How Virtual Work Environments Convey Perceptual Cues to Foster Shared Intentionality During Covid-19 for Blind and Partially Sighted Employees

Erin Lee (erin.lee@ocadu.ca)
Perceptual Artifacts Lab, Faculty of Design, OCAD University, 100 McCaul Street, Toronto, ON M5T 1W1

Mahadeo Sukhai (mahadeo.sukhai@cnib.ca)
Accessibility, Research & International Affairs, The Canadian National Institute for the Blind, 1929 Bayview Avenue, Toronto, ON M4G 3E8

Peter W. Coppin (pcoppin@ocadu.ca)
Perceptual Artifacts Lab, Faculty of Design, OCAD University, 100 McCaul Street, Toronto, ON M5T 1W1

Abstract

The Covid-19 pandemic altered workplaces. For those with 'office jobs,' this meant working 'virtually,' or remotely, from home. This transition forced organizations and workplaces to exercise flexibility, adapt workflows and rely on Information and Communication Technologies (ICTs) to work remotely. However, Blind and Partially Sighted Individuals (BPSI) face challenges accessing work digitally and remote communications through ICTs. In response, we report on the results of our longitudinal participatory design study investigating the impact of working and training over a distance for BPSI. What emerged is a conceptual model to assist in understanding how ICT interfaces convey spatial-topological cues for the construction of shared intentionality in virtual work environments. The implications of our model could be significant, as it aids understanding of what is lost and gained when transitioning to virtual work environments. This could inform the development of ICTs with cross-sensory interaction and national accessibility policies for the workplace.

Keywords: Inclusive design; Perception; External Representations; Diagrams; Shared intentionality

Introduction

The Covid-19 pandemic resulted in a transition to working virtually and remotely from home. This transition more greatly impacted Blind and Partially Sighted Individuals (BPSI), who faced challenges with digital accessibility, setting up their home offices, financing assistive devices, remote communications and employer support (Ginley, 2020). In order to investigate the impact of the transition to working and training over a distance for BPSI, a longitudinal participatory design study that consisted of semi-structured interviews, observational research and co-design sessions was conducted. What emerged from this study is a conceptual model to assist in understanding the degree to which Information and Communication Technologies (ICTs) convey **spatial-topological (S-T) cues** for the construction of **shared intentionality** in virtual work environments.

Shared intentionality

Shared intentionality is an area of research and philosophical inquiry since ancient times and across many disciplines, which is the capacity to engage with others in collaborative interactions with joint goals, intentions and cooperative roles for pursuing said goals (Gilbert, 1989; Searle, 1995; Tomasello et al., 2005; Schweikard and Schmid, 2020). Shared intentionality is prevalent in our lives, for example, it is what motivates two or more individuals to raise a child, compete in team sports, play in an orchestra or work in an office. It can be demonstrated in our everyday actions, for instance, stopping at an intersection with another car that is flashing their turn signal. The other car signals a left turn, this light flashing is enough of an implicit statement for you to assume the intention to turn left and therefore understanding the goal of the interaction. For this reason, you wait for them to turn before proceeding (Broz, 2008).

Spatial and topological properties of everyday interactions

In contrast to a real-world situation, spatial and topological properties of everyday interactions, such as gestures, facial expressions, pictures, diagrams or schematics, are conveyed through ICTs in varying ways. For example, consider a scenario where you decide to stop by your manager's office to discuss a task with them. When you arrive, their posture, affect, tone and disarray of objects in the room, allow you to infer that they may be stressed, and it may not be the best time to discuss something with them. These perceptual cues of the physical environment (by means of the light reflecting from objects and people) and communication (implicitly) when perceived inform your understanding of interactions that would be unavailable if you only had access to what is explicitly stated through spoken language or text chat.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Virtual work environments for blind and partially sighted employees

Physical and virtual environments differ in how everyday interactions convey S-T properties. These are cues derived from objects and people that occupy space, and the relations between these. For example, consider a meeting held in the brick-and-mortar workspace; you may enter a room with other people, furniture, a whiteboard and office supplies. These people and objects are in spatial and topological relation so that the individual may apply meaning or act upon when perceived. This aids in understanding that a meeting is occurring. On the other hand, virtual environments are limited in the S-T properties conveyed due to a lack of or limited transmission of implicit communications cues such as gesture, tone, body

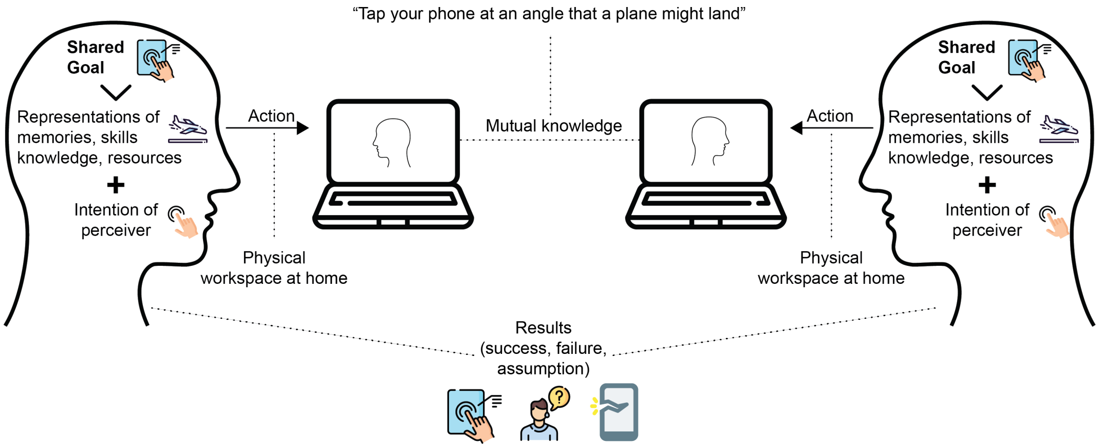

Figure 1: Tomasello et al.'s (2005) shared intentionality adapted for virtual work environments depicting an example of descriptive language use during smartphone training for BPSI

language that people use to infer the intentions of others. These implicit communication cues represent S-T information that can be used to infer the intentions of others (Furlanetto, 2013; Cavallo et al., 2017; Koul et al., 2019; Tversky, 2021). In the absence of these cues, what users can perceive of the people, objects and environment is restricted to what the interface facilitates; in most cases, they are limited to what video conferencing software provides.

ICT-based virtual work environments are composed of external representations we perceive through our senses. Examples include interface displays for videoconferencing that runs on a computer; it is composed of rows of buttons labelled with text and icons. The computer's beeps, alarms, and speech are also external representations, as are the moving images of a video conference call. These external representations play a powerful role in thinking and sensemaking, otherwise individuals would have to rely solely on their internal representations to make sense of the world (Kirsh, 2010; Tversky, 2010). Furthermore, in virtual work environments, the representations you may experience might be both diagrammatic (e.g., charts, 3D objects) and sentential representations (e.g., text chat) (Larkin and Simon, 1987). This study builds on Tomasello et al.'s concept of shared intentionality (2005), however, that model assumes in-person interaction. Therefore, a way to discuss the construction of shared intentionality through ICT-mediated interactions is introduced. For this, Larkin and Simon's concept of diagrammatic and sentential representations was used. They define diagrammatic representations as those that preserve the topological and geometric relations among the components of a data set, while sentential representations are those that preserve temporal or logical sequence within a data set (Larkin and Simon, 1987).

Figure 1 (adapted from Tomasello et al., 2005) demonstrates how shared intentionality emerges in virtual work environments. Each employee's understanding of the shared goal is informed by their internal representations based on previous experiences, the S-T properties, or diagrammatic properties, conveyed through the external representations via ICTs (the laptop) and informed by mutual knowledge (MK) between both parties. This results in the success or failure of the shared goal. Shared intentionality requires individuals to infer the intentions of others. In situations where they cannot draw upon knowledge, based on their experience, as a resource for inferring the individual's intentions, they may form assumptions of what the other person intended.

Individuals with sight loss vary in their abilities to access diagrammatic representations that are presented visually. For example, a BPSI who relies on screen-reader technology has access to text-based tags and labels, text, and the spoken language of the video conference (if they are not deaf or hard of hearing). Without non-visual diagrammatic properties in ICTs, spatial-topological ambiguity can impede the construction of shared intentionality. For example, consider the experience of a blind and partially sighted employee in a meeting where the screen share function is used to demonstrate a chart. Charts contain spatial relations, or diagrammatic properties, between plot points critical to infer value and meaning (Coppin et al., 2015). The conceptual model provides a better understanding of this S-T ambiguity for BPSI, for whom diagrammatic properties are more difficult to perceive and rely on the conceptual specificity of language, or sentential properties.

Methods

Participants

Participants (N= 19; age range = 18-64) of the study included clients and staff from a vision loss organization, and experts from the vision loss community. Participants of the study were both sighted and reported some degree of sight loss. Participants were part of an employment program which offers BPSI job connections, job-readiness workshops, technology training and connections to employers; clients of a national provider of rehabilitation services for people who are blind or who have experienced significant loss of vision; and participants of technology workshops.

Design methodology

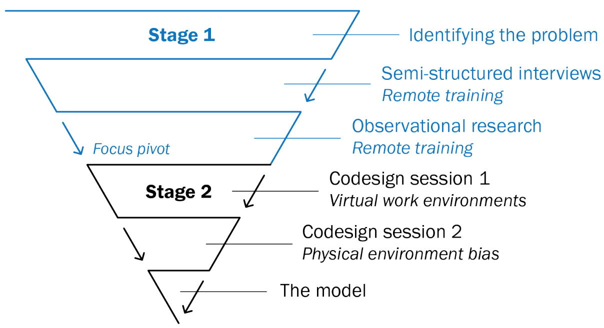

Figure 2: Study design methodology

This study was conducted using longitudinal participatory design. Participatory design is an iterative approach involving exploration, design discovery, prototyping and assessment that allows "[participants] and researchers to critically examine the impacts of redesigns in progress" (Spinuzzi, 2005). This methodology was particularly effective for this project which focused on deriving observations from an evolving and dynamic situation that resulted from the Covid-19 pandemic. Each stage of the research process was designed based on previous steps, as demonstrated in Figure 2, and the methodology was fluid and iterative as the study responded to the needs of the virtual work environment of BPSI. The study involved ten semi-structured interviews, two observational research training sessions and two co-design sessions, leading to the development of our conceptual model.

Results: The model

The model is composed of three dimensions of a Cartesian coordinate system. This includes **Spatial-topological (S-T) synchrony** (Fig. 3, x-axis) which is the degree to which diagrammatic cues, through video, spatial audio, or haptics, offer implicit communication cues, such as gestures, body location, or visual-spatial representations (e.g., diagrams). 0 denotes sentential *descriptions of* 3; 3 denotes diagrammatic representations of 5, for example, as in a video recording; 5 denotes spatial-topological properties of a physical event. **Temporal synchrony** (Fig. 3, y-axis) is the degree to which interactions are synchronous or asynchronous. 0 denotes an asynchronous sentential description of 3, for example as in a letter; 3 denotes a recording of 5; 5 denotes a synchronous event. **Mutual knowledge (MK) creation** (Fig. 3, z-axis) is the degree to which diverse perspectives facilitate the joint construction of knowledge. 0 denotes no MK of working with BPSI (in the context of this study); 3 denotes some MK of working with BPSI; 5 denotes lived experience of sight loss, the extent of MK. All three axes combined, 0,0,0 denotes a scenario where shared intentionality is low, 5,5,5 denotes a scenario where shared intentionality is high. The model will be demonstrated through five case studies collected during this study.

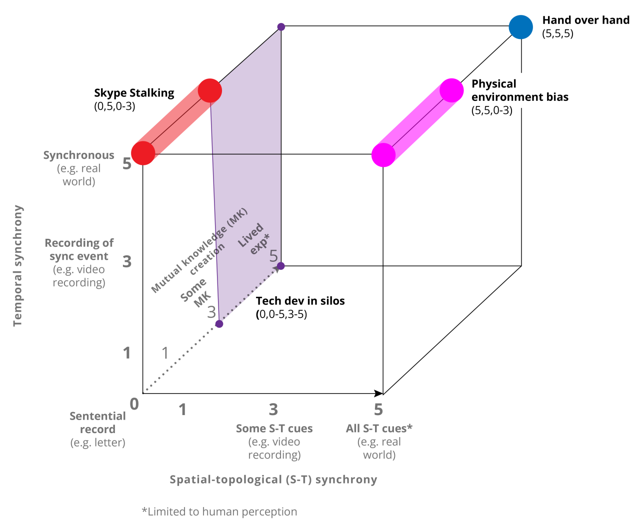

Figure 3: Demonstration of the model through case studies

Case study 1: Skype Stalking. Participants reported a practice in the virtual work environment called "Skype Stalking." This practice is when managers would infer whether employees were "at work" based on whether or not an information display on an ICT, (such as in Skype) displayed the employee's status as "online." The inference that employees were online was then further used to infer whether they were working on their assigned tasks. Skype Stalking is placed at 0,5,0-3 (Fig. 3) in the cartesian coordinate system. This case study depicts a situation in which the construction of shared intentionality is low. It is placed at 0 on the 5-point scale of spatial-topological synchrony because the indicator is a sentential description of a diagrammatic representation and thus provides no S-T cues. In other words, remote work meant that the manager did not observe if the employee was "at work," or how hard they were working through perceptual cues of objects and people in the room. This practice appears to be an attempt to compensate for the lack of cues by relying on internal representations based on memories of previous interactions. Further, Skype Stalking relies on an ICT display that claims to indicate the employee's status synchronously. For this reason, it is placed at 5 on the 5-point scale, however the indicator conveys no information about whether the employee is engaged in work. Participants reported that managers may not possess the knowledge of disability, or had previous conversations with BPSI, and is relying on assumptions of how work is done, thus placing this example at a range between 0 and 3 on the 5-point scale of MK creation, indicating that the manager has no or some experience working with BPSI. Consider how the lack of shared intentionality could impact the working relationship in the long-term. This employee could feel surveilled. This could impact trust amongst team members, which in turn could affect how the team cooperatively works together to complete tasks.

Case study 2: The Hand over hand method. Instructors of vision loss rehabilitation services use "Hand over hand" when teaching life skills, such as cooking or gestures on a smartphone. An instructor places their hands directly on a BPSI's hands to show how to perform actions. Hand over hand falls at 5,5,5 (Fig. 3), a scenario where shared intentionality is high. BPSI use sensory modalities outside of vision to perceive S-T cues of the physical environment. The instructor's hand-over-hand takes this further, as this takes place in the real-world and thus possesses all the S-T and temporal synchrony it affords, placing it at 5. Participants found instructors with lived experience of sight loss relatable as they have the internal representations of learning activities for the first time. There is high MK creation from these representations, placing it at five on the 5-point scale.

Case study 3: Physical environment bias. Participants referred to a "physical environment bias" (Fig. 3) which describes the preconception that working in-person and in the same space is required to work. For this reason, physical presence was the default way of thinking to inform decision-making related to remote work policies and accessibility measures and practices. Physical Environment Bias is denoted at 5,5,0-3 of the cartesian coordinate system. Physical environments, or the real-world, possess all S-T cues limited to human perception putting it at five on the 5-point scale of S-T synchrony and involve synchronous interactions, placing it at five on the 5-point scale of temporal synchrony. This bias may result from defaulting to perceptually rich environments, making the need to rely on internal representations less likely. In addition, virtual work environments result from an evolving situation caused by the Covid-19 pandemic. For this reason, this bias is denoted by 0-3 on the 5-point scale of MK, where the phenomenon may arise from no or some experience working with BPSI in virtual environments.

Case study 4: "Technology developed in silos" (inconsistent interfaces). Participants expressed concerns that "technology is being developed in silos" (Fig. 3). This was their way of referring to the many inconsistent and therefore difficult to learn interface designs in their workplaces (e.g., Zoom versus Microsoft Teams). This inconsistency is denoted via the 3-5 range of the purple plane on the z-axis, indicating the lack of MK amongst ICT developers. In addition, these interfaces afforded very few accessible (non-visual) diagrammatic cues, placing the purple plane at 0 on the x-axis. Interfaces of this case study included both synchronous and asynchronous features, placing the purple plane at the 0-5 range on the y-axis.

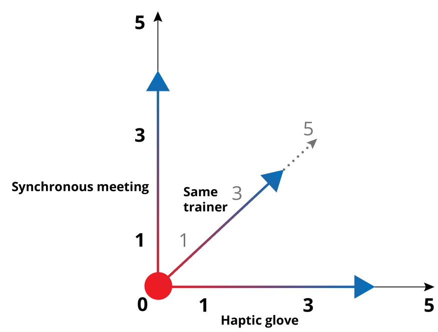

Figure 4: Virtual working solutions

Case study 5: Virtual working solutions (VWS) refer to workarounds in response to insufficient perceptual cues available through ICTs (such as a lack of cross-sensory diagrammatic cues for a BPSI learner using a graphically oriented video conferencing system). **Increasing spatial-topological synchrony:** As a result of the Covid-19 pandemic, life skills training (e.g., cooking) for BPSI was delivered through ICTs rather than in-person. Translating this training to an ICT-mediated interaction (e.g., video conferencing) meant using language to describe embodied actions (e.g., chopping an onion). Sentential representations, or language, can result in conceptual ambiguity (Coppin et al., 2015). In physical environments, life skills instructors indicated using the 'Hand over hand' method to demonstrate embodied actions. In the virtual context, individuals with sight loss are more impacted as they would not have access to S-T cues provided through video conferencing. Consider how a haptic glove can simulate haptic experiences through the sense of touch and is thus a way for S-T cues to be transmitted. A haptic glove could transmit interactions, like Hand-over-hand, gestures or pointing (Perret & Vander Poorten, 2018). In this scenario, it could simulate the action of chopping. This would result in a shift along the x-axis from

0 (Fig. 4), sentential record (e.g., using language to describe an action), to 4, more S-T cues conveyed, not as many as the real-world, however more than what a video conference would provide. **Increasing temporal synchrony:** Email miscommunications can occur in the absence of S-T cues such as tone, body language, facial expressions that convey implicit communication cues that inform an individual's understanding of the interaction. It is also an asynchronous format, reducing the ability to explain one's intentions in real-time. For this reason, email miscommunications are placed at 0 temporal synchrony, sentential record. A synchronous meeting would put this case study at 4. It would convey more S-T properties than a recording of the synchronous event and fewer than what an interaction in the real-world provides (Fig. 4). **Increasing mutual knowledge creation:** Clients shared that consistency in trainers and personalized learning materials were helpful for their learning. For this reason, having an inconsistent trainer for technology sessions is placed at 0, no MK, on the 5-point scale of the z-axis. Over time, clients were able to work with technology trainers to develop a mutually agreed toolkit of representations of technical language to overcome the ambiguity of diagrammatic representations when working together through ICTs. Working with a consistent trainer increases MK to 4 on the 5-point scale of the z-axis (Fig. 4).

Discussion

This section outlines how the model can be used to inform the development of recommendations to improve shared intentionality in the workplace.

Increasing spatial-topological synchrony for better remote communication and collaboration

The brick-and-mortar workspace affords the ability for managers to gain implicit cues through S-T properties that provide information as to whether the employee is at work, or how hard they are working, as well as their focus and intensity toward a task (through posture, gaze, body orientation, etc.). BPSI may vary in their abilities to perceive visual information, however they may have access to other S-T cues, such as spatial audio in the physical space, that allows them to navigate the work environment. Some S-T properties have become conventions for workplace activities. For example, a door slightly ajar often signifies that the employee is at work, but probably should not be disturbed because they are in a meeting, on a phone call, or engaged in a focused task. An open door, as suggested by the phrase "my door is always open," can signify how the employee is at work, is engaged in work-related tasks, but is open to productive interruptions, questions, and relevant work-related conversations from colleagues as they arise. This "door open" practice was a common way to foster shared intentionality prior to Covid-19 induced social distancing. Further, pre-meeting chit chat or what the participants identified as the 'water cooler effect,' was more difficult to replicate or non-existent in the virtual work environment. These are the types of communication opportunities that are possible but unplanned. They can transpire at the water cooler, at the office coffee pot, or while passing colleagues in the hall.

In response to this impact on remote communication, recent work in the design of ICTs has demonstrated how S-T properties of the brick-and-mortar environment have been translated for communicating over a distance through ICTs. Large-scale conferences have used platforms such as Wonder (Wonder, 2021), a virtual space where people can network and talk. This platform displays an overhead map of the "room" and its participants, it allows proximity-based conversations, where participants can get physically closer to attendees they would like to speak to in the "room." However, Wonder relies on visual diagram cues to navigate the "room." Alternatively, for BPSI, audio and tactile cues (such as 3D models) can be used to convey maps to enable navigation through spatial relations (Biggs et al., 2019). Mozilla Hubs is another virtual communication space where participants gather in a room via avatars, it allows for full immersion through virtual reality or a 2D experience using your browser. Mozilla Hubs offers more S-T cues using spatial audio, where conversations closer to you are louder and those further are quieter (Mozilla Hubs, 2021). Consider how these communication platforms afford the possibility of the 'water cooler effect,' colleagues can pass each other in virtual space, engage in unplanned conversations and identify when a person is "in the office." The use of Social VR, getting together in simulated space, could augment the S-T properties currently offered through videoconferencing or phone calls (Li et al., 2020). Other areas of work that build on Social VR are the addition of haptic interactions, for example the ability to give another person a 'high five' or pass documents in virtual space (Fermoselle et al., 2020). The addition of haptic interaction provides a cross-sensory mode of engagement for BPSI, where force feedback provides S-T cues, this could be particularly informative for accessing 3D visualizations in digital spaces. A recent work by Huisman et al. (2021) introduced a wearable system that can translate head gestures of a remote participant into a handheld haptic display of a user during remote communication (Huisman et al., 2021). This advancement would allow the wearer to gain access to implicit communication cues such as nodding and gestures during conversations over a distance and would have a greater impact for BPSI for whom visual cues are not accessible.

Organizational training and education to improve mutual knowledge creation

The participant who described their experiences with Skype Stalking reported how an assumption is formed based on how managers believe the job should be completed. However, these assumptions might fail in many cases. Consider the possibility that the individual could be logging out of Skype to focus on a specific task, this would mean that their indicator light would signify that they are not online, however they are engaged in a focused task. Furthermore, an employee might not be aware that the manager has these expectations, this study found that employees sought clarity around expectations and responsibilities when they transitioned to working virtually and remotely. For example, a participant shared that a coworker kept their video off during video conferencing because they had caregiving duties, which could have been interpreted that the individual wasn't working. In these scenarios, variability in the representations of how work should be completed based on memories of previous work could result in the increased likelihood for assumption forming about what the employee is doing. These assumptions could be alleviated over time as the employee and manager explicitly describe their expectations, but then also iteratively explore practices that, over time, cause them to engage in MK creation, a precursor to shared intentionality. In this study, participants shared the need for training related to accessibility in the context of organizational onboarding of new employees, as well as embedded in ongoing management skill building. This is an opportunity for staff to engage in MK creation and skill development, through shared and agreed upon representations that they can call upon later to counter assumption building and make sense of interactions.

'Hand-over-hand' in virtual spaces through cross-sensory interactions

This study found that life skills instructors faced difficulties when translating the 'hand-over-hand' method to ICT-based virtual environments. Transmitting these actions requires the instructor to heavily rely on how effectively ICTs can convey S-T properties via sentential and diagrammatic representations, or spoken language, which, in most ICTs, is via video streaming. However, if the learner is blind and partially sighted, access to diagrammatic perceptual cues is limited. Figure 1 demonstrates an example from this study by adapting Tomasello et al.'s (2005) depiction of shared intentionality. In this figure, the shared goal between the instructor and blind and partially sighted client is to tap on a smartphone screen. This is a new action for the client and therefore they lack representations to draw on based on memories or skills. To demonstrate this action, the instructor states, "Tap your phone at an angle that a plane might land." This use of metaphorical language is a mutually developed representation that the BPSI has access to. In this way, using language, the instructor and client can develop the MK required to carry out the shared goal. It was found that instructors and staff often compensated by relying on metaphorical language to describe S-T relations that otherwise would have been conveyed via diagrammatic representations to sighted participants. The use of language, when describing gestures and actions or concrete structures can result in ambiguity as a result of the lack of S-T properties that are conveyed (Coppin et al., 2015).

As described in case study 5, haptic interaction provides one way to access S-T properties through the sense of touch. Haptic interaction can provide a tangible way for distance learners in virtual classrooms to access 3D visualizations (Nestor, 2021), for example students in design, engineering and health where 3D models are prevalent. Neto et al. (2020), used tabletop robots to teach students with and without visual impairments how to draw basic shapes used in geometry and handwriting using a hand-held robot that provides haptic feedback (Neto et al., 2020). Further, video and animation are examples of visual information that are not accessible to BPSI, Guinness et al. (2018), developed a Haptic Video Player, that authors audio-haptic content from videos using a mobile robot that can annotate videos and be touched as it moves across the screen (Guinness et al., 2018). This form of haptic interaction provides S-T cues through ICTs for demonstrating actions or providing training, that would otherwise only be available in-person through something like the 'hand-over-hand' demonstration.

Conclusion: Responding to the model

ICTs are critical to how individuals communicate in virtual environments. Our conceptual model allows technology developers, organizations, and IT specialists to understand how ICTs facilitate S-T properties in an accessible way and how they could be enhanced through strategic use of tangible and embodied interactions. Moreover, emerging technologies such as haptic interaction and social VR can simulate experiences, suggesting further ways to transmit S-T properties effectively. Future research and development for accessible ways to integrate these emerging technologies in the workplace and in the education sector would significantly impact the perceptual cues that BPSI accesses through ICTs.

Currently, we are engaged in a project that seeks to develop a cross-sensory authoring tool for 3D visualizations used in e-learning. Spatial skills are critical for understanding relations among objects, people and the environment and inform how we make sense of the world. However, educational content delivered through virtual platforms lacks a way to convey 3D visualizations, particularly in health sciences, industrial design, engineering, and game design. Further, we are engaged in a second project that seeks to develop accessibility standards for accessible and inclusive ICTs that address the needs of individuals with sensory disabilities and differences in sense processing (e.g. neurodiversity, learning disabilities). Our conceptual model provides a theoretical framework from which to inform the development of both interfaces for e-learning and the design of national accessibility policies. For this reason, future work on the model will be completed, specifically the model will continue to be evaluated by collecting empirical evidence of more experiences of BPSI and other disability dimensions of working remotely through the aforementioned future work. Cross-sensory (e.g., audio, haptic) design of ICTs would allow a diversity of humans interact with interfaces that afford S-T synchrony, to improve shared intentionality in contexts where we learn, work and socialize through ICTs.

References

Broz, F. (2008). Planning for human-robot interaction: Representing time and human intention (CMU-RI-TR-08-49) [Doctoral dissertation, Carnegie Mellon University]. Carnegie Mellon University Libraries.

Cavallo, A., Ansuini, C., Capozzi, F., Tversky, B., & Becchio, C. (2017). When Far Becomes Near: Perspective Taking Induces Social Remapping of Spatial Relations. *Psychological Science, 28(1)*, 69–79.

Biggs, B., Coughlan, J. M., & Coppin, P. (2019). Design and Evaluation of an Audio Game-inspired Auditory Map Interface. Proceedings of the ... International Conference on Auditory Display. International Conference on Auditory Display, 2019, 20–27.

Coppin, P.W., Li, A., & Carnevale, M. (2016). Iconic Properties are Lost when Translating Visual Graphics to Text for Accessibility. Cognitive Semiotics. http://openresearch.ocadu.ca/id/eprint/1035/1/Coppin_Iconic_2016_preprint.pdf

Fermoselle, L., Gunkel, S., ter ter Haar, F., Dijkstra-Soudarissanane, S., Toet, A., Niamut, O., & van van der Stap, N. (2020). Let's Get in Touch! Adding Haptics to Social VR. *ACM International Conference on Interactive Media Experiences (IMX '20). Association for Computing Machinery*, USA, 174–179.

Furlanetto, T., Cavallo, A., Manera, V., Tversky, B., & Becchio, C. (2013). Through your eyes: incongruence of gaze and action increases spontaneous perspective taking. *Frontiers in human neuroscience, 7,* 455.

Gilbert, M. (1989) On social facts. Routledge.

Ginley, B. (2020). Working remotely if you are visually impaired. *British Journal of Visual Impairment*, 1-4.

Guinness, D., Muehlbradt, A., Szafir, D., & Kane, S.K. (2018). The Haptic Video Player: Using Mobile Robots to Create Tangible Video Annotations. *In Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces (ISS '18). Association for Computing Machinery*, USA, 203–211.

Huisman, G., Lisini Baldi, T., D'Aurizio, N., & Prattichizzo, D. (2021). Feedback of Head Gestures in Audio-haptic Remote Communication. *2021 International Symposium on Wearable Computers. Association for Computing Machinery*, USA, 135–137.

Kirsh, D. (2010). Thinking with external representations. *AI & Society 25*, 441–454.

Koul, A., Soriano, M., Tversky, B., Becchio, C., & Cavallo, A. (2019). The kinematics that you do not expect: Integrating prior information and kinematics to understand intentions. *Cognition, 182*, 213–219.

Larkin, J., & Simon, H. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science, 11(1)*, 65–100.

Li, J., Vinayagamoorthy, V., Williamson, J., Shamma, D.A., & Cesar, P. (2021). Social VR: A New Medium for Remote Communication and Collaboration. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery*, USA, *81*, 1–6.

Mozilla Hubs. (2021). Meet, share and collaborate together in private 3D virtual spaces. https://hubs.mozilla.com/

Nestor., J.A. (2021). Experiences with Remote Teaching an Embedded Systems Course. *Proceedings of the 2021 on Great Lakes Symposium on VLSI. Association for Computing Machinery*, USA, 437–442.

Neto, I., Johal, W., Couto, M., Nicolau, H., Paiva, A., & Guneysu, A. (2020). Using tabletop robots to promote inclusive classroom experiences. *In Proceedings of the Interaction Design and Children Conference (IDC '20). Association for Computing Machinery*, USA, 281–292.

Perret, J. & Vander Poorten, E. (2018). Touching Virtual Reality: A Review of Haptic Gloves. *ACTUATOR 2018; 16th International Conference on New Actuator*, Germany, 203–211.

Schweikard, D.P., & Schmid, H.B. (2020). Collective Intentionality. In E. N. Zalta (Ed.), Stanford encyclopedia of philosophy (Winter 2020 ed.). Stanford University.

Searle, J. (1983) Intentionality. Cambridge University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *The Behavioral and brain sciences, 28(5)*, 675–735.

Tversky, B. (2010). Visualizing Thought. *Topics in Cognitive Science, 3*, 499-535.

Tversky, B. (2021). Thinking Tools: Gestures Change Thought About Time. *Topics in Cognitive Science, 28(1)*, 69–79.

Wonder. (2021). Wonder – Online events that are fun. https://www.wonder.me/