Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20

Can blindfolded users replace blind ones in product testing? an empirical study

Shi Qiu, Jun Hu, Ting Han & Matthias Rauterberg

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2023.2226768

Can blindfolded users replace blind ones in product testing? an empirical study

Shi Qiu [a,b], Jun Hu[b], Ting Han[a] and Matthias Rauterberg[b]

[a]Department of Design, Shanghai Jiao Tong University, Shanghai, People's Republic of China; [b]Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands

ABSTRACT
During the design, it is important to evaluate the user experience of representative users in many human product interactions. But, in some cases, it is difficult or even impossible to recruit representative users because they have disabilities that do not allow them to take part in such investigations. Thus, alternative populations are widely studied. The most common way to replace real blind people is to use sighted but blindfolded users when studying design solutions. To test whether such alternative or proxy users can be used to represent blind people in social interactions, we examined the communication quality of 20 blind-sighted pairs and 20 blindfolded-sighted pairs in two different experiments. A prototype named E-Gaze glasses was evaluated as the testing tool. Results clearly show that the blindfolded participants achieved significantly higher communication quality than the blind participants. In qualitative data analysis, the blindfolded participants also reported their user experience of being blindfolded in conversations. Our qualitative results strengthen the conclusion that blindfolded users' behaviour is different from real blind users' behaviour. We recommend that blind users should not be substituted for blindfolded users in human product evaluations when communication quality is measured.

ARTICLE HISTORY
Received 11 October 2019
Accepted 12 June 2023

KEYWORDS
Accessibility; alternative users; communication quality; proxy user; representative user; visual impairments

1. Introduction

Explorations and evaluations of ideas with representative users are critical in the human-product interaction design community. Recruiting representative users in usability tests and controlled experiments can be regarded as a general rule for interaction design researchers and practitioners (Sears and Hanson 2011). However, a fundamental challenge is that for most disability studies, representative users will be an issue, as researchers and participants with disabilities will face many difficulties. Some difficulties are addressed as below: (1) Difficult to recruit people with disability, although contacting some institutions (e.g. special education schools) might provide a solution; (2) Difficult to bring people with disability (e.g. physical disabilities) to a specific location for laboratory experiments; (3) Difficult to find a sufficient number of people with disability for running rigorous experiments (e.g. considering experimental control and statistical power); (4) Difficult to get verbal feedback from such users due to reduced introspection capacity (e.g. dementia patients). Due to all these reasons, getting enough data from such people with disability is too costly to obtain and they may be too exhausted to continue in a lab-based setting. Sometimes due to the privacy issue, it is even not possible to collect data from people with disability. In sum, it is not easy to involve people with disability in user experiments. Therefore, proxy and alternative users provide a feasible solution for running such user experiments. If target users are difficult to study, their caregivers or experts could be treated as proxy users. Researchers can investigate proxy users who are most familiar with target users. For example, Internet proxy users refer to older adults' grandchildren to help them do online shopping on behave of older adults (Dolničar et al. 2018). Here, we are more interested in alternative users. They refer to able-bodied people who simulate a given disability in a certain situation (e.g. using noise-dampening headphones to remove audio input or special glasses that distort a person's vision) (Sears and Hanson 2011). Among alternative users, the most common way is to invite sighted but blindfolded (hereafter blindfolded) participants to participate in an experiment as an alternative for the target

CONTACT Shi Qiu qiushi11@sjtu.edu.cn Department of Design, Shanghai Jiao Tong University, 800 Dongchuan RD. Minhang District, Shanghai, People's Republic of China

blind users (Mihara et al. 2005; Huang et al. 2012; Tang and Li 2014; Balata, Mikovec, and Neoproud 2015; Anagnostakis et al. 2016).

One of the typical examples is regarding investigating the shared interfaces in the collaboration between blindfolded and sighted people (Huang et al. 2012). Twenty-eight university students participated in their interface evaluation and collaborated in 14 pairs. All participants were sighted of which 14 participants in each pair were blindfolded. The results demonstrated that adding audio cues in a shared interface made collaboration between a sighted and a blindfolded person more efficient. The authors argued that the effects found in this evaluation can be generalised to visually impaired people. But, is it correct to generalize?

Silverman (2015) claimed that the disability simulation of alternative users gives misleading information. Sears and Hanson (2011) also emphasised that studying non-representative users might cause inaccurate conclusions or missed insights. Zeng and Weber (2015) integrated a collaborative approach to a navigation system, aiming at helping blind people find entrances independently in an unfamiliar environment. They tested the system with 10 participants (five blind and five blindfolded). The results indicated that because of having enough orientation and mobility skills, the blind participants significantly outperformed the blindfolded participants during evaluations. Therefore, in interaction design, the use of information provided by alternative users must be approached with caution. In certain cases, differences in user behaviours can be found between representative users and their alternative ones.

Earlier literature discussed blindfolded participants tested in many different kinds of contexts of use (e.g. shared interfaces (Huang et al. 2012), providing navigation (Kammoun, Bouhani, and Jemni 2016). Some cases support the results of alternative population testing, while others argue that alternative users provided misleading information. Since this topic is still controversial, it is worth being investigated. It is valuable to explore the representativeness of the alternative-users approach in more detail, especially in a certain context of use. Specifically, in our study, we aim at exploring whether blindfolded users can replace blind ones in communication-quality evaluations. We are also interested in investigating their user experience. The term 'user experience' has a wide range of meanings, from traditional usability to the aesthetic, hedonic, emotional, or experiential aspects of technology use (Forlizzi and Battarbee 2004). Hassenzahl and Tractinsky (2006) identify user experience takes a 'human' perspective, focusing more on the emotional impact on people than technology. They stated that user experience is the result of the user's internal state, the system being designed, and the environment in which the interaction occurs (e.g. the social settings). The communication experience is part of the user experience that takes place in a social setting. Thus, in our case, we want to investigate the user experience of blind and blindfolded people when they communicate face to face, especially their perceptual and emotional aspects.

The central focus of this study is the contrast in communication quality between blindfolded users and blind ones. However, blind people not only include people with total blindness but also include some intermediate levels (i.e. visual impairments of different degrees). From a broader perspective, people with visual impairments also include some older adults who gradually lose their sight as they age. In our study, to increase the heterogeneity in samples of people with impairments, we recruit blind participants who cover a wide range of blindness, from total blindness to low vision. Our findings could be beneficial for design researchers and designers who study visually impaired people and even older adults.

To be noted, although other contexts of use have been explored, to the best of our knowledge, communication quality in a face-to-face conversation scenario between blindfolded users and blind ones is still unknown. Conversation scenario is very common and important in real life and many product designs (e.g. designing the assistive device to enhance communication quality) are aimed at this scenario. However, this has been ignored in previous studies. It may be partly due to the preconceived impression that conversation mostly relies on hearing and speaking, so blind people are thought to have the similar or same abilities as sighted or blindfolded people. In our study, we identify this research gap and clarify the main research question: Is there a significant behavioural difference in the communication quality between blind and blindfolded people? To answer this research question, we developed a gaze simulation prototype called E-Gaze glasses as a testing tool in communication-quality evaluation (Qiu et al. 2020; Qiu et al. 2021, 2022; Qiu et al. 2020a, 2020b). Due to the reason that gaze is known to be an important social cue in face-to-face communication, simulating gaze for blind people can efficiently promote communication between blind and sighted people. More specifically, E-Gaze can simulate the visual gaze of a blind person, thereby establishing 'eye contact' between a blind person and a sighted person. Meanwhile, corresponding tactile feedback enables the blind person to sense when 'eye contact' occurs. In the user experiment, E-Gaze can provide certain stimuli according to the experimental conditions. The validity and reliability of the alternative-users approach are carefully examined via the E-Gaze in this research study.

2. Related work

2.1. Measuring communication quality

In this paper, we aim to investigate whether there is a significant difference in communication quality between blind and blindfolded people. In social science, communication quality was used in the research fields of social communication and personal relationships (Montgomery 1988). Particularly, the quality of face-to-face communication is highly related to social presence, which measures the perception of the other with whom one is interacting (Biocca, Harms, and Burgoon 2003). Biocca and Harms (2002) defined social presence as a 'sense of being with another', not only considering the face-to-face interactions of human to human but also considering the mediated experience of human to the artificial agent. Biocca and Harms (2002) categorised social presence into three distinct levels: (1) the perceptual level (i.e. one becomes aware of the co-presence of the other), (2) the subjective level (i.e. attentional engagement, emotional state, comprehension, and behavioural interaction), as well as (3) the intersubjective level (i.e. access the level of correlation between one's feelings of social presence and their impressions towards the other's psychological sense of social presence). In our experiments, we measured the level of communication quality through the specific levels of social presence based on the research from Biocca and Harms (2002).

It is appropriate and reasonable to use conversational context/scenario to measure communication quality. For example, Kemp and Rutter (1986) used controlled experiments to explore how blind participants behave in conversations. Ten blind-blind, ten blind-sighted, and ten sighted-sighted pairs were observed in a series of laboratory discussions. Finally, questionnaires were used to measure the behaviours of their own and their conversation partners. Another example is, Przybylski and Weinstein (2013) used a dyadic conversation scenario to test whether the presence of mobile communication technology could influence face-to-face conversation quality. In the controlled experiments, dyads were asked to spend 10 min discussing the topic together, trying to emulate the content of many real-life conversations. After conversations, dyads completed the questionnaires used to measure their conversation quality. These related works on the method are important to provide a good foundation for our research. In our experimental design, we borrow similar context, tasks, and procedures as studies described in (Kemp and Rutter 1986; Przybylski and Weinstein 2013). In our study, all the participants were paired. One pair included a sighted participant and a non-sighted participant. They discussed a daily topic for 10 min. After discussions, the participants filled in the questionnaires to measure their communication quality.

2.2. Blindfolded participants used in evaluation studies

Blindfolded participants used in evaluation studies (see Appendix) can be summarised in three categories: (1) only including blindfolded participants, (2) including both blind and blindfolded participants, but without comparison, and (3) a comparison of blind and blindfolded participants. We focus on the studies in category 3 because we want to investigate whether there is a significant difference in the communication quality between blind and blindfolded people and under which contexts of use.

2.2.1. Category 1: only including blindfolded participants

Research studies in this category substituted blind participants with blindfolded participants when studying technical solutions intended for blind people (Mihara et al. 2005; McGookin, Brewster, and Weiwei 2008; Tinwala and MacKenzie 2008; Gomez et al. 2011; Poláček, Grill, and Tscheligi 2012; Tang and Li 2014; Balata, Mikovec, and Neoproud 2015; Wilson et al. 2015; Alkhanifer and Ludi, 2015; Kammoun, Bouhani, and Jemni 2016; Qiu et al. 2016a). The contexts of use of these studies are quite diverse, and they are summarised below:

- Providing navigation (Gomez et al. 2011; Poláček, Grill, and Tscheligi 2012; Alkhanifer and Ludi, 2016; Kammoun, Bouhani, and Jemni 2016);
- Enhancing spatial cognition (Tang and Li 2014; Wilson et al. 2015);
- Accessibility of Braille and touch screen (Mihara et al. 2005; McGookin, Brewster, and Weiwei 2008);
- Blind camera system (Balata, Mikovec, and Neoproud 2015);
- Text entry techniques (Tinwala and MacKenzie 2008);
- Enhancing gaze signals perception to promote social interactions (Qiu et al. 2016a).

These studies reported the working progress and the preliminary findings, focusing on system improvements during the design and implementation. Blindfolded participants participated in such preliminary evaluations to examine the feasibility of the products. For instance, Mihara et al. (2005) presented a Braille-recognition system to help blind people who cannot read Braille, identify the desired button from an elevator or a ticket vending machine. In the preliminary evaluation, six blindfolded participants were able to recognise the meaning of the buttons that they identified. Observations regarding their behaviours provided insights for the system improvements. Tang and Li (2014) presented an assistive system named EyeWear, aiming at helping blind people to perceive the objects in their surroundings through spatial audio. They reported the work-in-progress and tested the system with twelve blindfolded participants. Preliminary findings revealed that EyeWear was able to provide a useful level of object localisation assistance, but the accuracy of grasping should be improved.

2.2.2. Category 2: including both blind and blindfolded Participants, but without a comparison

Research studies in this category include both blind and blindfolded participants in evaluations but without a direct comparison of their performances (Apostolopoulos et al. 2014; Smith and Nayar 2018). The contexts of use of these two studies are the indoor navigation system (Apostolopoulos et al. 2014) and the audio-based racing game (Smith and Nayar 2018). In these studies, researchers recruited the blind and blindfolded participants respectively at different stages of design. For instance, Apostolopoulos et al. (2014) developed a system to guide blind people through indoor environments. Three studies were conducted. In the first study, one blind and nine blindfolded participants validated the early stage of the system with a simple and interactive sensing approach achievable on a smartphone. In the second study, eight blindfolded participants evaluated different methods for estimating the user's step length. Finally, researchers recruited eight blind participants and collected their comments on the final system. Smith and Nayar (2018) introduced an audio-based user interface to allow blind people to play racing games. The evaluation included three blind and twelve blindfolded participants. Data from blind and blindfolded participants were analyzed and reported separately. Because the blindfolded participants might be biased by the initial challenge of being blind, they judged the capability of blind people as much less than it is (Silverman, Gwinn, and Van Boven 2015).

2.2.3. Category 3: a comparison of blind and blindfolded participants

These studies reported a comparison of blind and blindfolded participants in evaluations (Kamel and Landay 2002; Douglas and Willson 2007; Awada et al. 2013; Zeng and Weber 2015; Anagnostakis et al. 2016; Leo et al. 2018). The contexts of use of these studies are regarding graphic access (Kamel and Landay 2002; Awada et al. 2013), accessibility of museum collections (Anagnostakis et al. 2016) and visual symbols (Leo et al. 2018), navigation (Zeng and Weber 2015), as well as haptic perception (Douglas and Willson 2007).

Some of these studies reported a non-significant difference was observed between the performance of the blind and blindfolded participants (Kamel and Landay 2002; Douglas and Willson 2007; Anagnostakis et al. 2016). It has two possible interpretations: (1) Blind and blindfolded participants indeed performed in a similar manner; (2) Lack of statistical power. There is not a sufficient number of participants to achieve a large N, so a type II error may occur. It means the alternative hypothesis (H1) is true, but we remain H0. For instance, Kamel and Landay (2002) presented a method to transform a mouse-based graphic interface into an auditory interface. Researchers evaluated the interface with eight blind and eight blindfolded participants. The blind participants were observed to perform as well or better than the sighted participants on three tasks. But there was not a significant difference in statistics. This is possibly due to the small sample size ($n = 16$).

Other studies claimed that the blind and blindfolded participants behaved differently in evaluations (Awada et al. 2013; Zeng and Weber 2015; Leo et al. 2018). For instance, Leo et al. (2018) presented a set of tactile symbols that can be used on small-size tactile displays. They tested such symbols with 19, 20 blindfolded low vision, and 22 blindfolded full vision participants. The results demonstrated that the blind participants were much faster to identify tactile symbols than the blindfolded low vision and the blindfolded full vision participants.

Based on the above-mentioned literature, the results of comparing behaviours of blind and blindfolded people are still inconclusive. Thus, we need additional research studies to investigate how a behavioural difference in the communication quality between blind and blindfolded people.

3. Experimental study

We reported already the experimental design of the E-Gaze in Experiment 1 (E1) and Experiment 2 (E2) in detail elsewhere (Qiu et al. 2020a; Qiu et al. 2020b). Here, we report the results extracted from E1 and E2 together, aiming at answering our methodological research question: Is there a significant difference in communication quality between blind and blindfolded people when they talk with sighted people? We combined the dataset from E1 and E2 to enhance the overall test power (Judd, McClelland, and Ryan 2017), which is a technique in statistical analysis. It is better to combine than to analyze each dataset separately. In the current analysis, the sample size is 80. This is a sufficient sample size for doing quantitative data analysis. E1 and E2 share similarities. Both of them have four test conditions, from low to high, simple to complex, sharing the same pattern and the same contexts of use (i.e. E-Gaze was used in a conversational scenario). Most importantly, E1 and E2 can do repeated measurements with the same dependent variables. The same dependent variables are necessary, otherwise, two datasets cannot be combined. All the above properties of both datasets make the combination possible.

In our data analysis, we have one independent variable, three confounding factors, and one co-variate, as described in more detail below. The whole idea of this kind of data analysis is to exclude as much as possible the variance from controlling or other confounding factors. Take 'gender' as an example. We do not specify the hypothesis about gender effects, because we do not ask a research question about the male and female differences in both experiments, although several studies have shown gender differences in perception and cognition (Lawton and Hatcher 2005; Nario-Redmond 2010). We include 'gender' as one of the confounding factors. According to Rosel et al. (2005), age has an impact on the verbalism of blind and sighted people. So, we also include 'age' as a co-variate.

3.1. Experimental design

In experimental design, we have one independent variable (IV) 'sight-capacity', three confounding factors: (1) 'experiment', (2) 'gender', and (3) 'test-condition', as well as one co-variate 'age'.

Table 1. Combining test conditions from the E1 and E2 datasets.

Combination of Test Conditions	Conditions in E1	Conditions in E2
condition 1	No Gaze	Tactile Feedback and Interactive Gaze are not active
condition 2	Constant Gaze	Tactile Feedback is not active but Interactive Gaze is active
condition 3	Random Gaze	Tactile Feedback is active but Interactive Gaze is not active
condition 4	Interactive Gaze	Tactile Feedback and Interactive Gaze are both active

Our *independent variable* is 'sight-capacity'. This is the prime targeted factor in our data analysis. It is a between-subjects factor, including three conditions: (1) blind people, (2) blindfolded people, and (3) sighted people. We are interested in the difference between blind and blindfolded people.

Our *first confounding factor* is 'experiment'. By combining both datasets, we added the factor 'experiment'. It is a between-subjects factor, including two conditions: (1) E1, and (2) E2.

Our *second confounding factor* is 'gender'. It is a between-subjects factor, including two conditions: (1) male, and (2) female.

Our *third confounding factor* is 'test-condition'. It is a within-subjects factor, including four conditions: (1) condition 1, (2) condition 2, (3) condition 3, and (4) condition 4. This factor comes from combining test conditions of the E1 dataset and E2 dataset. In E1, it has four test conditions: (1) No Gaze, (2) Constant Gaze, (3) Random Gaze, and (4) Interactive Gaze. In E2, it also has four test conditions: (1) Tactile Feedback and Interactive Gaze are not active; (2) Tactile Feedback is not active but Interactive Gaze is active; (3) Tactile Feedback is active but Interactive Gaze is not active; (4) Tactile Feedback and Interactive Gaze are both active. Because both E1 and E2 had four test conditions and the same dependent variables, we can combine them and include the factor 'test-condition' in our data model. Table 1 shows the combination of test conditions from the E1 and E2 datasets.

Our *co-variate* is 'age' measured in years of age (rational scale).

To measure the *communication quality*, we used an adapted version of the 'Networked Minds Social Presence Inventory' (NMSPI) (Harms and Biocca 2004). NMSPI has 36 items with a seven-point Likert scale ranging from one (1 = strongly disagree) to seven (7 = strongly agree). We have six dependent variables: (1) *co-presence* (i.e. the degree of awareness of the partner), (2) *attentional allocation* (i.e. the amount of attention that a person provides to, and receives from the partner), (3) *perceived message understanding* (i.e. the ability to understand the message from the partner), (4) *perceived affective understanding* (i.e. the ability to understand the partner's emotion and attitudes), (5) *perceived emotional interdependence* (i.e. the extent that a person's emotional state affects, and is affected by the partner), and (6) *perceived behavioral interdependence* (i.e. the extent that a person's behaviour affects and is affected by the partner).

At the end of the experiment, we asked in a semi-structured interview about the user experience of blindfolded participants only (i.e. 'What is your perception when you are blindfolded in conversations?'). We are interested in investigating differences in the communication experience of blindfolded participants compared to their everyday conversations without the blindfold. Sight can significantly influence their behaviours. Even if they are blindfolded, they can still maintain some habits (e.g. forming visual images in their minds through verbal descriptions) that differ from those of truly blind users. Their reflections may provide insights into interpreting quantitative results. For blind and sighted participants, just talk as usual and they do not need to change their behaviours in conversations. So we do not ask about their communication experience.

3.2. Prototype

Our prototype for gaze simulation that is used in the present study is called E-Gaze glasses (Qiu et al. 2018). E-Gaze was designed for simulating gaze behaviours for visually impaired people, to enhance the communication quality between blind and sighted people during social interactions. The function of gaze simulation consists of two major aspects: the first is to simulate gaze behaviours for blind people as a visual reaction, and the second is to assist blind people to perceive the tactile feedback when establishing 'eye contact' with sighted people. Our design philosophy embodies the ideas mentioned by Schmutz, Sonderegger, and Sauer (2017), moving from an 'accessibility for users with disabilities' approach to an 'inclusive design' approach. The joint consideration of users with and without disabilities is of great importance. Visual and tactile gaze feedback of the E-Gaze can benefit both sighted and blind people in conversational scenarios. The design of the E-Gaze and the rationale behind it have been extensively reported in (Qiu et al. 2020; Qiu et al. 2016a; Qiu et al. 2018; Qiu et al. 2015b; Qiu, Rauterberg, and Hu 2016b; Qiu et al. 2020a, 2020b); the design is briefly summarised as follows.

Our gaze simulation technology was inspired by AgencyGlass (Osawa 2014). Originally AgencyGlass was designed for sighted people to reduce a potential emotional load in a societal context. In our research, we developed the AgencyGlass further into our E-Gaze system to provide means for blind people to react to sighted people by simulating eye contact. The interactive gaze that is displayed via E-Gaze is based on an eye-contact and a turn-taking strategy in human-to-human conversations. The implemented interactive gaze model has been described in detail elsewhere (Qiu et al. 2020b).

Our test system includes both the E-Gaze glasses and a tactile wristband (Figure 1). The E-Gaze helps a blind person react to a sighted person through a simulated gaze via artificial eyes, and a tactile wristband provides the blind person with the corresponding feedback when establishing 'eye contact'. The tactile feedback allows the blind person to realise that a sighted person is looking at the artificial eyes. The blind person wears the tactile wristband on the dominant hand since the tactile perception of people's dominant hands is more sensitive than non-dominant hands (Ghent 1961).

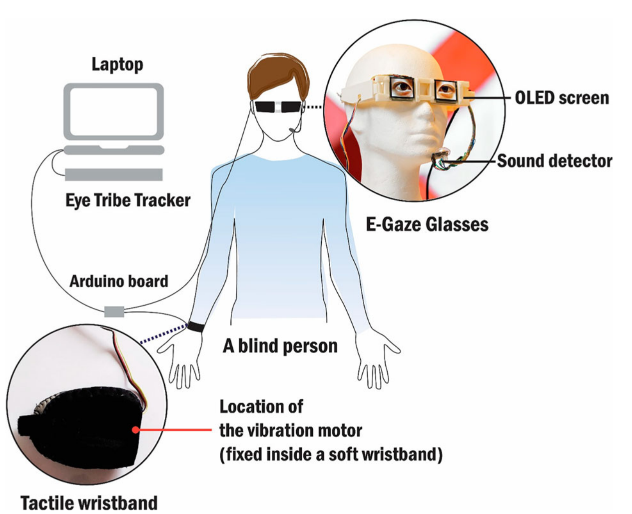

Figure 1. Overview of the test system, including both the artificial eyes displayed via E-Gaze glasses and a tactile wristband for feedback about eye contact moments.

3.3. Participants

Overall in E1 and E2, the participants were 80 students in China. They were divided into two conversation groups: (1) blind-sighted conversation, and (2) blindfolded-sighted conversation. Each of these two groups had the same number of participants. Blind-sighted conversation included 20 blind-sighted pairs (M_{age} = 16.80, SD = 1.04, N = 16 females and 24 males) while blindfolded-sighted conversation consisted of 20 blindfolded-sighted pairs (M_{age} = 22.55, SD = 2.66, N = 24 females and 16 males). The profiles of participants in these two groups are shown in Tables 2–4.

Our experiments were conducted at two locations in China: (1) Shanghai for blindfolded-sighted pairs, and (2) Yangzhou for blind-sighted pairs. Sighted participants were recruited from Shanghai Jiao Tong

Table 2. Participants' age.

Conversation Group	N	Mean	Std. Deviation	Minimum	Maximum
Blind-sighted	40	16.80	1.04	15	20
Blindfolded-sighted	40	22.55	2.66	18	30
Total	80	19.67	3.52	15	30

Table 3. Participants' gender.

Conversation Group	Gender	N
Blind-sighted	Male	24
	Female	16
Blindfolded-sighted	Male	16
	Female	24
Total	Male	40
	Female	40
	Total	80

University and Jiangsu College of Tourism, while the blind participants were from Yangzhou Special Education School. We selected the blind participants based on two criteria: (1) blindness is the only significant disability, and (2) the participants should be registered blind in the China Association of the Blind (CAB 2013). During the recruitment of blind participants in the special education school, we did find a small number of blind students with cognitive disabilities. So a school teacher helped us preliminarily screen blind students who could attend classes normally and communicate with teachers and other students normally. Among these blind students, blindness is their only significant disability.

A total of 20 blind participants covered a wide range of blindness, from total blindness to low vision. Among them, 11 participants became blind by birth. Vision conditions of all blind participants are reported in Table 4. A total of 20 blind participants covered a wide range of blindness, from total blindness to low vision. Among them, 11 participants became blind by birth. Vision conditions of all blind participants are reported in Table 5. According to the International Classification of Diseases 11 (World Health Organization 2022), visual impairment fall into four categories: (1) mild (i.e. visual acuity worse than 6/12 to 6/18), (2) moderate (i.e. visual acuity worse than 6/18 to 6/60), (3) severe (i.e. visual acuity worse than 6/60 to 3/60), and (4) blindness (i.e. visual acuity worse than 3/60).

Due to the difficulty of recruiting blind participants, a total of 80 participants, including 20 blind, can be considered a large sample size. In our study, the large sample size and extensive visual impairment of blind participants would provide a good basis for the generalizability of the results. Kukull and Ganguli (2012) emphasised that a fair sample must provide a valid estimate of the population characteristics being studied. Recruiting blind participants with extensive visual impairment is particularly useful for avoiding selection bias. For example, if the visual impairment in the sample includes only total blindness and excludes moderate or severe visual impairment, the experimental results may be biased. The larger samples and heterogeneity in samples of people with impairments can efficiently increase statistical power.

In practice, during the recruitment of blind participants, individuals with various causes of blindness and varying degrees of visual impairment were observed. This is common in studies that involve blind participants and cannot be fully avoided. While we could not completely rule out the possibility that varying degrees of visual impairment influenced the results, we were able to ensure that all blind participants with varying degrees of impairment during the experiment were temporarily at the same level as total blindness, equivalent to sighted participants wearing blindfolds. This made the experimental design more rigorous and removed

Table 4. Participants' education.

Conversation Group	Education	N
Blind-sighted	The third grade	1
	The eighth grade	11
	The tenth grade	23
	The eleventh grade	5
	Total	40
Blindfolded-sighted	Bachelor	18
	Master	21
	Ph.D. student	1
	Total	40

Table 5. Vision conditions of the blind participants.

Gender	Age	Vision Conditions (WHO Standard)[a]	Congenital Blindness	Able to Sense Light and Colour
Male	17	Blindness	Yes	No
Male	16	Blindness	Yes	No
Male	16	Blindness	Lost sight at 12	No
Male	15	Blindness	Lost sight at 2	No
Female	16	Blindness	Yes	No
Female	16	Blindness	Yes	No
Male	18	Blindness	Lost sight at 7 or 8	Yes
Female	16	Blindness	Lost sight at 10	Yes
Female	15	Blindness	Lost sight at 9	Yes
Male	19	Severe visual impairment	Yes	Yes
Male	18	Severe visual impairment	Lost sight at 4	Yes
Male	17	Severe visual impairment	Yes	Yes
Male	16	Severe visual impairment	Lost sight at 12	Yes
Male	20	Moderate visual impairment	Yes	Yes
Female	19	Moderate visual impairment	Yes	Yes
Male	18	Moderate visual impairment	Yes	Yes
Female	17	Moderate visual impairment	Yes	Yes
Female	16	Moderate visual impairment	Yes	Yes
Male	16	Moderate visual impairment	Not mentioned	Yes
Female	16	Moderate visual impairment	Not mentioned	Yes

[a]Vision impairments are sorted from low to high.

any potential influence of varying degrees of visual impairment on the results.

A more detailed explanation of the prototype system and experimental setup are provided. The E-Gaze glasses are made of ABS plastic and were 3D printed. Since the E-Gaze glasses is only a prototype, the shell design of the glasses has not been designed to fit the face perfectly. The inner side of the glasses that is close to the skin still needs to add soft materials to make the wearer more comfortable. So, in the experiment, both blind and blindfolded participants wore eye masks and E-Gaze glasses. The black eye mask is opaque, and participants achieved the same level of blindness in the experiment whether they were blind with moderate vision impairment, completely blind or blindfolded.

Prior to conducting the experimental study, we interviewed 20 blind participants to investigate their perception of social signals (e.g. gaze, smile) (Qiu et al. 2020). We recruited participants with a broad range of visual impairments, ranging from mild to total blindness. This included individuals with both congenital and acquired blindness. We found that in face-to-face communication, all blind participants were unable to perceive the eye behaviour of their conversation partners. Their understanding of gaze behaviour was largely based on their own communicative difficulties, such as handling conversation turns. Little was known about how the act of gaze expresses different emotions and feelings, and some relied on exaggerated depictions in romance novels to imagine its function. Overall, their understanding of eye behaviour was indirect and vague. In the experiments, we also asked blind participants about their perception and understanding of eye contact. All 20 blind participants reported no direct experience and no clear concept of eye contact, including those who lost their vision after birth. After a few years, those who later went blind gradually forgot the act of making eye contact. Based on the above investigations on their understanding of eye contact, we believe that the effect of visual impairment of the blind participants on experimental results, whether mild or acquired, is not as significant as initially expected.

Additionally, the participants in each pair were matched of the same gender and similar age. Same-sex participants can effectively avoid confounding factors that improve communication quality due to the emotional or romantic attraction of the opposite sex. Since this experiment took place in China, this aspect is more sensitive than in western world (Ge 2017). In a conversational test scenario, participants of similar age can generate topics more easily. Each participant was awarded 100 CNY at the end of the experiment.

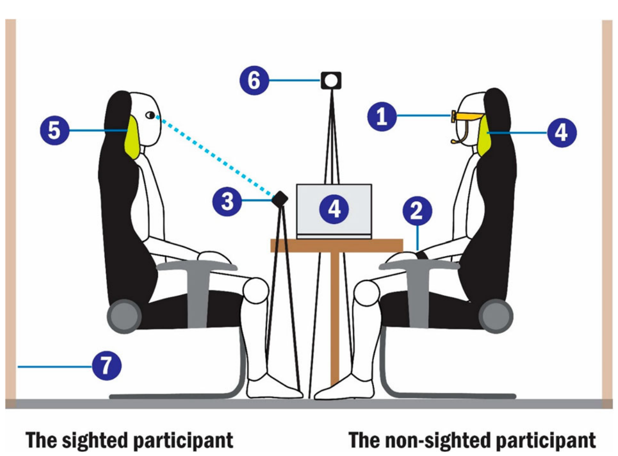

Figure 2. Side view of the experimental setup: (1) the E-Gaze glasses, (2) the tactile band, (3) the Eye Tribe Tracker, (4) the laptop, and (5) the pillow to support the neck of the participant to stabilise and track the gaze accurately, (6) the observation camera to capture the whole scene, and (7) the folding screens.

3.4. Setup

Our participants consisted of blind, blindfolded, and sighted ones. From now on, we call blind and blindfolded participants *non-sighted participants*. In E1, a non-sighted participant wore the E-Gaze and sat in front of a sighted participant. In E2, a non-sighted participant sat in the same position with both E-Gaze and the tactile wristband. The sighted participant was about 1.8 m away from the non-sighted participant, regarded as a comfortable social distance for people sitting in chairs or gathered in a room (Hall 1963). The Eye Tribe tracker[1] was aligned and adjusted towards the sighted participant's face for the maximum trackability of his/her gaze behaviour (see also Figure 2). With this eye tracker, we can measure the location of the sighted one's gaze. When this gaze matches the gaze of the virtual eyes of the non-sighted participants, these non-sighted participants of E2 can perceive this moment of 'eye contact' through the tactile feedback.

3.5. Procedure

E1 and E2 were conducted according to the Declaration of Helsinki. Informed consent was signed by all the participants. They were informed about the experiment and gave their consent to participate. In the blind-sighted pairs, a volunteer orally presented the consent form and allowed the blind participant enough time for the questions to be asked and answered. With a clear understanding, the blind participant gave oral consent for participation. The whole consent procedure was audio recorded as part of the documentation of the consent forms. After filling out the pre-experimental questionnaire regarding demographic information, non-sighted participants wore the E-Gaze in E1 and wore both E-Gaze and the tactile wristband in E2. For each blindfolded-sighted pair, one participant was randomly selected to be blindfolded. He or she should wear the blindfold during the overall experiment, including answering questionnaires. Next, all the participants discussed a daily topic after three minutes of preparation.

The topic was randomly selected from the 14 topics in an IELTS oral exam ('IELTS Speaking Module,' 2012). Blind students experience differences in educational and skill levels compared to their sighted peers. For example, in the special education school, music education is often emphasised, and blind students may learn to play an instrument. In regular schools, academic performance is usually a priority. This can lead to differences in the topics they are good at discussing. Therefore, we specially selected topics in an IELTS oral exam related to daily life, such as '*your favorite family member*' or '*the most important decision you have made in your life*'. These topics are easy to arouse the interest of talking with each other.

After this ten-minute discussion, they finished the post-experimental questionnaires. The experimenter orally presented the questionnaires to non-sighted participants and finished questionnaires according to their oral answers. Answering questionnaires took about 20-25 min for the blind participants and 15-20 min for the blindfolded participants. The blind participants spent more time than the blindfolded participants understanding and responding to all the questions. To avoid carry-over effects all four conversations were taken place in a counterbalanced order. The overall experiment of the blind-sighted pairs lasted about 150-180 min, and of the blindfolded-sighted pairs lasted about 120-150 min.

Table 6. Means and standard deviations on 'sight-capacity' in the experiment.

Measure	Sight capacity in the experiment (blind, blindfolded, sighted)	Mean	Std. Error
Co-presence	Blind	4.97	0.17
	Blindfolded	5.83	0.16
	Sighted	5.24	0.11
Attentional allocation	Blind	4.93	0.19
	Blindfolded	5.58	0.18
	Sighted	4.63	0.12
Perceived message understanding	Blind	4.67	0.21
	Blindfolded	6.31	0.21
	Sighted	4.74	0.13
Perceived affective understanding	Blind	4.00	0.25
	Blindfolded	5.58	0.24
	Sighted	4.29	0.16
Perceived emotional interdependence	Blind	4.83	0.27
	Blindfolded	5.26	0.26
	Sighted	4.31	0.17
Perceived behavioural interdependence	Blind	4.50	0.23
	Blindfolded	5.77	0.22
	Sighted	4.53	0.14

Note. The mean score ranges from 1-7.

4. Results

4.1. Quantitative analysis

We used General Linear Models (GLM) to analyze data collected from the subjective questionnaire Because 'sight-capacity' is our prime independent variable, we only report the main effect of 'sight-capacity' and all interaction effects with 'sight-capacity' for each dependent variable separately. Table 6 illustrates the means and standard deviations for the factor 'sight-capacity' in our dataset. Figure 3 shows bar charts with standard errors of the participants' communication quality measured with our six dependent variables. We present now the results of each dependent variable separately.

Co-presence. A significant main effect was observed for *sight capacity* [$F(2, 67) = 6.319$, $p \leq 0.003$, $\eta_p^2 = 0.159$] (Table 7). The post hoc contrast analysis revealed that the blindfolded participants ($M = 5.83$, $SE = 0.16$) perceived significantly higher co-presence than the blind participants ($M = 4.97$, $SE = 0.17$) (Table 6). Because Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(5) = 13.625$, $p \leq 0.018$, we report the results with the Greenhouse-Geisser correction. A significant two-way interaction effect was observed between *sight capacity* and *test condition*, [$F(5.308, 177.810) = 7.432$, $p < .001$, $\eta_p^2 = 0.182$] (Table 8). The three-way interaction between *sight capacity, test condition* and *experiment* was also significant [$F(5.308, 177.810) = 2.456$, $p \leq 0.032$, $\eta_p^2 = 0.068$] (Table 8).

Attentional allocation. A significant main effect was observed for *sight capacity* [$F(2, 67) = 10.290$, $p < .001$, $\eta_p^2 = 0.235$] (Table 7). The post hoc contrast analysis revealed that the blindfolded participants ($M = 5.58$, $SE = 0.18$) perceived significantly higher attentional allocation than the sighted participants ($M = 4.63$, $SE = 0.12$) (Table 6). The two-way interaction between *sight capacity* and *test condition* was not significant [$F(6, 201) = 1.912$, $p \leq 0.081$] (Table 8). The three-way interaction between *sight capacity*, *test condition,* and *experiment* was not significant [$F(6, 201) = 1.972$, $p \leq 0.071$] (Table 8).

Perceived message understanding (PMU). A significant main effect was observed for *sight capacity* [$F(2, 67) = 22.213$, $p < .001$, $\eta_p^2 = 0.399$] (Table 7). The post hoc contrast analysis revealed that the blindfolded participants ($M = 6.31$, $SE = 0.21$) perceived

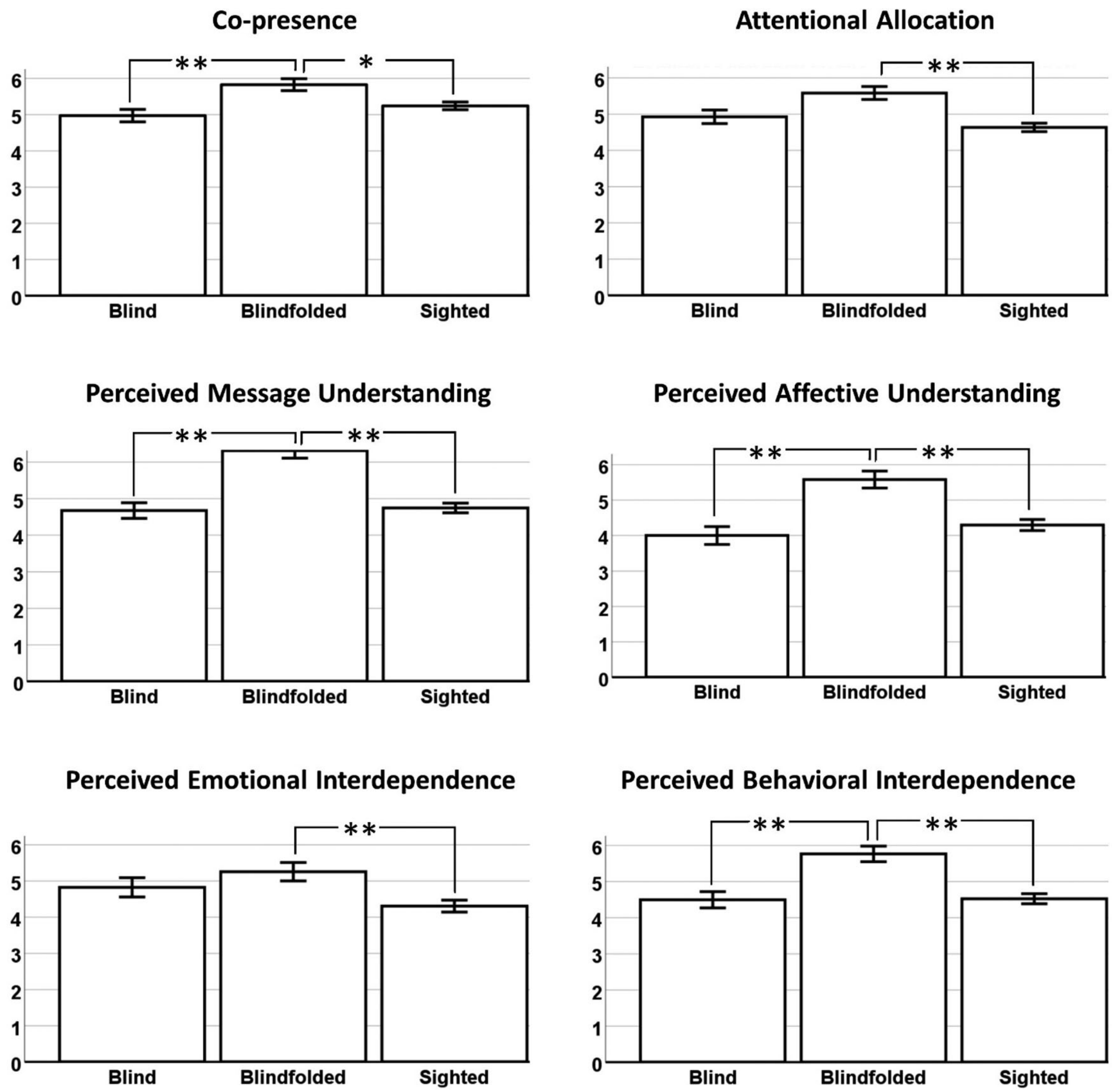

Figure 3. Bar charts with standard error (SE) bars of the participants' communication quality measured with six dependent variables (1) co-presence, (2) attentional allocation, (3) perceived message understanding, (4) perceived affective understanding, (5) perceived emotional interdependence, and (6) perceived behavioural interdependence, across three different sight capacities: blind people, blindfolded people and sighted people. Significant group difference are indicated as * = $p < .05$, ** = $p < .01$.

significantly higher PMU than the blind participants (M = 4.67, SE = 0.21) (Table 6). The two-way interaction between *sight capacity* and *test condition* was also significant [F(6, 201) = 2.749, $p \leq 0.014$] (Table 8). The three-way interaction between *sight capacity*, *test condition*, and *experiment* was not significant [F(6, 201) = 1.760, $p \leq 0.109$] (Table 8).

Perceived affective understanding (PAU). A significant main effect was observed for *sight capacity* [F(2, 67) = 11.861, $p < .001$, $\eta_p{}^2 = 0.261$] (Table 7). The post

Table 7. GLM results summary of main effects of 'sight-capacity' on the communication quality among sighted, blindfolded, and blind participants with varying degrees of visual impairment (N = 80).

Source	Measure	*SS*	*df*	*MS*	*F*	*p*	Observed Power[a]
'sight-capacity'	Co-presence	5.693	2	2.847	6.319	**0.003****	0.885
	Attentional allocation	10.999	2	5.499	10.290	**0.001****	0.984
	Perceived message understanding	31.138	2	15.569	22.213	**0.001****	1.000
	Perceived affective understanding	23.019	2	11.509	11.861	**0.001****	0.993
	Perceived emotional interdependence	12.159	2	6.080	5.585	**0.006****	0.841
	Perceived behavioural interdependence	19.260	2	9.630	12.345	**0.001****	0.995

Significant group difference; * $p < .05$, ** $p < .01$.
[a]Computed using alpha = .05

Table 8. GLM results summary of 2-way interaction effects between 'test-condition' and 'sight-capacity' on the communication quality; 3-way interaction effects between 'test-condition', 'sight-capacity' and 'experiment' on the communication quality.

Source	Measure	*SS*	*df*	*MS*	*F*	*p*
'test-condition' * 'sight-capacity'	Co-presence	16.068	5.308	3.027	7.432	**0.001****
	Attentional allocation	5.639	6	0.940	1.912	0.081
	Perceived message understanding	6.457	6	1.076	2.749	**0.014***
	Perceived affective understanding	5.770	5.180	1.114	2.167	0.058
	Perceived emotional interdependence	4.005	6	0.668	2.156	**0.049***
	Perceived behavioural interdependence	8.786	5.421	1.621	4.279	**0.001****
'test-condition' * 'sight-capacity' * 'experiment'	Co-presence	5.311	5.308	1.001	2.456	**0.032****
	Attentional allocation	5.817	6	0.970	1.972	0.071
	Perceived message understanding	4.134	6	0.689	1.760	0.109
	Perceived affective understanding	2.514	5.180	0.485	0.944	0.456
	Perceived emotional interdependence	3.021	6	0.503	1.626	0.142
	Perceived behavioural interdependence	4.135	5.421	0.763	2.014	0.073

Significant group difference; * $p < .05$, ** $p < .01$.

hoc contrast analysis revealed that the blindfolded participants (M = 5.58, SE = 0.24) perceived significantly higher PAU than the blind participants (M = 4.00, SE = 0.25) (Table 6). Because Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(5) = 16.752$, $p \leq 0.005$, we report the results with the Greenhouse-Geisser correction. The two-way interaction between *sight capacity* and *test condition* was not significant [F(5.180, 173.519) = 2.167, $p \leq 0.058$] (Table 8). The three-way interaction between *sight capacity*, *test condition,* and *experiment* was also not significant [F (5.180, 173.519) = 0.944, $p \leq 0.456$] (Table 8).

Perceived emotional interdependence (PEI). A significant main effect was observed for *sight capacity* [F(2, 67) = 5.585, $p \leq 0.006$, $\eta_p^2 = 0.143$] (Table 7). The blindfolded participants (M = 5.26, SE = 0.26) perceived a significantly higher PEI than the blind participants (M = 4.83, SE = 0.27) (Table 6). A significant two-way interaction effect was also observed between *sight capacity* and *test condition,* [F(6, 201) = 2.156, $p \leq .049$, $\eta_p^2 = 0.060$] (Table 8). The three-way interaction *sight capacity*, *test condition,* and *experiment* was not significant [F(6, 201) = 1.626, $p \leq 0.142$] (Table 8).

Perceived behavioural interdependence (PBI). A significant main effect was observed for *sight capacity* [F(2, 67) = 12.345, $p < .001$, $\eta_p^2 = 0.269$] (Table 7). The post hoc contrast analysis revealed that the blindfolded participants (M = 5.77, SE = 0.22) perceived significantly higher PBI than the blind participants (M = 4.50, SE = 0.23) (Table 6). Because Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(5) = 11.104$, $p \leq 0.049$, we report the results with the Greenhouse-Geisser correction. A significant two-way interaction effect was also observed between *sight capacity* and *test condition*, [F(6, 201) = 4.279, $p \leq .001$, $\eta_p^2 = 0.113$] (Table 8). The three-way interaction between *sight capacity*, *test condition,* and *experiment* was not significant [F(5.421, 181.597) = 2.014, $p \leq 0.073$] (Table 8).

4.2. Qualitative analysis

We analyzed the qualitative data using a standard analysis method named Qualitative Content Analysis (QCA) (Hsieh and Shannon, 2005). Using QCA we can interpret the content of textual data set through systematic coding and categorising. Since we focused on the communication experience of all blindfolded, the relevant quotes were selected from the interview transcripts. Finally, we reported 24 quotes from 20 blindfolded participants regarding their experience of being blindfolded. We sorted all 24 quotes into seven categories: (1) enthusiastic to speak in conversations (five quotes), (2) nonverbal behaviours (three quotes), (3) rely on hearing (five quotes), and (4) differences between familiar and unfamiliar people (three quotes), (5) visualise the conversation partner (three quotes), (6) perceive conversation duration (two quotes), and (7) others (three quotes). We used 'BFx' as an abbreviation to describe a quote from a blindfolded participant with the number x as follows:

Enthusiastic to speak in conversations: Five quotes show the blindfolded participants tended to be more engaged in conversations than usual. The typical quote is 'I listen to the partner more attentively than usual, and I am fully engaged in conversations' (BF19). BF9 emphasised that she concentrated on answering in discussions, and even ignored self-image (e.g. appropriate facial expressions and postures that she usually cares about a lot). BF11 also said: 'I am enthusiastic to talk to my conversation partner. I try to guess his first impression on me through talking.' BF17 explained the reasons for enthusiastic talking: 'I try to explain everything merely relying on verbal communication. If I am not blindfolded, I can express myself by using gaze or eye contact to emphasize an intention. Now I cannot see, so I become enthusiastic to speak in conversations.'

Rely on hearing: Five quotes describe the blindfolded participants heavily relied on hearing in conversations, and even ignored the sense of smell. Just as BF9 said: 'I feel I communicate with a loudspeaker. It does not matter whether my partner is present or absent, female or male […] I only rely on hearing in conversations, and almost lost the sensitivity toward other nonverbal cues. For example, I do not realise the perfume scent from my partner in conversations.' The other two quotes show a similar viewpoint. 'Hearing is the only way for me to sense my conversation partner' (BF17) and 'I rely on hearing in conversations' (BF19).

Nonverbal behaviors: Three quotes are regarding the nonverbal behaviours of the blindfolded participants in conversations. Two quotes describe the blindfolded participants still keeping gaze behaviours as usual in talking. BF15 said: 'I am talking with my eyes open despite being blindfolded.' 'Although I cannot see anything, I still keep gaze behaviours (e.g. look down unconsciously)' (BF9). Besides, BF5 mentioned if he could see in conversations, he might have some body language (e.g. showing an agreement by imitating behaviours of the conversation partner and synchronising behaviours of both sides).

Differences between familiar and unfamiliar people: Three quotes show the different attitudes of the blindfolded participants towards familiar and unfamiliar conversation partners. If the conversation partner is a stranger, BF3 said: 'I am a little worried about my appearance of wearing the E-Gaze.' If he or she is a familiar person, 'I become relaxed in conversations, since we have a good relationship […]' (BF3). BF13 also mentioned: 'I feel almost the same as usual in conversations. Because I am familiar with my conversation partner, and often communicate with her.'

Visualize the conversation partner: Three quotes describe the blindfolded participants visualising the conversation partner in their minds. One typical example is 'as being familiar with her voice, I start to imagine that she is sitting there, and visualize her appearance according to her voice. In my imagination, she has wavy, brown, and shoulder-length cut hair. Her face is a little bit bigger, and she looks like a sporty girl. Her voice is similar to one of my classmates, so I can visualize her face by using the facial appearance of my classmate' (BF7). BF5 also said: 'I cannot see the conversation partner, but I imagine that he sits there and looks like Rodin's Thinker.' 'I can imagine a conversation partner [is sitting in front of me], so [being blindfolded] does not affect me too much' (BF15).

Perceive conversation duration: Two quotes describe how the blindfolded participants perceived conversation duration. BF7 said: 'I feel time becomes very fast in each conversation.' BF5 explained in detail: 'We have four conversations with the same duration in the test. However, I do not feel each conversation has the same duration. I think my perception of the duration depends on the amount of information exchanged in conversations. If a large amount of information has been exchanged in conversations, I feel the duration is long. Similarly, the duration is perceived shorter if we have less information exchanged in conversations.'

Others: Three quotes report other special perceptions of being blindfolded in conversations. For example, BF15 said: 'We keep silent at the beginning of the conversation. If I can see my partner is hesitant to speak, I will speak first without any hesitation.' 'If my conversation partner deliberately hides her emotion, it is not easy for me to observe' (BF15).

5. Discussion

In our research studies, we investigate whether blindfolded people can be used to represent blind people in social interactions, especially in conversation scenarios. We conducted an experimental study to answer our research question: Is there a significant behavioural difference in the communication quality between blind and blindfolded people?

In the experimental study, we contribute to exploring the specific context of use and find out whether the communication quality of blindfolded people could be representative of that of blind people. Our quantitative findings demonstrate that there is a significant main effect of *sight capacity* on communication quality (Table 6). More specifically, the blindfolded participants experienced significantly higher co-presence, perceived message understanding, perceived affective understanding, and perceived behavioural interdependence (Figure 3). We also observed that the communication quality of blind people was almost equal to sighted people (Figure 3). It indicates that the results of testing sighted people on communication quality can be largely generalised to blind people. In addition to the main effect, we also report some significant findings regarding two-way and three-way interaction effects. However, these interaction effects included the factor *test condition*. It is not a real test condition for E1 and E2. Instead, the factor *test condition* comes from combining real test conditions of the E1 dataset and E2 dataset, which has been described in Section 4.1 in detail. Thus, the interaction effects including *test condition* are not very useful to explain the results. We still focus on the analysis of the main effect of *sight capacity*.

According to Silverman (2015), a short-term blindness simulation often takes a negative impact on blindfolded people, making them feel frustrated and fearful. The user experience of blindfolded people is even worse than real blind people. Because blind people are familiar with performing tasks without sight. Some blind people never have sight, particularly a young infant. Generally, blindfolded people perceive a worse user experience than blind people due to a short-term disability simulation. But, in our case, the situation is the opposite. Our quantitative results found that although most blind participants in our study were born blind, and all of them lost sight before age of 12 (Table 4), they still perceived lower communication quality than the blindfolded participants. This could be partly explained by experimental findings from Sak-Wernicka (2016). Compared with blindfolded and sighted participants, blind participants experienced more difficulties in recognising and judging emotions during natural communication, especially concealed and negative emotions. Blind participants' difficulties in detecting emotions would affect their communication quality.

In line with the quantitative results, qualitative findings also indicate a behavioural difference between blind and blindfolded people in the conversation scenario. Here, we fully discuss the user experience of blindfolded participants, which strongly supports the significant behavioural differences between blind and blindfolded people in conversations. Specifically, user experience includes the inner states of the participants, such as predispositions, expectations, needs, motivations, and emotions (Hassenzahl and Tractinsky 2006). We gained relevant insights from the blindfolded participants' perceptions of being blindfolded (Section 4.2.2). In conversations, the blindfolded participants completely relied on listening, and almost lost sensitivity towards the nonverbal cues. They became enthusiastic to speak in conversations, and earnest to get feedback through the partner's utterances. Although they could not see, some of them still kept their gaze behaviours as usual. Some participants even tried to visualise the partner's face based on his/her voice. Such behaviours were quite different from blind people. According to previous studies (Qiu et al. 2015; Qiu et al. 2020), in face-to-face communication, blind people are very sensitive to smell and even could distinguish subtle olfactory differences in their friends. Due to the loss of vision, the human brain can gradually learn to process information normally acquired through vision by using other sensory modalities (e.g. smell and hearing) (Bach-y-Rita and W Kercel 2003; Jóhannesson et al. 2016; Buimer et al. 2018). Such a long-term behavioural change cannot occur to the blindfolded participants in a controlled experiment for a short period. During a short term of losing sight, blindfolded participants became more concentrated on conversations, possibly causing an increase in communication quality. Overall, the qualitative findings from the blindfolded participants provide us with evidence that the perception and behaviours of the blind participants differ from the blindfolded participants during face-to-face communication.

Our experimental study yielded rich quantitative and qualitative data. Nevertheless, there are several limitations to this research. First, the age, education background, and level of spoken language were not well balanced in the between-group tests. Thus, in data analysis, we treat 'gender' as a confounding factor and 'age' as a co-variate. In our future work, we also need to take the level of the spoken language of the participants into account. Second, four conditions of study one and study two should require 24 orders of treatments ($4 \times 3 \times 2 \times 1$) in a within-subjects design, and the number of the participant pairs must be a multiple of 24. In our study, there was a limited number of truly blind students that could meet all criteria (e.g. age, intelligence quotient, and without any other disabilities). Thus, 10 orders of treatments for the blind-sighted pairs was a compromise in our study. In future work, we can use the Latin square design if the number of blind participants is not sufficient. Third, due to a limited number of blind participants, we did not distinguish between early blind and late blind people. They might have different communication experience. We should consider this point in our future work. Fourth, the level and quality of education, life skills, and social interactions in schools for blind students in many places are indeed significantly lower than schools attended by sighted students. We acknowledge this important observation and have documented it for our future work. Going forward, we aim to dig deeper to identify other factors that may affect the quality of communication between blind and blindfolded people.

In sum, our quantitative findings demonstrated that there was a significant difference in the communication quality between blind and blindfolded participants. We argue that blind people cannot be replaced by blindfolded people in evaluations with conversations, since their behaviours and perceptions quite differ.

6. Conclusion

This paper presents an experimental study to investigate the alternative-user approach. Specifically, we conducted our experimental studies with 20 blind-sighted pairs and 20 blindfolded-sighted pairs. The results demonstrated that blind and blindfolded participants perceived significantly different communication quality when they talked to sighted participants. We suggest that if research studies want to replace blind with blindfolded people, make sure the proper context of use provides the evidence. For all contexts of use without proper validation studies, be careful to replace blind with blindfolded people. We need more validation studies for these contexts of use.

We tested a smaller sample that had more similar visual impairment. Based on the International Classification of Diseases 11 (World Health Organization 2022), vision impairment is classified into four categories: (1) mild, (2) moderate, (3) severe, and (4) blindness. In the new data analysis, we filtered participants with (1) mild, (3) moderate, and (3) severe visual impairments and only kept (4) blindness, which fulfilled the requirement of a smaller sample with more similar visual impairments.

In the SPSS dataset we defined a variable named 'VisionConditions' which included five values (i.e. 0 = 'Sighted'; 1 = 'Moderate visual impairment'; 2 = 'Severe visual impairment'; 3 = 'Blindness'; 4 = 'Blindfolded'). In the full-cases condition, we include all five values in the dataset (N = 80). In the selected-cases condition, we include the value of 0, 3, and 4 of 'VisionConditions' (N = 69). Since 'sight-capacity' is our prime independent variable, we report the main effect of 'sight-capacity'. Table 7 and Table A2 report GLM results in summary of the main effects of 'sight-capacity' on the communication quality in the full-cases condition and the selected-cases condition, respectively. We found that the test power for each dependent variable in Table 7 and Table A2 are quite similar and do not change too much. Since test power ranges from 0-1, the observed power in Table 7 has already been very high. Results indicate that no matter a smaller sample that had a more similar visual impairment or a broad sample of visual impairment, it does not influence the test power extensively. Our findings about the whole sample are well-suited to be generalised.

Note

1. https://theeyetribe.com/

Acknowledgments

Many thanks to Xiang Cheng, Liang Zang, and Zhi Liu for helping us organise the participants from Yangzhou Special Education School and Jiangsu College of Tourism.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Humanity and Social Science Youth foundation of Ministry of Education of China: [grant no 22YJC760073]; Shanghai Pujiang Program: [grant no 2020PJC071]; Shanghai Jiao Tong University: [grant no WF220543011].

ORCID

Shi Qiu http://orcid.org/0000-0001-6699-1693

References

Alkhanifer, A. A., and S. Ludi. 2015. "Developing SAGAT Probes to Evaluate Blind Individuals' Situation Awareness When Traveling Indoor Environments." In *Proceedings of the 17th International Conference on Computers and Accessibility*, 305–306. New York, USA: ACM Press. doi:10.1145/2700648.2811350.

Anagnostakis, G., M. Antoniou, E. Kardamitsi, T. Sachinidis, P. Koutsabasis, M. Stavrakis, S. Vosinakis, and D. Zissis. 2016. "Accessible Museum Collections for the Visually Impaired: Combining Tactile Exploration, Audio Descriptions and Mobile Gestures." Proceedings of the 18th international conference on human-computer interaction with mobile devices and services adjunct, MobileHCI 2016 (pp. 1021–1025). doi:10.1145/2957265.2963118.

Apostolopoulos, I., N. Fallah, E. Folmer, and K. E. Bekris. 2014. "Integrated Online Localization and Navigation for People with Visual Impairments Using Smart Phones." *ACM Transactions on Interactive Intelligent Systems (TiiS)* 3 (4): 1–28. doi:10.1145/2499669.

Awada, A., Y. Bou Issa, J. Tekli, and R. Chbeir. 2013. "Evaluation of Touch Screen Vibration Accessibility for Blind Users." In *Proceedings of the 15th International Conference on Computers and Accessibility*, 1–2. New York, USA: ACM Press. doi:10.1145/2513383.2513430.

Bach-y-Rita, P., and S. W Kercel. 2003. "Sensory Substitution and the Human-Machine Interface." *Trends in Cognitive Sciences* 7 (12): 541–546. doi:10.1016/j.tics.2003.10.013.

Balata, J., Z. Mikovec, and L. Neoproud. 2015. "BlindCamera: Central and Golden-Ratio Composition for Blind Photographers." Proceedings of the mulitimedia, interaction, design and innnovation (pp. 8:1–8:8). doi:10.1145/2814464.2814472.

Biocca, F., and C. Harms. 2002. "Defining and Measuring Social Presence - Contribution to the Networked Minds Theory and Measure." Proceedings of PRESENCE 2002 (pp. 1–36).

Biocca, F., C. Harms, and J. K. Burgoon. 2003. "Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria." *Presence: Teleoperators and Virtual Environments* 12 (5): 456–480. doi:10.1162/105474603322761270.

Buimer, H. P., M. Bittner, T. Kostelijk, T. M. Van Der Geest, A. Nemri, R. J. A. Van Wezel, and Y. Zhao. 2018. "Conveying Facial Expressions to Blind and Visually Impaired Persons Through a Wearable Vibrotactile

Device." *PLoS ONE* 13 (3): e0194737. doi:10.1371/journal.pone.0194737.
China Association of the Blind. 2013. Classification and grade of the Chinese disabled people. Retrieved from www.zgmx.org.cn.
Dolničar, V., D. Grošelj, M. Filipovič Hrast, V. Vehovar, and A. Petrovčič. 2018. "The Role of Social Support Networks in Proxy Internet use from the Intergenerational Solidarity Perspective." *Telematics and Informatics* 35 (2): 305–317. doi:10.1016/j.tele.2017.12.005.
Douglas, S. A., and S. Willson. 2007. "Haptic Comparison of Size (Relative Magnitude) in Blind and Sighted People." In *Proceedings of the 9th International Conference on Computers and Accessibility*, 83–90. New York, USA: ACM Press. doi:10.1145/1296843.1296859.
Forlizzi, J., and K. Battarbee. 2004. "Understanding Experience in Interactive Systems." In *DIS2004 - Designing Interactive Systems: Across the Spectrum*, 261–268. Association for Computing Machinery. doi:10.1145/1013115.1013152.
Ge, F. 2017. *The Role of Culture in Close Relationships: East-West Differences in Communication and Emotion Regulation*. University of Massachusetts Amherst.
Ghent, L. 1961. "Developmental Changes in Tactual Thresholds on Dominant and Nondominant Sides." *Journal of Comparative and Physiological Psychology* 54 (6): 670–673. doi:10.1037/h0047319.
Gomez, J. D., S. Mohammed, G. Bologna, and T. Pun. 2011. "Toward 3D Scene Understanding via Audio-Description: Kinect-IPad Fusion for the Visually Impaired." In *Proceedings of the 13th International Conference on Computers and Accessibility*, 293–294. New York, USA: ACM Press. doi:10.1145/2049536.2049613.
Hall, E. T. 1963. "A System for the Notation of Proxemic Behavior." *American Anthropologist* 65 (5): 1003–1026. doi:10.1525/aa.1963.65.5.02a00020.
Harms, C., and F. Biocca. 2004. "Internal Consistency and Reliability of the Networked Minds Measure of Social Presence." In *Seventh Annual International Workshop: Presence 2004*, 246–251. Valencia: Universidad Politecnica de Valencia.
Hassenzahl, M., and N. Tractinsky. 2006. "User Experience - A Research Agenda." *Behaviour and Information Technology* 25 (2): 91–97. doi:10.1080/01449290500330331.
Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687
Huang, Y. Y., J. Moll, E. L. Sallnäs, and Y. Sundblad. 2012. "Auditory Feedback in Haptic Collaborative Interfaces." *International Journal of Human Computer Studies* 70 (4): 257–270. doi:10.1016/j.ijhcs.2011.11.006.
IELTS Speaking Module. 2012. Retrieved from http://www.goodluckielts.com/IELTSspeaking-%0Atopics-2.html.
Jóhannesson, ÓI, O. Balan, R. Unnthorsson, A. Moldoveanu, and Á Kristjánsson. 2016. "The Sound of Vision Project: On the Feasibility of an Audio-Haptic Representation of the Environment, for the Visually Impaired." *Brain Sciences* 6: 3. doi:10.3390/brainsci6030020.
Judd, C. M., G. H. McClelland, and C. S. Ryan. 2017. *Data Analysis: A Model Comparison Approach to Regression, ANOVA, and Beyond*. Routledge.
Kamel, H. M., and J. A. Landay. 2002. "Sketching Images Eyes-Free: A Grid-Based Dynamic Drawing Tool for the Blind." In *Proceedings of the Fifth International ACM Conference on Assistive Technologies*, 33–40. New York, USA: ACM Press. doi:10.1145/638249.638258.
Kammoun, S., W. Bouhani, and M. Jemni. 2016. "Sole Based Tactile Information Display for Visually Impaired Pedestrian Navigation." In *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 1–4. New York, USA: ACM Press. doi:10.1145/2769493.2769561.
Kemp, N. J., and D. R. Rutter. 1986. "Social Interaction in Blind People: An Experimental Analysis." *Human Relations* 39 (3): 195–210. doi:10.1177/001872678603900302.
Kukull, W., and M. Ganguli. 2012. "Generalizability: The Trees, the Forest, and the low-Hanging Fruit." *AAN Enterprises* 78 (23): 1886–1891.
Lawton, C. A., and D. W. Hatcher. 2005. "Gender Differences in Integration of Images in Visuospatial Memory." *Sex Roles* 53 (9–10): 717–725. doi:10.1007/s11199-005-7736-1.
Leo, F., C. Baccelliere, A. Waszkielewicz, E. Cocchi, and L. Brayda. 2018. "Tactile Symbol Discrimination on a Small Pin-Array Display." In *Proceedings of the 2018 Workshop on Multimedia for Accessible Human Computer Interface*, 9–15. New York, USA: ACM Press. doi:10.1145/3264856.3264858.
McGookin, D., S. Brewster, and J. Weiwei. 2008. "Investigating Touchscreen Accessibility for People with Visual Impairments." Proceedings of the 5th nordic conference on human-computer interaction: building bridges. (pp. 298–307).
Mihara, Y., A. Sugimoto, E. Shibayama, and S. Takahashi. 2005. "An Interactive Braille-Recognition System for the Visually Impaired Based on a Portable Camera." In *CHI'05 Extended Abstracts on Human Factors in Computing Systems*, 1653–1656. doi:10.1145/1056808.1056989.
Montgomery, B. M. 1988. "Quality Communication in Personal Relationships." In *Handbook of Personal Relationships:Theory, Research and Interventions*, 343–359. Oxford, England: John Wiley & Sons.
Nario-Redmond, M. R. 2010. "Cultural Stereotypes of Disabled and non-Disabled men and Women: Consensus for Global Category Representations and Diagnostic Domains." *British Journal of Social Psychology* 49 (3): 471–488. doi:10.1348/014466609X468411.
Osawa, H. 2014. "Emotional Cyborg: Human Extension with Agency for Emotional Labor." *Proceedings of the 2014 ACM/IEEE international conference on human-robot interaction* (p. 108). doi:10.1145/2559636.2559647.
Poláček, O., T. Grill, and M. Tscheligi. 2012. "Towards a Navigation System for Blind People: A Wizard of oz Study." *ACM SIGACCESS Accessibility and Computing* 104: 12–29. doi:10.1145/2388818.2388820.
Przybylski, A. K., and N. Weinstein. 2013. "Can you Connect with me now? How the Presence of Mobile Communication Technology Influences Face-to-Face Conversation Quality." *Journal of Social and Personal Relationships* 30 (3): 237–246. doi:10.1177/0265407512453827.

Qiu, S., P. An, J. Hu, T. Han, and M. Rauterberg. 2020. "Understanding Visually Impaired People's Experiences of Social Signal Perception in Face-to-Face Communication." *Universal Access in the Information Society* 19 (4): 873–890. doi:10.1007/s10209-019-00698-3.

Qiu, S., P. An, K. Kang, J. Hu, T. Han, and M. Rauterberg. 2021. "Investigating Socially Assistive Systems from System Design and Evaluation: A Systematic Review." *Universal Access in the Information Society*, 1–25. doi:10.1007/s10209-021-00852-w.

Qiu, S., P. An, K. Kang, J. Hu, T. Han, and M. Rauterberg. 2022. "A Review of Data Gathering Methods for Evaluating Socially Assistive Systems." *Sensors* 22: 1. doi:10.1109/JSEN.2022.3226932.

Qiu, S., S. A. Anas, H. Osawa, M. Rauterberg, and J. Hu. 2016a. "E-Gaze Glasses: Simulating Natural Gazes for Blind People." Proceedings of the tenth international conference on tangible, embedded, and embodied interaction (pp. 563–569). doi:10.1145/2839462.2856518.

Qiu, S., T. Han, M. Rauterberg, and J. Hu. 2018. "Impact of Simulated Gaze Gestures on Social Interaction for People with Visual Impairments." *Transdisciplinary engineering methods for social innovation of industry 4.0: proceedings of the 25th ISPE Inc. international conference on transdisciplinary engineering* (Vol. 7, pp. 249–258). doi:10.3233/978-1-61499-898-3-249.

Qiu, S., J. Hu, T. Han, H. Osawa, and M. Rauterberg. 2020a. "An Evaluation of a Wearable Assistive Device for Augmenting Social Interactions." *IEEE Access* 8: 164661–164677. doi:10.1109/ACCESS.2020.3022425.

Qiu, S., J. Hu, T. Han, H. Osawa, and M. Rauterberg. 2020b. "Social Glasses: Simulating Interactive Gaze for Visually Impaired People in Face-to-Face Communication." *International Journal of Human-Computer Interaction* 36 (9): 839–855. doi:10.1080/10447318.2019.1696513.

Qiu, S., J. Hu, and M. Rauterberg. 2015a. "Nonverbal Signals for Face-to-Face Communication Between the Blind and the Sighted." International conference on enabling access for persons with visual impairment (pp. 157–165).

Qiu, S., H. Osawa, J. Hu, and M. Rauterberg. 2015b. "E-Gaze: Create Gaze Communication for People with Visual Disability." International conference on human-agent interaction (pp. 199–202). doi:10.1145/2814940.2814974.

Qiu, S., M. Rauterberg, and J. Hu. 2016b. "Tactile Band: Accessing Gaze Signals from the Sighted in Face-to-Face Communication." In *Proceedings of International Conference on Tangible, Embedded, and Embodied Interaction*, 556–562. New York, USA: ACM Press. doi:10.1145/2839462.2856520.

Rosel, J., A. Caballer, P. Jara, and J. C. Oliver. 2005. "Verbalism in the Narrative Language of Children who are Blind and Sighted." *Journal of Visual Impairment and Blindness* 99 (7): 413–425. doi:10.1177/0145482X0509900704.

Sak-Wernicka, J. 2016. "Differences in Recognition and Judgement of Negative Emotions Between Blind and Sighted People." *Roczniki Humanistyczne* 64 (10): 113–124. doi:10.18290/rh.2016.64.10-7.

Schmutz, S., A. Sonderegger, and J. Sauer. 2017. "Implementing Recommendations from Web Accessibility Guidelines: A Comparative Study of Nondisabled Users and Users with Visual Impairments." *Human Factors* 59 (6): 956–972. doi:10.1177/0018720817708397.

Sears, A., and V. L. Hanson. 2011. "Representing Users in Accessibility Research." Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2235–2238). doi:10.1145/1978942.1979268.

Silverman, Arielle Michal. 2015. "The Perils of Playing Blind: Problems with Blindness Simulation and a Better Way to Teach About Blindness." *Journal of Blindness Innovation and Research* 5: 2. doi:10.5241/5-81.

Silverman, Arielle M., J. D. Gwinn, and L. Van Boven. 2015. "Stumbling in Their Shoes: Disability Simulations Reduce Judged Capabilities of Disabled People." *Social Psychological and Personality Science* 6 (4): 464–471. doi:10.1177/1948550614559650.

Smith, B. A., and S. K. Nayar. 2018. "The RAD: Making Racing Games Equivalently Accessible to People Who Are Blind." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. New York, USA: ACM Press. doi:10.1145/3173574.3174090.

Tang, T. J. J., and W. H. Li. 2014. "An Assistive EyeWear Prototype That Interactively Converts 3D Object Locations Into Spatial Audio." Proceedings of the 2014 ACM international symposium on wearable computers (pp. 119–126). doi:10.1145/2634317.2634318.

Tinwala, H., and I. S. MacKenzie. 2008. "Letterscroll: Text Entry Using a Wheel for Visually Impaired Users." In *Proceeding of the Twenty-Sixth Annual CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 3153–3158. New York, USA: ACM Press. doi:10.1145/1358628.1358823.

Wilson, G., S. Brewster, H. Caltenco, C. Magnusson, S. Finocchietti, G. Baud-Bovy, and M. Gori. 2015. "Effects of Sound Type on Recreating the Trajectory of a Moving Source." In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, 1645–1650. New York, USA: ACM Press. doi:10.1145/2702613.2732821.

World Health Organization. 2022. ICD-11 for Mortality and Morbidity Statistics. Retrieved October 28, 2022, from https://icd.who.int/browse11/l-m/en.

Zeng, L., and G. Weber. 2015. "A Pilot Study of Collaborative Accessibility: How Blind People Find an Entrance." Proceedings of the 17th international conference on human-computer interaction with mobile devices and services (pp. 347–356). doi:10.1145/2785830.2785875.

Appendix

Table A1. Summary of the paper lists and features.

Reference	Contexts of Use	Feedback Modality	Participants	Measures	Results	Category (1, 2; 3)[a]
Zeng and Weber 2015	The system to help blind people find entrances independently in unknown regions	Auditory and vibration	Five blind and five blindfolded participants	- User task performance (e.g. time spent, effective walking speed, success rate) - Questionnaires (e.g. evaluate the method by four five-point scales and two open questions)	The performances of the blindfolded participants were worse than the blind participants.	Category 3
Anagnostakis et al. 2016	The system to explore exhibit replicas presented in an exhibition room	Auditory and touch	Two blind, four visually impaired, and four blindfolded participants	User task performance (e.g. task success, average task time, total user interaction Time, and the total number of errors)	The blindfolded participants did not fully concentrate on the audio instructions in comparison to the blind participants.	Category 3
Tang and Li 2014	An assistive prototype to detect objects on planar surfaces and find their 3D locations using spatial audio	Auditory (3D sounds)	Twelve blindfolded participants	User task performance (e.g. user accuracy)	The prototype was able to provide a useful level of object localisation assistance.	Category 1
Mihara et al. 2005	Braille recognition system	Auditory	Six blindfolded participants	- Functional validation (i.e. recognition rates) - User task performance (e.g. usability)	Six blindfolded participants were able to interact with the system and recognised the meaning of the buttons that were identified.	Category 1
Balata, Mikovec, and Neoproud 2015	Blind camera system	Auditory and vibration	Twelve blindfolded participants	- User task performance (e.g. completion time of taking a photo) - Questionnaires (e.g. aesthetic quality)	The combination of golden ratio composition with voice feedback produced the best aiming time and completion time, as well as the best aesthetic quality of photos.	Category 1
Alkhanifer and Ludi 2015	Situation Awareness Global Assessment Technique (SAGAT)	Auditory	Three blindfolded participants	User task performance	The system could be beneficial to facilitating blind people's situational awareness when traveling in unfamiliar indoor environments.	Category 1
Wilson et al. 2015	An audio bracelet is worn by a blind person to inform his/her movements to aid spatial cognition rehabilitation. It is also worn by a sighted person to inform the blind person of others' movement in space.	Auditory	Six blindfolded participants	User task performance (e.g. endpoint distance, trajectory deviation, total distance travelled)	All of the initial sounds facilitate the recreation of 2D horizontal movement trajectories similarly well, although *birdsong* was problematic and *speech* and *waves* were more promising.	Category 1
Awada et al. 2013	The system to access simple contour-based images	Vibration	Six blind participants since birth, seven blind participants after birth, and sixteen blindfolded participants	User task performance (e.g. percentage of correct answers)	Blindfolded participants performed better: with 8.11% more correct answers than blind participants since birth, and 28.29% more correct answers than blind participants after birth.	Category 3
Douglas and Willson 2007	3D wall system to measure human performance	Haptic	Eleven blind and eleven blindfolded participants	User task performance (e.g. accuracy)	Blind participants took 50% longer with equivalent accuracy to sighted participants.	Category 3
Apostolopoulos et al. 2014	Indoor navigation system	Auditory	Study 1: One blind and nine blindfolded participants Study 2: Eight	- Questionnaires - Observational data - User task performance	Study 1: determine the feasibility of guiding blind people through an indoor environment using a	Category 2

(*Continued*)

Table A1. Continued.

Reference	Contexts of Use	Feedback Modality	Participants	Measures	Results	Category (1, 2; 3)[a]
			blindfolded participants Study 3: Six blind participants		simple and interactive sensing approach achievable on a smartphone. Study 2: evaluate the different methods for estimating the user's step length. Study 3: 85% of the paths were completed successfully by real blind participants.	
McGookin, Brewster, and Weiwei 2008	Touch screen	Auditory	One blind and twelve blindfolded participants	User task performance (e.g. time taken, errors)	The qualitative evaluation provided guidelines for future designers, to help them exploit the potential of touchscreen technology for blind people.	Category 1
Kamel and Landay 2002	A drawing tool for transforming a mouse-based graphical user interface into an auditory interface	Auditory	Eight blind and Eight blindfolded participants	User task performance (e.g. task completion time, level of confidence, and quality of drawing)	Blind participants performed as well or better than the sighted participants on all three tasks for all the measurements.	Category 3
Kammoun et al., 2016	The tactile sole with 5 vibrators presents spatial information for blind people as an orientation aid	Haptic	Two blindfolded participants	User task performance (e.g. task completion time, tactile instructions detection)	Participants have expressed their satisfaction with the intensity and frequency used.	Category 1
Qiu et al. 2016a	The tactile band that enables a blind person to feel attention (gaze signals) from a sighted person	Tactile	Fifteen sighted and fifteen blindfolded participants	Questionnaires (e.g. relationship quality and partner closeness)	No significant difference in engagement was found across three test conditions.	Category 1
Leo et al. 2018	Tactile symbols that can be used by engineers, designers, and rehabilitation practitioners in representing tactile maps and diagrams on small-size tactile displays	Tactile	Nineteen blind, twenty blindfolded low vision and twenty-two blindfolded sighted participants	User task performance (e.g. accuracy, response time)	The blind participants were much faster to identify tactile symbols than the blindfolded low vision and the blindfolded full vision participants.	Category 3
Tinwala and MacKenzie 2008	Four text entry techniques for blind users	Auditory	Seven blindfolded participants	- User task performance (e.g. speed and accuracy) - Questionnaire	An experiment compared two methods (M1 and M4) with seven blindfolded participants who entered a total of 84 phrases of text. M4 had a significant improvement in text entry rate.	Category 1
Smith and Nayar 2018	An audio-based user interface that allows blind players to play the same racing games as sighted players	Auditory	Three blind and twelve blindfolded participants	Questionnaires (e.g. user interface, awareness of upcoming terms)	Racing Auditory Display (RAD) allowed a blind gamer to race as well on a complex racetrack as casual sighted players.	Category 2
Gomez et al. 2011	An electronic travel aid	Auditory	One blindfolded participant	User task performance	Objects' location with sound enables the blindfolded user to build a mental perception of the environment.	Category 1
Poláček, Grill, and Tscheligi 2012	An indoor navigation system	Auditory and vibration	Twenty blindfolded participants	- Interviews - Observations (e.g. evaluator notes) - System log (e.g. all interactions with the Mobile Wizard) - Questionnaires (e.g. ease of use, satisfaction)	The voice commands chosen for navigation are almost complete and can be used for the follow-up study.	Category 1

[a]Category 1: Only Including Blindfolded Participants; Category 2: Including Both Blind and Blindfolded Participants, But without a Comparison; Category 3: A Comparison of Blind and Blindfolded Participants.

Table A2. GLM results summary of main effects of 'sight-capacity' on the communication quality among sighted, blindfolded, and blind only participants (N = 69).

Source	Measure	*SS*	*df*	*MS*	*F*	*p*	Observed Power[a]
'sight- capacity'	Co-presence	24.948	2	12.474	7.733	**0.001****	0.939
	Attention allocation	43.199	2	21.600	9.978	**<0.001****	0.980
	Perceived message understanding	118.085	2	59.043	20.251	**<0.001****	1.000
	Perceived affective understanding	79.203	2	39.601	9.727	**<0.001****	0.977
	Perceived emotional interdependence	46.783	2	23.391	5.379	**0.007****	0.823
	Perceived behavioural interdependence	73.883	2	36.941	11.284	**<0.001****	0.990

Significant group difference; * $p < .05$, ** $p < .01$.
[a]Computed using alpha = .05