



HAL
open science

Guidelines for Inclusive Avatars and Agents: How Persons with Visual Impairments Detect and Recognize Others and Their Activities

Lauren Thévin, Tonja Machulla

► **To cite this version:**

Lauren Thévin, Tonja Machulla. Guidelines for Inclusive Avatars and Agents: How Persons with Visual Impairments Detect and Recognize Others and Their Activities. ICCHP International Conference on Computers Helping People with Special Needs, Nov 2020, Virtual, Italy. pp.164-175, 10.1007/978-3-030-58796-3_21 . hal-02992013

HAL Id: hal-02992013

<https://hal.science/hal-02992013v1>

Submitted on 6 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Guidelines for Inclusive Avatars and Agents: How Persons with Visual Impairments Detect and Recognize Others and their Activities

Lauren Thevin^{1,2}[0000-0002-5273-2794] and Tonja Machulla²

UCO, Bretagne , France
LMU Munich, Munich 80337, Germany
lthevin@uco.fr, tonja.machulla@ifi.lmu.de

Abstract. Realistic virtual worlds are used in video games, in virtual reality, and to run remote meetings. In many cases, these environments include representations of other humans, either as stand-ins for real humans (avatars) or artificial entities (agents). Presence and individual identity of such virtual characters is usually coded by visual features, such as visibility in certain locations and appearance in terms of looks. For people with visual impairments (VI), this creates a barrier to detecting and identifying co-present characters and interacting with them. To improve the inclusiveness of such social virtual environments, we investigate which cues people with VI use to detect and recognize others and their activities in real-world settings. For this, we conducted an online survey with fifteen participants (adults and children). Our findings indicate an increased reliance on multimodal information: **vision** for silhouette recognition; **audio** for the recognition through pace, white cane, jewelry, breathing, voice and keyboard typing; **sense of smell** for fragrance, food smell and airflow; **tactile information** for length of hair, size, way of guiding or holding the hand and the arm, and the **reactions of a guide-dog**. **Environmental and social cues** indicate if somebody is present: e. g. a light turned on in a room, or somebody answering a question. Many of these cues can already be implemented in virtual environments with avatars and are summarized by us in a set of guidelines.

This is a preprint, see the final version at https://link.springer.com/chapter/10.1007/978-3-030-58796-3_21

1 Introduction

Virtual reality (VR) technology allows users to visit virtual environments and to interact with them. This increasingly includes interactions with virtual representations of other human users and artificial characters controlled by the computer, termed avatars and agents, respectively. Much recent research focuses on improving the appearance of such virtual humans and how different characteristics

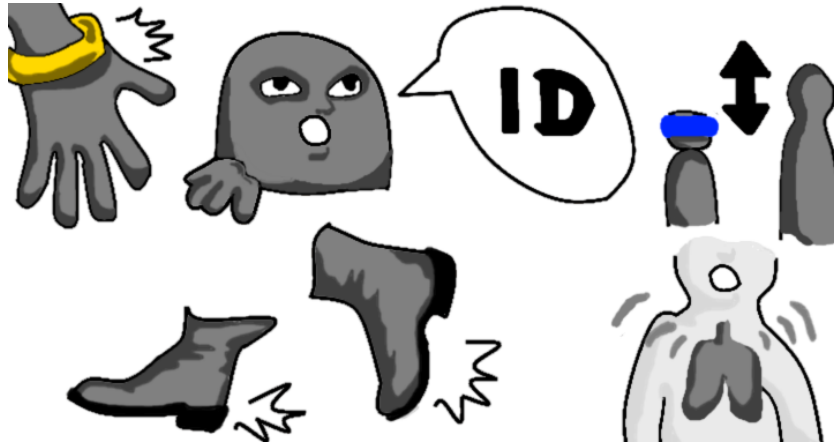


Fig. 1. Avatars with cues usable by sighted and people with visual impairments: wrist-band and jewelry making noise with the movement of the avatar, distinctive voice or laughter and telling the identity, different size (audio and tactile cues), pace, loud or specific breathing pattern.

such as emotional expressiveness, personalization, or motion profiles influence the quality of the interpersonal interaction. As with many other features of VR environments, current representations of human characters rely heavily on visual aspects. As a result, they are less accessible to users with visual impairments. Difficulties include not being aware of the presence of other characters, or not being able to discern their location, identity, or activity. As a result, persons with visual impairments face a higher threshold when it comes to joining in social or collaborative VR experiences. Recently, there have been efforts to increase the accessibility of VR for people with VI [22]. So far, these efforts have focused on improving interactions with physical features of the VR environment. Our work aims to also improve inclusiveness in terms of social and collaborative aspects of VR. To this end, we investigated how people with VI determine presence, identity, and activity of other humans in real-world settings. These factors are likely to influence whether a person will approach others with the goal of engaging in a social interaction: first, a person with VI has to be aware of the presence of potential interaction partners; second, the identity of co-present persons might play a role for certain types of interactions (such as a banter between friends); and third, the activity might determine whether the other person is available for an interaction or not (e.g., a working colleague should not be interrupted). We created an online survey answered by fifteen adults and minors with VI. We present the results of this survey and propose guidelines for the creation of inclusive virtual characters, that can be implemented at present with currently available technological solutions.

2 State of the Art

Representation of Avatars and Agents in Interactive Virtual Environments. Virtual environments support more and more interactive and collaborative activities [14], such as physical and cognitive therapies [10,3], training[11,16,13], gaming, telepresence for meeting [17], and communication with conversational agents [15,12]. All of these applications require the representation of oneself or others in virtual space, in the form of (self-)avatars or virtual agents. These representations rely strongly on visual cues, as these are considered essential for the realism and faithfulness of avatars [18]. Hence, many research efforts focus on photo-realistic avatars as well as visual rendering of life-like motions and non-verbal cues. For example, research on therapy to counter perceived body image distortion in anorexia has used biometrically correct virtual bodies, so the visual representation is realistic [10,3]. Crisis management training [11,16] and surgery training [13] aims to teach collaboration with virtual agents and other users to improve collective work. The representation of others and virtual agents has mostly made use of visual feedback.

Occasionally, multisensory cues are used to increase the realism and individuality of avatars, e. g., in the context of communication. Here, it is increasingly common to additionally implement auditory cues (synthetic voices, Voice Over Internet Protocol (VoIP), etc.; e.g., [16] and [13]). Voices rendering is known as a key part of collaborative skills and avatar realism, combined with visual cues [18]. Rarely, social interactions in VR include other auditory cues and tactile feedback to the user, such as a touch to the forearm or shoulder [5,6].

As of yet, it is unclear whether these developments are conducive to creating inclusive virtual environment. Improving on fine visual detail is unlikely to create strong cues for people with VI, and the current focus on voices ignores other potential auditory sources as cues to the presence and identity of others.

Accessibility of Virtual Reality for People with VI. Multiple projects have specifically investigated VR for people with VI. It is now possible to navigate auditory, sometimes audio-visual, virtual environments without vision using a force-feedback joystick and environmental cues through audio-tactile feedback [9], with a keyboard to explore an virtual audio environment [2], with smartphone applications using and gesture [4,?], and by physically walking in a virtual or augmented room [8,19]. Two works propose virtual white canes: one simulates a whitecane with a braking system [21], and another work designed a virtual white cane through vibration in the hand controller of the VR system [7]. In all of these works, the user interacts with physical aspects of the VE only, such as exploring virtual environments and interact with them. As far as we are aware, there are no virtual characters present, thus ignoring the potential for collaborative experiences.

3 Method and Participants

3.1 Design of the Survey

We created a questionnaire to gather insights into the strategies and cues used by persons with VI in real-world settings when attempting to detect and identify others with the goal of initiating an interaction. To this end, we developed 20 questions, falling into several different categories. The full set of questions is provided table 1. A first set of categories covers four types of mutual awareness in social and collaborative settings: environmental awareness, user and interaction awareness, action awareness, and organization awareness [1]. As summarized in [20], environmental awareness refers to sharing the same space (Q1), action awareness is being aware of the concurrent actions of the other users (Q6), interaction and user awareness concerns knowing the available interactions and knowing who the other users are (Q2, Q3, Q4, Q5, Q8, Q9), and organizational awareness is about knowing and sharing the implications of context on the group (Q7). The question Q8, Q9, Q19 and Q20 aim to identify the good practices for both representing people and initiating interaction in virtual environment. Further, we asked about the use of the senses (Q10 to 15), as well as potential other information found in literature and blogging (Q16-17) and one open question (Q18).

3.2 Protocol

The survey was self-administered. It was organized as follow:

- **presentation of the goal of the study, confidentiality and consent form.** For minors, this part was read and agreed upon by the participants' legal guardian, for legal reasons (it was the only part that required another person to answer for the minor participants).
- **demographic information.** Information regarding age, gender, and the level of vision was collected.
- **survey.** the twenty questions presented table 1.

3.3 Participants

We send our survey to three blind adults first, so they could verify the accessibility and the questions. We then distributed our survey to professionals of two schools for people with visual impairments in France. The schools send our surveys to former students and to parents of current students. Questionnaire responses were returned by fifteen participants in total, i.e., eight adults and seven minors under the age of eighteen years old. For confidentiality reasons and to respect the consent form terms, we do not provide the individual demographics. However, we will describe here individually each demographics dimension (gender, age, and level of vision). Regarding the adults' group, 6 participants were male, 2 were female, and 0 other. The average age is 38 years old, with a

Category	Question
<i>Environmental</i>	Q1. When you enter a room, how do you notice a person is present?
<i>Interaction and User</i>	Q2. When you meet up with people you know for a social appointment in a public space, how do you recognize them?
<i>Interaction and User</i>	Q3. When you meet people you know by coincidence in a public space, how do you recognize them?
<i>Interaction and User</i>	Q4. When you meet people you do not know, how do you know you have never met them?
<i>Interaction and User</i>	Q5. In the first few second after meeting a person for the first time, which factors (such as perceptual cues) influence your first impression of this person (a nice person, an unsympathetic person, a shy person, an extrovert person)?
<i>Action</i>	Q6. How do you identify that a group of people is present in a room you enter? How do you identify the activity of the group?
<i>Organizational</i>	Q7. How do you know if it is socially appropriate to join the activity of a group of people, for instance a conversation?
<i>Interaction and User, Good practices</i>	Q8. When do you find a person easy to recognize?
<i>Interaction and User; Good practices</i>	Q9. When do you find a person difficult to recognize?
<i>Use of the senses</i>	Q10. Do you use vision for it? If you use vision to identify people, what cues do you use to recognize them?
<i>Use of the senses</i>	Q11. Do you use touch for it? If you use touch to identify people, what cues do you use to recognize them?
<i>Use of the senses</i>	Q12. Do you use temperature perception for it? If you use temperature perception to identify people, what cues do you use to recognize them?
<i>Use of the senses</i>	Q13. Do you use air flow perception for it? If you use air flow perception to identify people, what cues do you use to recognize them?
<i>Use of the senses</i>	Q14. Do you use audio perception for it? If you use audio perception to identify people, what cues do you use to recognize them?
<i>Use of the senses</i>	Q15. Do you use smell perception for it? If you use smell perception to identify people, what cues do you use to recognize them?
<i>Other potential cues</i>	Q16. Do you use location and context for it? If you use location and context to identify people, what cues do you use to recognize them?
<i>Other potential cues</i>	Q17. If you have one, do you use information from your guide dog for it? If you use information from your guide dog to identify people, what cues do you use to recognize them?
<i>Open question</i>	Q18. What other cues do you use to recognize people?
<i>Good practices</i>	Q19. When somebody approaches you and you can't identify the person because you either don't know or remember the person, what are the most helpful cues to identify who it is?
<i>Good practices</i>	Q20. What do you want people not to do to alert you from their presence or to identify themselves? For example, because the action startles or annoys you or is not helpful.

Table 1. The twenty questions of the survey, and the categories. Awareness categories: Environmental, Interaction and User, Action and Organizational. Other categories: Use of senses, Other potential cues, Open question, and Good practices.

standard deviation of 18 years. The range is from 19 to 67 years old, with the following values: 19, 20, 21, 26, 48, 48, 52 and 67. 2 participants were low vision, 5 were blind with light perception, and 1 was blind without light perception. Regarding the minors' group, 5 participants were male, 1 was female, and 1 was other. The average age is 12 years old, with a standard deviation of 4 years. The range is from 6 to 16 years old, with the following values: 6, 10, 10, 10, 14, 15, 16. 1 participant was low vision, 2 were blind with light perception, and 4 were blind without light perception.

4 Results

In the following, we analyze participants' responses in terms of their own observational and behavioral strategies as well as strategies by others that are used to: detect and subsequently identify a person as known or unknown (4.1-4.3), and identify their activity (4.4). In section 4.5, we summarize strategies that others might use to initiate an interaction. Numbers in round brackets indicate the number of participants who mentioned a particular fact.

4.1 Detecting the presence of others

Observation strategies: Vision is used to identify silhouettes (2 participants), although audition is more commonly used, by paying attention to noises people make (9 participants) through their activities (indistinct noise, voice, breathing, typing on a computer, pace noise) or by being directly greeted by somebody in the room. Participants also use smells (4 participants), such as perfume, food, hygiene products such skin care and aftershave, touch for airflow perception and clothes, and, interestingly, a sense of a co-presence (5 participants) in form of passive echolocation or a feeling of being observed. Smell can help to estimate the distance of another person.

Behavioral strategies: Participants mention saying hello and then waiting whether somebody answers. The direction of the answer can help to visually search for the other person. An indirect strategy is trying to detect whether the light in a room is turned on since rooms containing people are often illuminated. This is useful when visual capabilities are insufficient to recognize silhouettes but light perception is possible.

4.2 Identifying a known person

Observation strategies: People can be identified visually by the individual shape of their silhouettes (2 participants), and their face if close enough (2 participants). However, eleven of the participants do not use vision at all to identify others. Visual or tactile feedback is used to recognize somebody by their hair (tactile, 2 participants; visually, 2 participants), their clothes (3 participants), the size (silhouette, tactilely when guiding by the height of the leading arm, with

audio when talking). Audition is used to pay attention to the voice (14 participants; sample needs to be long enough to disambiguate identity), to person-specific sounds like laughter (2 participants), way of breathing (3 participants), moving and walking (3 participants). Other person-specific sounds mentioned were the way to playing with a pen, typing on a keyboard, and sounds from jewelry. Jewelry can also help to identify a person by touch. Two participants said they do not dare too much using active tactile exploration, as it may make others uncomfortable. Passive tactile feedback is used to recognize close relatives (e.g., by the way how another person grabs one's arm or how they guide, how their hands feel, musculature). Smell and fragrance can help (10 participants), but do not always give a clear cue to identity. Participants also mentioned using passive echo-location signature for identification (2 participants). Lastly, guide-dog reactions indicate that a person is known; however, this does usually not help to identify the person (2 participants).

Behavioral strategies: Some participants let people approach and present themselves (5 participants). To this end, they define a precise meeting point (3 participants) and make their arrival known (4 participants; by phone, calling out loud, or hitting the white cane on the floor). The context can help to disambiguate persons (colleagues in the workplace, landlord inside a specific house). Other people can draw attention to themselves by calling out or waving at the person with VI.

Miscellaneous: One participant reports never going out without a sighted guide; two others report they are not able to identify people and have to rely on others approaching them. A person that is easy to identify has a recognizable voice (9 participants) or breath, a particular smell or fragrance (3 participants), makes distinctive noises (wrist bands, keys, heels, personal style of walking (3 participants), person in a wheelchair). A person is difficult to recognize if it is difficult to identify their voice (7 participants; person not talking, or too quiet, strong similarity to the voice of another person). Similarly, a person that does not move or is encountered out of their usual context is difficult to recognize, as well as encounters within large public spaces, with too many people (4 participants) and too much noise (3 participants). Lastly, familiarity plays a role—a person that is encountered less often is more difficult to recognize (4 participants) because there were insufficient opportunities to learn distinctive elements for identification (clothes, pace, voices, way of grabbing).

4.3 Identifying that a Person is unknown

Observation strategies: In general, the cues are the same as when identifying a known person; however, without any correspondence in memory: participants use the voice (10 participants), typical words or gestures, way of grabbing the arm or guiding, silhouette and smell.

Behavioral strategies: The fact that participants do not know a person is

usually confirmed during the conversation and by asking questions. Sometimes, the course of the conversation will lead to the realization that the participant wrongly assumed to be talking to a known person.

4.4 Identifying An Activity

Observation strategies and joining an activity: Ongoing activities can be identified visually from gestures and, in case of physical activities, movements of human silhouettes (2 participants). Audition is particularly helpful to pick up noises that humans or materials emit (9 participants; e.g., from a keyboard, specific sounds of a familiar activity such as board games tokens or cooking tools) as well as the density of the noise, the number of different voices and the content of the conversation are considered useful cues to the activity (6 participants), or the general atmosphere of the room (2 participants; e.g., tavern vs. meeting). Smell and touch give information about the material used (cooking, smell of a pen when drawing). Three participants indicated they would not participate in an activity or a conversation if they are not explicitly included by the others or invited to join.

4.5 Approaching a person with VI to start an activity

In the following, we summarize good and bad strategies mentioned by our participants regarding how others can identify themselves and initiate an activity.

Good practices: In many situations, it is helpful if the approaching person states their name (8 participants), their function or job (3 participants), and the context of previous encounters (5 participants). If the person is unknown, they should indicate why they are approaching and, if applicable, the name of the person who send them. When others did not present themselves properly, our participants' strategies are to ask who it is (4 participants), asking possible past memories or if there have been previous encounters (2 participants), and asking questions in general.

Bad practices: Several participants stated that sudden and unannounced physical contact is undesirable, such as catching somebody who is moving with a white cane by the arm to stop them (2 participants), pat them on the shoulder (2 participants), or touching otherwise without prior consent (4 participants). Approaching somebody silently can generate negative effects, like surprise or making the participants uncomfortable because they failed to hear the person's approach. Addressing a person with VI verbally has its pitfalls, as well: others may not speak up long enough for them to be identified (e.g., "Hello" is too short), a call out from too far away can lead to a visual search for the caller but without a successful localization, and calling and speaking loudly at close distance may have startling effects (3 participants).

First impression: What influences the first impression somebody makes on a person with VI? Participants stated that they rely on concrete cues: the hand-shaking, the tone and intonation of the voice (10 participants), if there is a warm

feeling in the voice, the way of speaking and the words used (three participants), the attention the person pays, the face, and the overall behavior (2 participants). One participant stated that he or she can notice if the visual handicap will be an obstacle from the voice and the intonation.

5 Creating Inclusive Virtual Characters

The responses to our questionnaire suggest a number of extensions that can be made to virtual characters to increase their accessibility for persons with VI with current technologies that we describe in the first subsection. The answers also open future perspectives that we describe in the second subsection.

5.1 Adaptation possible with Current Technologies

Visual attributes. Avatars and agents should have distinctive silhouettes, clothes (shapes, color), and faces, to be easily recognized even with low vision. They could have personalized movement patterns, e.g., with idle animation or when talking. A light turned on in a virtual room can indicate that an avatar or agent is present. Some adjustments make VR potentially more accessible, for instance avoiding the necessity of eye to have to adapt to differences in brightness, e.g., when changing virtual environments (e.g., rooms, or from outside to inside).

Audio attributes. Voice is a highly valuable cue to identify people, in particular if the avatar is talking long enough. Distinctive voices, e.g., in terms of pitch, differentiate between characters. With only three voices with referring to gender (feminine, masculine, mixed) a user can differentiate between three characters. Vocabulary, speed, and rhythm add specificity. Other person-specific sounds such as breathing, laughter, way of walking, or wearing different types of shoes (heels, sport shoes, squeaky shoes) improve the chances of identifying the virtual character. Pace can be a cue to decide whether to start an interaction or not (e.g., walking fast may imply little time). Identity and social cues could be conveyed by any sound-generating activity linked to a habit (playing with a pen, tapping finger nails on the table), to objects (keys, chains, sword), or to jewelry (wristbands, ear rings).

Multimodal attributes. The height of a person can be represented visually as well as via spatialized audio feedback. The cloth of virtual characters' garments can be represented visually by a texture as well as by distinctive sounds like rustling and swishing. Avatars in a wheelchair can be a very recognizable by their silhouette and associated sounds.

Airflow. Airflow from a fan could simulated that someone just came closer or walked by.

Interactions. Initiating an interaction is a crucial prerequisite for collaboration. Therefore, it is important that it is transparent to the user how they can join activities. The user should be able to address the entire virtual room, equivalent to "Hello?", using voice or through a controller. The virtual characters, either

agents or avatars, should detect it (broadcast to remote users or audio recognition for agents). The user may want to send a directed call to a particular character (virtual phone call, virtual poke) or semi-directed call (like making noise with the white can). The virtual characters should be able to answer, wave at the person or get closer. They may spontaneously address the user and present themselves or even invite the user to join. The virtual space should ease recognition of characters and activities (e.g., by being a meeting point, providing distinctive activity sounds, or good separation of multiple voices). The place of meeting (noisy, full of people or not) make the recognition difficult or not.

5.2 Future Implementations for Virtual Reality

Smell. Combined with airflow, a fragrance dispenser could indicate that someone is in the room; the distance from this person could be indicated by smell intensity. Many flavors, fragrances and spices exist, liquid or other (vanilla, chocolate, coconut, mint, banana, musk, cedar, cumin, curry, basil, cinnamon), even scent kits, and can be used to give smell feedback.

Tactile feedback for shapes. Tactile feedback helps to recognize jewelry by touch, the face and objects that help identifying the person.

Tactile feedback for materials. Tactile feedback could indicate hair (long, short, no hair) with only one actuator at the hand level for the user. Tactile feedback for cloth can be added to visual and audio-feedback.

Tactile feedback for human contact. Tactile feedback can give a lot of information, to identify a person and to make a first impression. In real life, tactile feedback for recognition is used during handshakes, guiding and even to know the size and the musculature of the person.

Guide dog A virtual guide dog could react differently depending on the time spent with a virtual character or a real person controlling the avatar.

Passive echo-location. Passive echo location is the ability to feel shapes and obstacles from reverberations without emitting a sound purposefully for echoing. If it can be more efficiently modeled in VR, it could be a new means to feel spaces and people.

6 Discussion & Perspectives

We presented the results of an online survey on the strategies used by people with VI for detecting and identifying people as well as on-going activities in real environments. We use the findings to propose solutions for inclusive collaborative VE, with inclusive representations of virtual characters and activities. Many of these solutions can already be implemented, while others will become achievable with future technological developments. Further, our work is applicable to any virtual environment such as for multiplayer games, applications with non-player characters, training environments, telepresence scenarios, or video conferencing. It may require new ways of recording and broadcasting people, to add information beyond the currently predominant video stream. Our future research will

focus on designing and evaluating such virtual characters with people with and without VI.

Acknowledgments

We thank the participants and the institutions that distributed our survey, especially IRSA and Ocons. This work was supported by the European Union's Horizon2020 Program under ERCEA grant no. 683008 AMPLIFY.

References

1. Ahn, H.J., Lee, H.J., Cho, K., Park, S.J.: Utilizing knowledge context in virtual collaborative work. *Decision Support Systems* **39**(4), 563–582 (2005)
2. Connors, E.C., Chrastil, E.R., Sánchez, J., Merabet, L.B.: Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches. *Frontiers in human neuroscience* **8**, 223 (2014)
3. Cornelissen, K.K., McCarty, K., Cornelissen, P.L., Tovée, M.J.: Body size estimation in women with anorexia nervosa and healthy controls using 3d avatars. *Scientific reports* **7**(1), 1–15 (2017)
4. Guerreiro, J., Ahmetovic, D., Kitani, K.M., Asakawa, C.: Virtual Navigation for Blind People: Building Sequential Representations of the Real-World. In: *ASSETS'17*. ACM, Baltimore, USA (2017)
5. Haans, A., IJsselstein, W.: Mediated social touch: a review of current research and future directions. *Virtual Reality* **9**(2-3), 149–159 (2006)
6. Hoppe, M., Rossmly, B., Neumann, D.P., Streuber, S., Schmidt, A., Machulla, T.K.: A human touch: Social touch increases the perceived human-likeness of agents in virtual reality. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–11 (2020)
7. Kreimeier, J., Götzelmann, T.: First steps towards walk-in-place locomotion and haptic feedback in virtual reality for visually impaired. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. p. LBW2214. ACM (2019)
8. Kunz, A., Miesenberger, K., Zeng, L., Weber, G.: Virtual navigation environment for blind and low vision people. In: *International Conference on Computers Helping People with Special Needs*. pp. 114–122. Springer (2018)
9. Lahav, O., Mioduser, D.: Construction of cognitive maps of unknown spaces using a multi-sensory virtual environment for people who are blind. *Computers in Human Behavior* **24**(3), 1139–1155 (may 2008). <https://doi.org/10.1016/j.chb.2007.04.003>, <http://dx.doi.org/10.1016/j.chb.2007.04.003>
10. Mölbert, S.C., Thaler, A., Mohler, B.J., Streuber, S., Romero, J., Black, M.J., Zipfel, S., Karnath, H.O., Giel, K.E.: Assessing body image in anorexia nervosa using biometric self-avatars in virtual reality: Attitudinal components rather than visual body size estimation are distorted. *Psychological medicine* **48**(4), 642–653 (2018)
11. Nakanishi, H., Koizumi, S., Ishida, T.: Virtual cities for real-world crisis management. In: *International Digital Cities Workshop*. pp. 204–216. Springer (2003)

12. Negrón, A.P.P., Vera, R.A.A., de Antonio Jimenez, A.: Collaborative interaction analysis in virtual environments based on verbal and nonverbal interaction. In: 2010 Ninth Mexican International Conference on Artificial Intelligence. pp. 129–133. IEEE (2010)
13. Paiva, P.V., Machado, L.S., Valença, A.M.G., Batista, T.V., Moraes, R.M.: Simcec: A collaborative vr-based simulator for surgical teamwork education. *Computers in Entertainment (CIE)* **16**(2), 1–26 (2018)
14. Pan, X., Hamilton, A.F.d.C.: Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology* **109**(3), 395–417 (2018)
15. Pejša, T., Gleicher, M., Mutlu, B.: Who, me? how virtual agents can shape conversational footing in virtual reality. In: *International Conference on Intelligent Virtual Agents*. pp. 347–359. Springer (2017)
16. Rudinsky, J., Hvanberg, E.T., Helgason, A.A., Petursson, P.B.: Designing soundscapes of virtual environments for crisis management training. In: *Proceedings of the Designing Interactive Systems Conference*. pp. 689–692 (2012)
17. Shachaf, P.: Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study. *Information & Management* **45**(2), 131–142 (2008)
18. Steed, A., Schroeder, R.: Collaboration in immersive and non-immersive virtual environments. In: *Immersed in Media*, pp. 263–282. Springer (2015)
19. Thévin, L., Briant, C., Brock, A.M.: X-road: Virtual reality glasses for orientation and mobility training of people with visual impairments. *ACM Transactions on Accessible Computing (TACCESS)* **13**(2), 1–47 (2020)
20. Thévin, L., Brock, A.: How to move from inclusive systems to collaborative systems: the case of virtual reality for teaching o&m. In: *CHI 2019 Workshop on Hacking Blind Navigation* (2019)
21. Zhao, Y., Bennett, C.L., Benko, H., Cutrell, E., Holz, C., Morris, M.R., Sinclair, M.: Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. p. 116. ACM (2018)
22. Zhao, Y., Cutrell, E., Holz, C., Morris, M.R., Ofek, E., Wilson, A.D.: Seeingvr: A set of tools to make virtual reality more accessible to people with low vision. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–14 (2019)