). According to the 7-point Likert scale, these scores correspond to the user somewhat agreeing (5), agreeing (6), or strongly agreeing (7) on the expected application's ease of use before trying it. It can be inferred that while participants did not expect any mode to be really hard to use, they were more confident about this statement for mode SH – where they are able to see and hear –, than for the modes with a disability.

Finally, the last dependent variable was the gap between experiences and expectations of the user for each of the three modes. As previously stated, a gap close to 0 suggests that the experience matched the expectations. On the contrary, a positive or negative gap indicates that the experience respectively exceeded or subceeded the user's expectations. The observable trend on Fig. 7c reveals that the experiences almost matched expectations for modes SH and S (slightly exceeding them on average), and they largely exceeded expectations in mode H.

The variance among participants appear to be smaller in the first study for the modes with a disability for the three dependent variables, as shown by the error bars for 95% confidence intervals in Fig. 7. The most obvious case is for the UMUX-Lite score. Indeed, Fig. 7a shows that the lower bound of the confidence interval for each mode is still greater than 70, meaning that we are 95% certain that the application would be considered accessible in all modes by any user.

Table 4

Quantitative comparison between the initial study and the follow-up study. The UMUX-LITE scores, expectation ratings, and gaps between experience and expectation in modes See and Hear (SH), only See (S), and only Hear (H) are reported

	Range	Mode	Study 1: initial study		Study 2: follow-up study	
			Mean	Standard deviation	Mean	Standard deviation
UMUX-LITE score	[22.9, 87.9]	SH	81.13	5.23	84.29	5.33
		S	69.84	13.53	77.52	8.15
		H	71.65	11.55	83.39	5.58
Expectation rating	[1, 7]	SH	6.25	0.75	6.50	0.67
		S	5.25	1.06	5.58	0.90
		H	3.83	1.34	5.42	0.79
Gap experience–expectation	[−6, 6]	SH	0.42	0.90	0.25	0.87
		S	−0.25	2.05	0.25	1.36
		H	1.75	1.49	1.33	0.65

Regarding the two open questions, no recurring elements were found, except users stating that their preferred mode was the one where they could see and hear as it is what they are accustomed to. Most of them also mentioned liking mode H. Two participants actually preferred modes with only one channel (either mode S or H) instead of the multimodal one, because they liked having to focus on only one channel for information. Two other participants mentioned that they would have appreciated having the speech balloon on screen in mode SH as well, rather than only in mode S.

During the experiment, we noticed that in mode S, it was sometimes hard for the participants to know when the robot was finished speaking, i.e. when they could press one button to play, even with the countdown from 3 to 1 from the robot. This might explain the lower usability score obtained by mode S compared to the other modes. One participant suggested in the open questions to highlight the text as Haru is speaking to follow its progress and solve this problem. We believe this would be a relatively easy and beneficial improvement, and intend to implement it in future work.

5.4. Evaluation: Study 1 and 2 compared

Both studies were run with the same number of 12 participants, equally distributed between genders. Table 4 displays the results obtained for both studies side-by-side for the three dependent variables: i) UMUX-LITE score, ii) expectation rating, and iii) gap between experience and expectation. The two studies exhibit several similarities. Notably, all UMUX-LITE scores were consistently at or above 70, indicating all three modes were deemed usable by the participants. Among the modes, SH attained the highest usability score, while S achieved the lowest, although still meeting the criteria for accessibility. In terms of expectation ratings, participants expressed higher expectations for mode SH, i.e. in the absence of disability, compared to modes S and H. The experience–expectation gap was nearly negligible for modes SH and S in both studies, whereas mode H had a positive gap, meaning that users' experience exceeded their initial expectations. Furthermore, the follow-up study's open questions mirrored those of the initial study by displaying personal preference: a few participants favored single-channel modes, or expressed a desire to keep speech balloons in mode SH, even though they were able to hear.

Table 4 also highlights some differences observed between the two studies. Firstly, as intended, there was overall less variance between participants in the second study, especially for the modes with a disability. For the UMUX-Lite score, the standard deviation dropped from 13.53 to 8.15 in mode S, and from 11.55 to 5.58 in mode H. For the gap between experience and expectation, the standard deviation dropped from 2.05 to 1.36 in mode S, and from 1.49 to 0.65 in mode H. The same trend is also visible for the expectation rating. This drop in variance shows the second design of the application lead to a more stable design in terms of user's satisfaction regarding accessibility. The follow-up study displayed higher usability scores for all modes, especially mode S improving by 7.68 and mode H by 11.74. Even though in the open questions the majority of participants still declared preferring mode SH, most of them added that mode H was equally enjoyable and easy to use, thus, the overall preference for mode SH was no longer obvious. Additionally, in Study 2, expectations for mode H were higher, similar to those for mode S.

This was in contrast to Study 1, where mode H was expected to be more challenging than mode S, possibly due to participants knowing they would have to state their choices out loud. Concerning the experience–expectation gap, the slightly negative gap obtained for mode S in the first study was eliminated for the second. This added to the increase in usability score suggest that the improved design was well received. Lastly, in the open questions of the initial study, most participants felt overwhelmed by mode S due to the multiple visual elements scattered across the screen. However, this concern was not mentioned by users in the second study, indicating that repositioning the elements on the screen and guiding the user’s gaze more helped in alleviating that feeling. The other common comment from participants in the initial study was that mode H felt unfair, as they needed to state their choice out loud before the robot, Haru. In Study 2, this feeling was addressed by changing the mode of interaction, allowing the robot to state its choice exactly when the user pressed the button of their choice, effectively removing the sense of unfairness.

5.5. Discussion

Once again, in a study with 12 participants, differences between the three modes: i) see and hear, ii) only see, and iii) only hear were unveiled. The results follow a similar trend as in the initial study, thus confirming our findings: the Rock–Paper–Scissors application’s was considered accessible in all modes. Expectations differed among the modes, with the modes with a disability reaching lower ones than the mode without one. With both see and hear and only see, the application met the participants’ expectations. With only hear, it performed above expectations. The results from this two user-studies clearly indicate that the social robot Haru can support people with a visual or hearing impairment.

Several alterations were brought to the application’s design between both studies. Notably, the interaction method (buttons board on screen or speech recognition depending on the modes) was replaced, with all modes transitioning to a control pad-based interaction. This change aimed at mitigating the feeling of unfairness when the user could only hear and had to state their choice out loud, before the robot. The fact that the participants could see the keys in the only hear mode is a limitation as not seeing the keys and having to recognize them through touch alone would surely influence their experience. Nevertheless, as mentioned before, individuals with sight are not accustomed to identifying elements solely through touch and would require training to recognize braille. Even with sufficient time, they may not attain the same level of reading fluency as braille users. Allowing them to visually identify keys rather than relying on braille reading, which blind participants could have engaged in, was necessary. The other modifications included relocating the score board and adding a blinking feature to signify score changes. The subtitles bar was replaced by a speech balloon for the mode in which users could only see. These changes surely influenced the ratings of the participants in the modes where they were able to see.

The new design with the revised interaction method helped all modes to reach higher usability scores, to reduce the variance among participants, and to reduce the gap between the modes with a disability and the mode without one. The mode where the user is only able to see had the lowest usability score, albeit a satisfactory one. This might be due to the fact that it seemed harder in this mode to know exactly when the robot was finished speaking. As a participant suggested, it could be mitigated by highlighting the text as Haru speaks to follow its progress.

In both studies, the users’ expectations differed among the modes, with see and hear having the highest expectations, and the modes with disabilities showing smaller ratings. In the initial study, the difference was more striking with the mode only hear having a very low expectation ratings compared to the other two modes, which was not the case in the follow-up study. Nevertheless, the participants were still expecting an application easier to use in the see and hear mode.

6. General discussion

We investigated the usability of the application in various modes where participants were able to assess the three modes, compare them, and express their preferences with their reasons and improvement suggestions. Crucially, we aimed to understand whether the modes designed for users with sensory disabilities could offer an equivalent experience in terms of completeness, enjoyability, and usability compared to the mode without any disability. To

achieve this objective, the first open question asked to the participants was “Which mode suited you the best/the least, and why?”. Hence, it was imperative that participants could thoroughly test all modes, thus making it necessary to recruit participants without any disability. However, we acknowledge that it constitutes a limitation and that conducting a future study with participants who have impairments would provide additional insights into the accessibility and user experience of the application. Nevertheless, both studies already highlighted several interesting aspects. The application was considered accessible in each mode, the expectations were lower when the participants were in a mode with a disability, and the experiences either matched or exceeded the expectations. The studies underline the importance of personal preference for the favorite mode of interaction, even though we can discern a general preference for the multimodal mode, with see suggested to be the most important aspect (cf. [59]) in the initial study.

In both user-studies, the number of participants was small with 12 participants for each study. However, the statistical tests we performed gave us significant results. Indeed, the convention is to take a significance level α of 0.05, which allows to generalize on the results with a 95% confidence if the obtained p -value is smaller than α . In our case, all the statistical tests resulted in a p -value largely smaller than 0.05. The highest p -value obtained across all the presented tests had a value of 0.023. In other words, we are more than 95% confident (precisely more than 97.7% confident, in relation to the highest p -value of 0.023) that the differences we found with our samples of 12 participants are representative of the population and can be generalized.

The social robot Haru can be considered as an embodied intelligent agent with a role of first point of contact, manager, or butler in an Intelligent Environment (IE). The virtual environment of the robot was used in this research, but all elements could be moved to a physical IE: the different windows on the computer can be replaced by actual screens and the robot’s camera and microphone would be used instead of the computer’s one. The simulation provides a different experience from the real robot regarding the interaction itself. However, it does not change the functionality we are investigating which is having an intelligent agent managing the modalities to use for inputs and outputs to adapt to user’s sensory abilities. Besides, using the simulated robot helps to avoid a “wow-effect” for participants who have never been in the presence of a robot before, and to avoid the user familiarity bias for participants who may have preconceived notions or biases towards robots. As user acceptance towards social robots is not the purpose of this research, removing or diminishing these biases can be beneficial. Future work in an actual IE with a design based on this prototype would give more insights into how the use of an IE can promote universal access with a social robot as intelligent agent. The control pad used in the second study helped removing the feeling of unfairness in the game and should therefore be kept. In the IE, they should be 3D printed and recognizable through touch to suit users with a visual impairment. Wearable sensors could be added to enhance Haru’s knowledge of the current physiological and affective state of the user [55].

Two sensory disabilities are included in this study, but universal access should consider many more and cover cases of multiple disabilities. A user could be able to see and hear, but not to click or press buttons, so speech recognition should be available in all modes. However speech recognition is not always adapted either, for instance if a user has a speech disorder [33,34]. Other ways to interact with the robot could be developed such as eye tracking, movement sensing technologies or haptic devices [23,33]. Several factors can be thought of for the manner of perceiving information, such as choosing visuals, sounds, or both. Other elements could be tuned such as paralinguistic aspects of speech [55] and the speed of speech of the robot to avoid an application going too fast for an individual, or too slow with too many long pauses [8].

Abilities are different for each individual. In addition to impairments, a neurodevelopmental condition, such as autism, can lead to an alteration in how sensory information is processed [10]. Physical and cognitive abilities are also not stable, they vary throughout life with ageing leading to a decrease in these abilities. IE should adapt to suit the user by meeting their changing requirements [50,52]. In Section 2, three methods to provide accessible technology devices were discussed: post-development accessibility features, assistive technologies, and ability-based design (cf Fig. 1). Post-development accessibility features impose a delay before the devices can be accessible and rarely provides a complete access to the system. Assistive technologies allow the user to employ the system, but the users themselves have the burden to adapt to the system. Ability-based design [60,61], which was the principle used in this paper, frees the user of this burden by having the system adapting to their abilities. Future work could focus on automated ability-based design via the automated assessment of user abilities. The user should also be able to communicate their preferences. The application should adapt accordingly by providing a suitable method

of interaction. In such an IE, Haru can act as manager and relay the information to the different devices to manage inputs and outputs selected modalities.

The need for universal access for next generation interfaces is paramount if we want to avoid social exclusion for different groups of users. Carrying on this research with more interaction methods and modalities can lead to covering more disabilities and help getting closer to universally accessible digital environments.

7. Conclusion

This study proposed a prototype with an adapted multimodal design for an application of Rock–Paper–Scissors with the social robot Haru. To be adaptive to different sensory disabilities, the game had three distinct modes where the user could either see and hear, only see, or only hear. Both of the performed user-studies showed significant results on several levels, each following the same trend which helped confirming our findings. It suggested that users expected the application to be harder to use when put into a condition with a disability. In the first study, it appeared to be especially the case with a visual impairment. In the end, the users' expectations were either matched or exceeded because the application was on average thought to be accessible in all modes.

At a time where technology has become omnipresent and does not cease to advance, it is essential to make sure that all users – including people with disabilities and older adults – are given the same access to these technologies and do not become socially excluded. The findings from this research indicate that making use of multimodality with a social robot in IE to promote accessibility is promising as it showed to perform well in terms of usability on our application. Besides, the open questions suggested that user preferences are crucial to take into account to ensure user satisfaction. An intelligent agent embodied as a social robot, such as Haru, can act as the manager of the IE by taking into account both the user's (dis)abilities and preferences, and selecting accordingly the inputs and outputs modalities for interaction to ensure a suitable communication with the user.

Acknowledgements

We thank our partners, members from the HONDA project between the Honda Research Institute in Japan and Utrecht University in the Netherlands, for making the virtual environment of Haru available to us for this research.

Conflict of interest

The authors have no conflict of interest to report.

References

- [1] J. Abascal, L. Azevedo and A. Cook, Is universal accessibility on track?, in: *Universal Access in Human–Computer Interaction. Methods, Techniques, and Best Practices*, M. Antona and C. Stephanidis, eds, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2016, pp. 135–143. doi:[10.1007/978-3-319-40250-5_13](https://doi.org/10.1007/978-3-319-40250-5_13).
- [2] N. Abdullah, K.E.Y. Low and Q. Feng, Sensory disability, in: *Encyclopedia of Gerontology and Population Aging*, D. Gu and M.E. Dupre, eds, Springer International Publishing, Cham, 2021, pp. 4468–4473. doi:[10.1007/978-3-030-22009-9_480](https://doi.org/10.1007/978-3-030-22009-9_480).
- [3] H.S. Ahn, I.-K. Sa, D.-W. Lee and D. Choi, A playmate robot system for playing the rock-paper-scissors game with humans, *Artificial Life and Robotics* **16**(2) (2011), 1614–7456. doi:[10.1007/s10015-011-0895-y](https://doi.org/10.1007/s10015-011-0895-y).
- [4] W. Albert and E. Dixon, Is this what you expected? The use of expectation measures in usability testing, in: *Proceedings of UPA 2003: 12th Annual UPA Conference, Usability Professionals Association*, Scottsdale, AZ, USA, 2003, p. 10.
- [5] J.C. Augusto, V. Callaghan, D. Cook, A. Kameas and I. Satoh, Intelligent environments: A manifesto, *Human-centric Computing and Information Sciences* **3**(1) (2013), 12, ISSN 2192-1962. doi:[10.1186/2192-1962-3-12](https://doi.org/10.1186/2192-1962-3-12).
- [6] A. Bangor, P.T. Kortum and J.T. Miller, An empirical evaluation of the system usability scale, *International Journal of Human–Computer Interaction* **24**(6) (2008), 574–594, ISSN 1044-7318. doi:[10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776).
- [7] N.M. Barbosa, J. Hayes, S. Kaushik and Y. Wang, “Every website is a puzzle!”: Facilitating access to common website features for people with visual impairments, *ACM Transactions on Accessible Computing* **15**(3) (2022), 19–11935, ISSN 1936-7228. doi:[10.1145/3519032](https://doi.org/10.1145/3519032).

- [8] J.G. Beerends, N.M.P. Neumann, E.L. van den Broek, A. Llagostera, J. Vranic, C. Schmidmer and J. Berger, Subjective and objective assessment of full bandwidth speech quality, *IEEE/ACM Transactions on Audio, Speech and Language Processing* **28**(1) (2020), 440–449. doi:[10.1109/TASLP.2019.2957871](https://doi.org/10.1109/TASLP.2019.2957871).
- [9] F.H.F. Botelho, Accessibility to digital technology: Virtual barriers, real opportunities, *Assistive Technology* **33**(sup1) (2021), 27–34, ISSN 1040-0435. doi:[10.1080/10400435.2021.1945705](https://doi.org/10.1080/10400435.2021.1945705).
- [10] L. Boyd, Sensory processing in autism, in: *The Sensory Accommodation Framework for Technology: Bridging Sensory Processing to Social Cognition*, L. Boyd, ed., Synthesis Lectures on Technology and Health, Springer Nature, Switzerland, Cham, 2024, pp. 27–40. doi:[10.1007/978-3-031-48843-6_3](https://doi.org/10.1007/978-3-031-48843-6_3).
- [11] H. Brock, S. Šabanović and R. Gomez, Remote you, Haru and me: Exploring social interaction in telepresence gaming with a robotic agent, in: *Companion of the 2021 ACM/IEEE International Conference on Human–Robot Interaction, HRI '21 Companion*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 283–287. doi:[10.1145/3434074.3447177](https://doi.org/10.1145/3434074.3447177).
- [12] A. Brunes, M.B. Hansen and T. Heir, Loneliness among adults with visual impairment: Prevalence, associated factors, and relationship to life satisfaction, *Health and Quality of Life Outcomes* **17**(1) (2019), 24. doi:[10.1186/s12955-019-1096-y](https://doi.org/10.1186/s12955-019-1096-y).
- [13] L. Burzagli, P.L. Emiliani, M. Antona and C. Stephanidis, Intelligent Environments for all: A path towards technology-enhanced human well-being, *Universal Access in the Information Society* **21**(2) (2022), 437–456, ISSN 1615-5289. doi:[10.1007/s10209-021-00797-0](https://doi.org/10.1007/s10209-021-00797-0).
- [14] N. Carbonell, Ambient multimodality: Towards advancing computer accessibility and assisted living, *Universal Access in the Information Society* **5**(1) (2006), 96–104, ISSN 1615-5297. doi:[10.1007/s10209-006-0027-y](https://doi.org/10.1007/s10209-006-0027-y).
- [15] T. Chen, J.R. Gil-Garcia and M. Gasco-Hernandez, Understanding social sustainability for smart cities: The importance of inclusion, equity, and citizen participation as both inputs and long-term outcomes, *Journal of Smart Cities and Society* **1**(2) (2022), 135–148. doi:[10.3233/SCS-210123](https://doi.org/10.3233/SCS-210123).
- [16] C. Creed, M. Al-Kalbani, A. Theil, S. Sarcar and I. Williams, Inclusive AR/VR: Accessibility barriers for immersive technologies, *Universal Access in the Information Society* (2023), ISSN 1615-5297. doi:[10.1007/s10209-023-00969-0](https://doi.org/10.1007/s10209-023-00969-0).
- [17] J.E. da Rosa Tavares and J.L. Victória Barbosa, Apollo SignSound: An intelligent system applied to ubiquitous healthcare of deaf people, *Journal of Reliable Intelligent Environments* **7**(2) (2021), 157–170, ISSN 2199-4676. doi:[10.1007/s40860-020-00119-w](https://doi.org/10.1007/s40860-020-00119-w).
- [18] K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. Koay and I. Werry, What is a robot companion – friend, assistant or butler?, in: *International Conference on Intelligent Robots and Systems (IROS 2005)*, 2005, pp. 1197. doi:[10.1109/IROS.2005.1545189](https://doi.org/10.1109/IROS.2005.1545189).
- [19] M.M.A. de Graaf and S. Ben Allouch, Exploring influencing variables for the acceptance of social robots, *Robotics and Autonomous Systems* **61**(12) (2013), 1476–1486, ISSN 0921-8890. doi:[10.1016/j.robot.2013.07.007](https://doi.org/10.1016/j.robot.2013.07.007).
- [20] P. Droege, *Intelligent Environments*, Elsevier, 1997. doi:[10.1016/B978-0-444-82332-8.X5000-8](https://doi.org/10.1016/B978-0-444-82332-8.X5000-8).
- [21] R. Dunne, T. Morris and S. Harper, A survey of ambient intelligence, *ACM Computing Surveys* **54**(4) (2022), 1–27, ISSN 0360-0300, 1557-7341. doi:[10.1145/3447242](https://doi.org/10.1145/3447242).
- [22] P.L. Emiliani and C. Stephanidis, Universal access to ambient intelligence environments: Opportunities and challenges for people with disabilities, *IBM Systems Journal* **44**(3) (2005), 605–619, ISSN 0018-8670. doi:[10.1147/sj.443.0605](https://doi.org/10.1147/sj.443.0605).
- [23] A. Flores Ramones and M.S. del-Rio-Guerra, Recent developments in haptic devices designed for hearing-impaired people: A literature review, *Sensors* **23**(6) (2023), ISSN 1424-8220. doi:[10.3390/s23062968](https://doi.org/10.3390/s23062968).
- [24] A. Foley and B.A. Ferri, Technology for people, not disabilities: Ensuring access and inclusion, *Journal of Research in Special Educational Needs* **12**(4) (2012), 192–200, ISSN 1471-3802. doi:[10.1111/j.1471-3802.2011.01230.x](https://doi.org/10.1111/j.1471-3802.2011.01230.x).
- [25] R. Gomez, D. Szapiro, K. Galindo and K. Nakamura, Haru: Hardware design of an experimental tabletop robot assistant, in: *Proceedings of the 2018 ACM/IEEE International Conference on Human–Robot Interaction, HRI '18*, Association for Computing Machinery, New York, NY, USA, 2018, pp. 233–240. doi:[10.1145/3171221.3171288](https://doi.org/10.1145/3171221.3171288).
- [26] H. Guner and C. Acarturk, The use and acceptance of ICT by senior citizens: A comparison of technology acceptance model (TAM) for elderly and young adults, *Universal Access in the Information Society* **19**(2) (2020), 311–330, ISSN 1615-5297. doi:[10.1007/s10209-018-0642-4](https://doi.org/10.1007/s10209-018-0642-4).
- [27] Y. Hasuda, S. Ishibashi, H. Kozuka, H. Okano and J. Ishikawa, A robot designed to play the game “Rock, Paper, Scissors”, in: *2007 IEEE International Symposium on Industrial Electronics*, 2007, pp. 2065–2070. doi:[10.1109/ISIE.2007.4374926](https://doi.org/10.1109/ISIE.2007.4374926).
- [28] M. Hilbert, Digital technology and social change: The digital transformation of society from a historical perspective, *Dialogues in Clinical Neuroscience* **22**(2) (2020), 189–194. doi:[10.31887/DCNS.2020.22.2/mhilbert](https://doi.org/10.31887/DCNS.2020.22.2/mhilbert).
- [29] S. Johansson, J. Gulliksen and C. Gustavsson, Disability digital divide: The use of the Internet, smartphones, computers and tablets among people with disabilities in Sweden, *Universal Access in the Information Society* **20**(1) (2021), 105–120, ISSN 1615-5297. doi:[10.1007/s10209-020-00714-x](https://doi.org/10.1007/s10209-020-00714-x).
- [30] M. Kandil, F. AlAttar, R. Al-Baghdadi and I. Damaj, AmIE: An ambient intelligent environment for blind and visually impaired people, in: *Technological Trends in Improved Mobility of the Visually Impaired*, S. Paiva, ed., Springer International Publishing, Cham, 2020, pp. 207–236. doi:[10.1007/978-3-030-16450-8_9](https://doi.org/10.1007/978-3-030-16450-8_9).
- [31] S. Kartakis and C. Stephanidis, A design-and-play approach to accessible user interface development in ambient intelligence environments, *Computers in Industry* **61**(4) (2010), 318–328, ISSN 0166-3615. doi:[10.1016/j.compind.2009.12.002](https://doi.org/10.1016/j.compind.2009.12.002).
- [32] H.M. Khan, K. Abbas and H.N. Khan, Investigating the impact of COVID-19 on individuals with visual impairment, *British Journal of Visual Impairment* (2023). doi:[10.1177/02646196231158919](https://doi.org/10.1177/02646196231158919).
- [33] S. Koch Fager, M. Fried-Oken, T. Jakobs and D.R. Beukelman, New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science, in: *Augmentative and Alternative Communication (Baltimore, Md.: 1985)*, Vol. 35, 2019, pp. 13–25, ISSN 1477-3848. doi:[10.1080/07434618.2018.1556730](https://doi.org/10.1080/07434618.2018.1556730).

- [34] K. Kuhn, V. Kersken, B. Reuter, N. Egger and G. Zimmermann, Measuring the accuracy of automatic speech recognition solutions, *ACM Transactions on Accessible Computing* **16**(4) (2024), 25–12523, ISSN 1936-7228. doi:[10.1145/3636513](https://doi.org/10.1145/3636513).
- [35] J.-H. Lee and H. Hashimoto, Intelligent space – concept and contents, *Advanced Robotics* **16**(3) (2002), 265–280, ISSN 0169-1864. doi:[10.1163/156855302760121936](https://doi.org/10.1163/156855302760121936).
- [36] J. Lewis, B. Utesch and D. Maher, UMUX-LITE: When there’s no time for the SUS, in: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2013, pp. 2099–2102. doi:[10.1145/2470654.2481287](https://doi.org/10.1145/2470654.2481287).
- [37] I. Mansutti, I. Achil, C. Rosa Gastaldo, C. Tomé Pires and A. Palese, Individuals with hearing impairment/deafness during the COVID-19 pandemic: A rapid review on communication challenges and strategies, *Journal of Clinical Nursing* **32**(15–16) (2023), 4454–4472, ISSN 1365-2702. doi:[10.1111/jocn.16572](https://doi.org/10.1111/jocn.16572).
- [38] N. Marangunić and A. Granić, Universal access in the information society international journal technology acceptance model: A literature review from 1986 to 2013, *Universal Access in the Information Society* **14** (2014), 1–15. doi:[10.1007/s10209-014-0348-1](https://doi.org/10.1007/s10209-014-0348-1).
- [39] O. Metatla, S. Bardot, C. Cullen, M. Serrano and C. Jouffrais, Robots for inclusive play: Co-designing an educational game with visually impaired and sighted children, in: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 1–13. ISBN 978-1-4503-6708-0. doi:[10.1145/3313831.3376270](https://doi.org/10.1145/3313831.3376270).
- [40] S. Ntoa, G. Margetis, M. Antona and C. Stephanidis, Digital accessibility in intelligent environments, in: *Human–Automation Interaction*, V.G. Duffy, M. Lehto, Y. Yih and R.W. Proctor, eds, Vol. 10, Springer International Publishing, Cham, 2023, pp. 453–475. doi:[10.1007/978-3-031-10780-1_25](https://doi.org/10.1007/978-3-031-10780-1_25).
- [41] Z. Obrenović, J. Abascal and D. Starčević, Universal accessibility as a multimodal design issue, *Communications of the ACM* **50**(5) (2007), 83–88. doi:[10.1145/1230819.1241668](https://doi.org/10.1145/1230819.1241668).
- [42] O. Ozmen Garibay, B. Winslow, S. Andolina, M. Antona, A. Bodenschatz, C. Coursaris, G. Falco, S.M. Fiore, I. Garibay, K. Grieman, J.C. Havens, M. Jirotko, H. Kacorri, W. Karwowski, J. Kider, J. Konstan, S. Koon, M. Lopez-Gonzalez, I. Maifeld-Carucci, S. McGregor, G. Salvendy, B. Shneiderman, C. Stephanidis, C. Strobel, C. Ten Holter and W. Xu, Six human-centered artificial intelligence grand challenges, *International Journal of Human–Computer Interaction* **39**(3) (2023), 391–437, ISSN 1044–7318. doi:[10.1080/10447318.2022.2153320](https://doi.org/10.1080/10447318.2022.2153320).
- [43] M. Pérez-Escolar and F. Canet, Research on vulnerable people and digital inclusion: Toward a consolidated taxonomical framework, *Universal Access in the Information Society* **22**(3) (2023), 1059–1072, ISSN 1615-5289, 1615-5297. doi:[10.1007/s10209-022-00867-x](https://doi.org/10.1007/s10209-022-00867-x).
- [44] L. Pettersson, S. Johansson, I. Demmelmaier and C. Gustavsson, Disability digital divide: Survey of accessibility of eHealth services as perceived by people with and without impairment, *BMC Public Health* **23**(1) (2023), 1471–2458. doi:[10.1186/s12889-023-15094-z](https://doi.org/10.1186/s12889-023-15094-z).
- [45] J. Pivin-Bachler, R. Gomez and E.L. van den Broek, “One for All, All for One”. A First Step Towards Universal Access with a Social Robot, 2023. doi:[10.3233/AISE230031](https://doi.org/10.3233/AISE230031).
- [46] C. Putnam, E.J. Rose and C.M. MacDonald, “It could be better. It could be much worse”: Understanding accessibility in user experience practice with implications for industry and education, *ACM Transactions on Accessible Computing* **16**(1) (2023), 9–1925, ISSN 1936-7228. doi:[10.1145/3575662](https://doi.org/10.1145/3575662).
- [47] R. Ragel, R. Rey, A. Páez, J. Ponce, K. Nakamura, F. Caballero, L. Merino and R. Gómez, Multi-modal data fusion for people perception in the social robot Haru, in: *Social Robotics*, F. Cavallo, J.-J. Cabibihan, L. Fiorini, A. Sorrentino, H. He, X. Liu, Y. Matsumoto and S.S. Ge, eds, Lecture Notes in Computer Science, Springer Nature, Switzerland, Cham, 2022, pp. 174–187. doi:[10.1007/978-3-031-24667-8_16](https://doi.org/10.1007/978-3-031-24667-8_16).
- [48] A. Shukla, M. Harper, E. Pedersen, A. Goman, J.J. Suen, C. Price, J. Applebaum, M. Hoyer, F.R. Lin and N.S. Reed, Hearing loss, loneliness, and social isolation: A systematic review, otolaryngology–head and neck surgery, *Official journal of American Academy of Otolaryngology-Head and Neck Surgery* **162**(5) (2020), 622–633. doi:[10.1177/0194599820910377](https://doi.org/10.1177/0194599820910377).
- [49] C. Stephanidis, Designing for all in ambient intelligence environments: The interplay of user, context, and technology, *Int. J. Hum. Comput. Interaction* **25** (2009), 441–454. doi:[10.1080/10447310902865032](https://doi.org/10.1080/10447310902865032).
- [50] C. Stephanidis, G. Salvendy, M. Antona, J.Y.C. Chen, J. Dong, V.G. Duffy, X. Fang, C. Fidopiastis, G. Fragomeni, L.P. Fu, Y. Guo, D. Harris, A. Ioannou, K.-A.K. Jeong, S. Konomi, H. Krömker, M. Kurosu, J.R. Lewis, A. Marcus, G. Meiselwitz, A. Moallem, H. Mori, F. Fui-Hoon Nah, S. Ntoa, P.-L.P. Rau, D. Schmorow, K. Siau, N. Streitz, W. Wang, S. Yamamoto, P. Zaphiris and J. Zhou, Seven HCI grand challenges, *International Journal of Human–Computer Interaction* **35**(14) (2019), 1229–1269, ISSN 1044-7318, 1532-7590. doi:[10.1080/10447318.2019.1619259](https://doi.org/10.1080/10447318.2019.1619259).
- [51] K. Tuinenbreijer, Selfish agents and the butler paradigm, in: *Social Intelligence Design 2004*, A. Nijholt and T. Nishida, eds, 2004, pp. 197–199, ISSN 0929–0672. ISBN 90-75296-12-6.
- [52] E.L. van den Broek, Robot nannies: Future or fiction?, *Interaction Studies* **11**(2) (2010), 274–282. doi:[10.1075/is.11.2.16van](https://doi.org/10.1075/is.11.2.16van).
- [53] E.L. van den Broek, ICT: Health’s best friend and worst enemy?, in: *BioSTEC 2017: 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, E.L. van den Broek, A. Fred, H. Gamboa and M. Vaz, eds, Proceedings Volume 5: HealthInf, SciTePress – Science and Technology Publications, Lda., Porto, Portugal, 2017, pp. 611–616. doi:[10.5220/0006345506110616](https://doi.org/10.5220/0006345506110616).
- [54] E.L. van den Broek, Monitoring technology: The 21st century’s pursuit of well-being? EU-OSHA discussion paper, European Agency for Safety and Health at Work (EU-OSHA), Bilbao, Spain, 2017. Available online, <https://osha.europa.eu/en/tools-and-publications/publications/monitoring-technology-workplace/view> (last accessed on 25/09/2023).
- [55] E.L. van den Broek, M.H. Schut, J.H.D.M. Westerink and K. Tuinenbreijer, Unobtrusive Sensing of Emotions (USE), *Journal of Ambient Intelligence and Smart Environments* **1**(3) (2009), 287–299. doi:[10.3233/AIS-2009-0034](https://doi.org/10.3233/AIS-2009-0034).
- [56] E.L. van den Broek, F. van der Sluis and T. Dijkstra, Cross-validation of bi-modal health-related stress assessment, *Personal and Ubiquitous Computing* **17**(2) (2013), 215–227. doi:[10.1007/s00779-011-0468-z](https://doi.org/10.1007/s00779-011-0468-z).

- [57] E.L. van den Broek and J.H.D.M. Westerink, Considerations for emotion-aware consumer products, *Applied Ergonomics* **40**(6) (2009), 1055–1064. doi:[10.1016/j.apergo.2009.04.012](https://doi.org/10.1016/j.apergo.2009.04.012).
- [58] F. van der Sluis and E.L. van den Broek, Feedback beyond accuracy: Using eye-tracking to detect comprehensibility and interest during reading, *Journal of the Association for Information Science and Technology* **74**(1) (2023), 3–16. doi:[10.1002/asi.24657](https://doi.org/10.1002/asi.24657).
- [59] F. van der Sluis, E.L. van den Broek, A. van Drunen and J.G. Beerends, Enhancing the quality of service of mobile video technology by increasing multimodal synergy, *Behaviour & Information Technology* **37**(9) (2018), 874–883. doi:[10.1080/0144929X.2018.1505954](https://doi.org/10.1080/0144929X.2018.1505954).
- [60] J.O. Wobbrock, K.Z. Gajos, S.K. Kane and G.C. Vanderheiden, Ability-based design, *Communications of the ACM* **61**(6) (2018), 62–71, ISSN 0001-0782. doi:[10.1145/3148051](https://doi.org/10.1145/3148051).
- [61] J.O. Wobbrock, S.K. Kane, K.Z. Gajos, S. Harada and J. Froehlich, Ability-based design: Concept, principles and examples, *ACM Transactions on Accessible Computing* **3**(3) (2011), 9–1927, ISSN 1936-7228. doi:[10.1145/1952383.1952384](https://doi.org/10.1145/1952383.1952384).
- [62] World Health Organization, *World Report on Vision*, World Health Organization, Geneva, 2019, ISBN 978-92-4-151657-0. Available online, <https://apps.who.int/iris/handle/10665/328717> (last accessed on 25/09/2023).
- [63] World Health Organization, *World Report on Hearing*, World Health Organization, Geneva, 2021, ISBN 978-92-4-002049-8. Available online, <https://www.who.int/publications/i/item/9789240020481> (last accessed on 25/09/2023).
- [64] M. Zajicek and W. Morrissey, Multimodality and interactional differences in older adults, *Universal Access in the Information Society* **2**(2) (2003), 125–133, ISSN 1615-5297. doi:[10.1007/s10209-003-0045-y](https://doi.org/10.1007/s10209-003-0045-y).