"I Don't Really Get Involved In That Way": Investigating Blind and Visually Impaired Individuals' Experiences of Joint Attention with Sighted People

Katherine Mary Jones
kj13690@bristol.ac.uk
School of Computer Science
University of Bristol
Bristol, UK

Ute Leonards
ute.leonards@bristol.ac.uk
School of Psychology
University of Bristol
Bristol, UK

Oussama Metatla
o.metatla@bristol.ac.uk
School of Computer Science
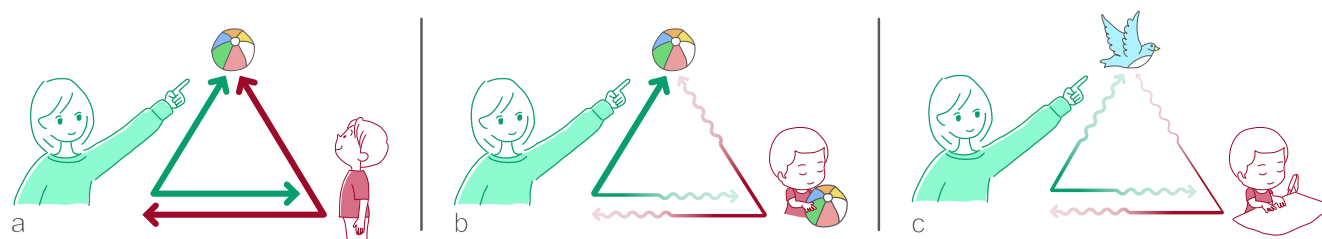University of Bristol
Bristol, UK

Figure 1: (a) Typical joint attention triad, (b) Potential breakdown of joint attention triad with lack of visual feedback for within-reach objects, and (c) Out-of-reach objects. The solid arrows indicate uninterrupted feedback avenues. The wavy lives represent the degradation of this feedback with the lines becoming thinner, and lighter as the disruption increases.

ABSTRACT

Joint attention (JA) is a crucial component of social interaction, relying heavily on visual cues like eye gaze and pointing. This creates barriers for blind and visually impaired people (BVI) to engage in JA with sighted peers. Yet, little research has characterised these barriers or the strategies BVI people employ to overcome them. We interviewed ten BVI adults to understand JA experiences and analysed videos of four BVI children with eight sighted partners engaging in activities conducive to JA. Interviews revealed that lack of JA feedback is perceived as voids that block engagement, exacerbated in group settings, with an emphasis on oneself to fill those voids. Video analysis anchored the absence of the person element within typical JA triads, suggesting a potential for technology to foster alternative dynamics between BVI and sighted people. We argue these findings could inform technology design that supports more inclusive JA interactions.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **HCI theory, concepts and models**.

KEYWORDS

Joint attention, Joint engagement, Blind, Vision impaired, Children, Adults

1 INTRODUCTION

You are in a park. You are playing ball with your sighted friend. "Catch!" *your friend calls, and you catch the ball, then throw it back. This depicts a classic joint attention triad (Fig 1, image a). You are doing an activity that depends on paying attention to each other and the ball. The interaction involves sight, movement, touch and verbal communication. You throw the ball back at your friend without warning, and they move quickly to catch it. Feeling a little competitive, they call,* "Head's up!" *and throw the ball back quickly, but you have become distracted by movement in the corner of your eye. The ball lands on the floor unheeded.* "Look!" *you exclaim, pointing at a bird launching from a tree. It is very colourful, unusual for a city park.* "Oooh pretty" *your friend says, and you start discussing what type of bird it might be. The joint attention triad for the ball broke down once you stopped paying attention to it, but the phenomena picked up again as you both started paying attention to and commenting on the bird. Here, the triad involved sight, physical gestures and verbal communication but no touch.*

A few days later, you are in the park again, playing ball with a different friend who has no sight. This game is achievable because the

ball jingles, so your friend can locate it using sound. The lack of vision may somewhat degrade the joint attention triad, but it persists successfully (Fig 1, image b). You notice the bird again, and momentarily forgetting your friend's lack of vision, you call, "Look!". In the case of the bird, vision and touch are not possible, resulting in a breakdown of the joint attention triad. At that moment, the bird lets out a trill, and your friend says, "It sounds nice!". The joint attention triad is re-established using audition (Fig 1. image c).

Joint Attention (JA) refers to the triangulation of attentional focus between two or more people and an object. JA is vital for social interaction and general development, including conceptualisation of the physical and social world around us through learning from others [15, 17, 22, 88]. The reliance on vision in JA activities hampers its development in blind and visually impaired (BVI) people. [1]

Accordingly, low or no vision from a young age delays the development of JA and similar skills (such as, spatial awareness). The consequences are social isolation and a lack of autonomy compared to sighted peers of the same age [27, 50, 56]. Later life outcomes are also impacted, for example, low employment opportunities (only one in four of BVI children will enter employment as adults, a statistic that has not changed since 1991), limited access to resources, depression and being cut off from the people and environment around them [69].

Despite the importance of JA for social and more general development and inclusion as we move through life, we need more research on how JA presents for BVI people beyond infancy. Indeed, previous research has mainly focused on examining BVI children between six months to five years old [3, 15], covering an age range where many JA tasks rely on physical contact between JA partners [26]. Furthermore, whilst, in the UK, children receive regular reports on their social development skills during nursery school, such rigorous records become rare once the child reaches primary school [53]. This lack of information about JA development in BVI children above the age of five coincides with a decrease in physical contact with others in a child's life [26, 36] and the expansion of social interaction beyond personal reaching space heavily relying on vision. For the BVI child, this decrease in physical contact and, with it, the use of touch provides even fewer opportunities to participate in JA activities [2].

From the above, we can see that JA is essential for social interactions and a person's general development; yet, we need to learn more about JA in BVI people beyond five years of age. Moreover, technology design exists for physical independence in an individual and collaboration in the workplace but does not explicitly support JA in social situations. We must tie these elements together within the technology design and development field: we need to characterise JA mechanisms in children and adults with visual impairment to inform design insights into technology that shifts its thinking from an individual's disability to an inclusive experience for everyone within a social context. This paper begins this process by presenting an analysis of semi-structured interviews to develop a better understanding of BVI people's experiences of JA in social

situations with sighted people. We complement these interviews with an observational analysis of home videos of four BVI children engaging in potentially rich JA activities with their sighted parents, teachers, siblings and friends. This paper makes three contributions:

(1) We characterise JA experiences through discussions with BVI adults.
(2) We provide an analysis of detailed interaction mechanisms that highlight indicators of successful initiation of JA between BVI children and sighted interaction partners via video observations.
(3) We highlight how this new knowledge can help to inform future technology design supporting JA between BVI and sighted people.

2 BACKGROUND AND RELATED WORK

2.1 Joint Attention experiences in sighted people

JA is a situation in which two individuals pay attention to the same object and are both aware of the other's attention to that object [89, 95, 96]. JA can use all our senses, from vision to touch, sound and smell [35]; yet, sighted people rely predominately on vision to extract JA information: a sighted person's eye gaze, head and body orientation indicate what their focus of attention is and whether they are engaged in JA. They will point and gesture effectively and make appropriate facial expressions and head movements [1, 70, 91]. In addition to such visual elements, sighted people will use speech to make the interaction successful [54], and the better their language skills, the better the JA opportunities [45]. Also, through visual feedback, sighted people know immediately whether their attempts to engage in JA are successful. Furthermore, when sighted children cannot see their JA partner, their spatial awareness, visual memory, and ability to follow voices will help them know whether they are being spoken to, what is being referred to or focused on [81].

2.2 Joint Attention experiences in blind and visually impaired people

JA begins in early infancy and has been well-researched [8, 12, 32, 68, 77, 80] in sighted cohorts. Yet, our understanding of JA in BVI individuals remains very rudimentary [3]. Developmental setbacks are a significant problem for BVI children; the less vision a child has, the bigger the setback [25]. BVI children display shorter attention spans with less ability to shift their attention between people and objects than their sighted peers [93]. The consequences of this lack of knowledge mean that external intervention is necessary for BVI people to tackle spatial awareness difficulties [14]. People with BVI may struggle to understand themselves within their physical environment, leading to a lack of goal-directed actions and behavioural responses to sighted people such as their caregivers. This lack of reactivity subsequently frustrates caregivers [6], affecting their behaviour. Such behaviour change in caregivers might not only further increase developmental delay [52], but also isolate BVI children further from engaging with sighted people unfamiliar with how to engage with BVI individuals socially [69].

Indeed, without early support for the caregivers [100], the lack of expected feedback from the BVI infant will affect the behaviour

---

[1]We chose to use a mix of identity-first and people-first language [33, 98]. The rationale is that research has suggested that identity-first language is the preferred approach for many individuals [18, 49]. However, people-first language is still recommended academically and medically [38]. We adopt both approaches in acknowledgement of this ongoing discussion.

of the sighted caregiver and might lead to lower responsiveness, greater negative reinforcement and less time spent by the caregiver looking at their babies' faces. Claims that this might lead to the caregiver missing important cues sent by the infant [78] are supported by an increase in negative vocalisations found in BVI infants.

Despite this, some BVI children have highly developed verbal delivery, possibly because it is the most powerful tool available to them to direct others' attention [15, 101]. Furthermore, the more frequently JA attempts are successful in BVI children, the more frequently BVI children will play [16].

We note that researchers agree that interventions should occur as early as possible to reduce developmental delay and promote any residual vision a child may have [3, 30, 100]. There is, however, an emerging school of thought positing that although BVI infants show slower JA development, we should not compare their development to sighted infants' typical development. A lack of vision logically results in different developmental mechanisms and timelines [3, 65, 82]. Yet, this does not mean that optimal JA development cannot occur without typical visual experiences, provided caregivers', societies' and environmental responses to differences in development are adequate [3]. It thus seems crucial to further our understanding of what leads to successful JA episodes for BVI children beyond infancy and early childhood: the more JA opportunities occur throughout the BVI child's life, the more opportunities for play [16]. This may in turn mean this would benefit their development through social interactions and social inclusion.

2.3 Technology supporting blind and visually impaired people

There is a plethora of assistive technology for BVI people [29, 85]. Design, however, focuses mainly on the functional issues of visual impairment in three main areas [42]: vision substitution such as screen readers to access content [99], navigation aids [58] and medical interventions [11]. Technical assistance for JA in social situations beyond personal space needs attention.

The number of publications has doubled every four years since the mid-1990s, with less than 50 publications per year in the 90s to around 400 per year in 2014 [11]. Yet, user acceptance remains low in adults and even lower in children, with most devices' designs remaining too big, heavy, or complex to use [29].

Historically, assistive technology for BVI people has focused on mobility, navigation, and object recognition [11]. Early technology for BVI people provided primarily verbal information about the environment as seen, for example, in the vOICe, or SmartVision [34, 44, 46]. More recent technology design has started to shift toward technical support for social interaction and inclusion in mainstream society [65]; yet, this is still rare [44] as current opinions in the field are that children primarily need technology for education and adults for routine living tasks [86]. To date, adults' most used assistance tools are the white cane and guide dogs [34]. For children, these include mobility devices, pre-canes, and virtual reality technology [29]. We are not the only ones to suggest that socialisation, independence, and control are the areas in need of development for technological support for BVI people, with collaborative design making a positive difference [20, 51, 64, 87]. To

successfully shift the focus of technology design toward attempting social inclusion, we propose that understanding barriers to and strategies for JA support in children and adult BVI individuals should be a key component of this research direction.

2.4 Technology supporting blind and visually impaired people in social interactions

There has been an exploration into technologies focusing on how to improve collaborative software to consider BVI people so that they can work alongside their colleagues [31, 57] with an emphasis on cultivating awareness of their physical environment [97] and software use [84]. These technologies provide valuable insights regarding the workplace that focus on sharing attention on the object and increasing the ease of collaborative work, including when vision is not the dominant modality of interaction [60]. However, there remains a lack of focus on the joint engagement between the individuals when the focus of attention is less physically immediate or obvious, for example, in everyday activities in classrooms, playgrounds or restaurants.

Researchers have created technology to support BVI children to make sense of their social environment whilst moving away from intimate space [24, 64]. The HoloLens [67] and its upgraded version, the PeopleLens [66], are pioneering examples of social systems explicitly designed to detect others nearby whilst providing feedback signals to bystanders. They aim to increase social sense-making, a person's understanding of and interaction with their social environment. Engaging both BVI and sighted people with the technology is a promising way toward greater acceptance of new technology in BVI people [59, 66]. Yet, we cannot effectively assess this new technology's usefulness because so little is known about JA for BVI people. Acknowledging this gap highlights the urgency for a deeper understanding of the details of JA experiences for BVI individuals. We argue that we can only design technology that supports JA episodes between BVI and sighted peers when we understand how both can meaningfully engage in JA.

3 SCOPE

We build on the literature presented above to address a gap in our knowledge of how to design technologies that foster inclusive JA between BVI and sighted people. As a starting point to help achieve this broader aim, we focus in the present work on two research questions:

- RQ1: How do BVI individuals perceive barriers to engaging in JA with sighted peers throughout their childhood and adult lived experience?
- RQ2: How should we understand and characterise mechanisms and indicators of successful initiation of JA between BVI and sighted people?

4 INTERVIEW STUDY

4.1 Aims

In addressing RQ1, we wanted to develop an understanding of BVI people's experiences of JA in social situations with sighted people. More specifically, to understand perceived barriers to JA and alternative strategies used by BVI people to overcome them

Table 1: Table showing interview participants' demographics.

ID	Participant	Age	Gender	Condition	Level of vision over time
1	Aled	40s	Male	Not specified	Began in childhood, leads to total sight loss
2	Bryn	60s	Male	Not specified	None
3	Ceri	30s	Female	Not specified	Began in childhood, progressive condition
4	Delyth	30s	Female	Congenital Aniridia and Glaucoma	None
5	Emyr	70s	Male	Congenital Nystagmus	Quite a lot
6	Ffion	40s	Female	Congenital Leber' s Amaurosis	None
7	Gaynor	40s	Female	Retinitis igmentosa	Progressive since early adulthood
8	Heulwen	40s	Female	Congenital Leber' s Amaurosis	None
9	Rhian	40s	Female	Not specified	Some useful vision but not significant
10	Wyn	60s	Male	Glaucoma	None

and to gain first-hand accounts from blind and visually impaired individuals.

4.2 Participants

We recruited ten BVI participants (four men and six women) via word of mouth and an advert in a local newsletter. Participants' ages ranged from thirty to seventy years old. Participants' levels of vision varied, from being congenitally blind with no sight perception to later onset and progressive sight loss (see Table 1 for more detail). Pseudonym names have been assigned to maintain the anonymity of participants.

4.3 Methods and procedure

We used semi-structured interviews that lasted between thirty and ninety minutes. Depending on preference, participants joined the interviewer via Teams, Zoom or a phone call. The interviews began with questions surrounding the participant's vision experience and what they understood about JA. We encouraged participants to describe a positive JA experience with a sighted peer. They then explained what made the experience positive. Next, we enquired about a negative experience and questioned what made it a negative one. We also encouraged participants to reflect on what they perceived as the most significant barrier to JA with sighted peers. We asked the participants to reflect on their childhood and describe a positive and negative JA experience from when they were younger. Participants reflected on what made these experiences positive or negative. In a final question, we asked if there was anything that they would like to change when considering situations involving JA and people they knew less well. We encouraged participants to describe as much as possible for each question. We asked probing questions throughout the interview to elicit more information from the participants.

4.4 Data analysis

We began by familiarising ourselves with the transcripts and creating initial codes. These codes quickly pointed to initial themes, which we refined, as reflected in the thematic map (see Figure 2). We coded the audio recordings using the reflexive thematic analysis approach described by Braun and Clarke (2019) [21]. An open inductive thematic analysis was applied to this study because research to date has been biased towards vision, and as pointed out in the

related works section, if exploring joint attention from the point of view of BVI people, then we need to approach the analysis from an alternative starting point. We conducted peer validation throughout the coding process [4], where the authors met regularly to review and clarify coding and grouping decisions. General themes were then determined through iterative discussions.

4.5 Results

We identified 5 themes which we broadly grouped into either positive or negative experiences of JA (see Figure 2). The first theme focused on a lack of feedback, which we labelled as an experience of a void. The second theme centred around the benefits of sighted people using clear and descriptive language. The third theme concentrated on participants preferring one-to-one JA experiences, expressing relief when the JA partner was nearby. The fourth theme addressed participants conveying discomfort in groups, which we interpreted as amplifying the experience of a void. Finally, participants expressed a sense of responsibility for themselves rather than their sighted peers to fill such voids.

4.5.1 The relief of proximity. Participants displayed a preference for having their JA partner close by. Delyth described a greater sense of connectivity: *"I like being close to someone when they're talking to me [...] it gives me a sense of someone, whereas the other side of the room they could be looking at their phone, looking out the window, [...] they might well be engaged [...], but I don't have a sense of their engagement. If I sit near them, I'm like, you're a real person; you're sat there next to me." [P4]* Rhian described the physical closeness as important to utilise her residual vision: *"I can see more when they're closer"* and to connect with the other person: *"You can really sense the presence of somebody else [she explained], and I think the shift of the air [...], that's a sense of connection for me." [P9]* Despite the observation by Rhian, no participants discussed touch as being important to maintaining a JA connection. Regardless of whether someone is in proximity, however, if they have poor language skills, this is felt to harm the social experience of the participants.

4.5.2 Good use of language to describe the world. One of the aspects that participants identified that helped with their JA experiences was when their sighted peers used straightforward language. Aled expressed the need for good use of language through frustration
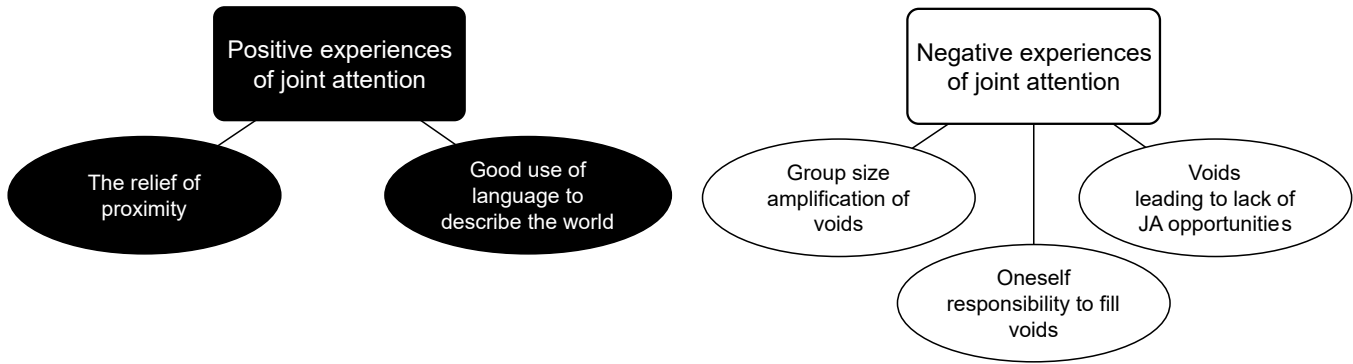
Figure 2: Thematic map of themes grouped by whether they correspond to positive or negative experiences of JA.

by telling the researcher that *"The vague-arity of language can be quite frustrating, [...] it does require a level of mindfulness a lot of people aren't capable of, mindfulness of the world in which they exist, the names of objects, how to estimate distance, how to describe it. [...]. This constant verbal recall that's required in order to communicate things that I think we're just not sort of trained to do in society." [P1]* There was a sense of regret about not being able to share in joint experiences with those who don't have the language skills: *"My cousin has a small child [Ceri conveyed] at the age where she can't necessarily describe things, but she wants to show you something [...] I can't see what it is she's pointing at. So, she'll come up and say, 'look', show me something, and I'm like, 'ohh, that's lovely'. [I have] no idea what it is, and she can't give that communication to say this is what I'm holding up, this what I'm doing." [P3]* This verbal feedback, then, was seen as crucial when describing JA scenarios.

From these first two themes, we summarise that proximity and language are essential for a sense of connection. Indeed, Rhian told the researcher that *"there's definitely that point in which my sight doesn't allow me to keep that connection, that it feels severed quite quickly," [P9]* describing a loss of connection as the person moves away. This observation leads us to the third theme, which explores participants' feelings about groups, a theme that arose unprompted by most of the participants.

4.5.3 Group size amplification of voids. The most prominent theme identified from the data was not liking groups, with six of ten participants mentioning it unprompted. The overarching feeling the participants gave us was profound discomfort when in a group. Participants were worried about turn-taking. Bryn told the researcher: *"In a larger group, I don't know if they're talking to me unless they say my name because I can't see who they're looking at." [P2]* Wyn was concerned about speaking over people or out of turn: *"If you were in a group of people, it could sometimes be difficult to gain their attention because you hadn't got the means of catching someone's eye. So, you just have to speak up, and of course, that runs all the risks of just clashing with somebody else or interrupting someone else if they haven't quite finished." [P10]*

This worry about taking part manifests in Ceri and Heulwen with a feeling of being left out: *"Usually then when the negative things happen, it's if you're in a group and there's a group of people looking at something, because you possibly then can't go [...] sorry,*

can you just stop the conversation to describe to me what's going on?" [P3], "With a group of sighted people, and suddenly they are all [...] commenting on something, exclaiming over something, and they haven't necessarily bothered to fill you in. You have to pay extra attention to keep up." [P8] Then, when you try to take part, you risk hitting a void, which, for Ffion, caused anxiety: *"If I say something, particularly in a group, and there's no answer. It's; what are the looks? Was it, a 'Wow, that was amazing!' Or was it, a 'Oh my God, she's such a, what a doughnut, why did you say that?' [...] It's that uncomfortable feeling of: Have I done something wrong?" [P6]* For Bryn, there were concerns about bullying, which meant he stopped trying to engage in groups: *"I can get intimidated by groups if I don't know people well" [P2].* He talked about feeling bullied; *"If you weren't good at what they were good at, they didn't want to know you, or they would bully the people [...] and I think that's why I didn't do so well in groups." [P2]*

We conclude from these observations that BVI people hesitate to initiate JA in groups because they need appropriate feedback to signal turn-taking or group approval. Once a group has formed, sighted individuals don't necessarily take on the responsibility to provide verbal feedback, so an information void develops. We see in the next section that the participants feel a sense of responsibility to mend this rift in groups and with individuals when in potential JA situations but need the tools to do so.

4.5.4 Oneself responsibility to fill voids. A strong theme through the interviews showed that participants felt a sense of responsibility for failures to communicate in potential JA situations. For some participants, this was self-imposed, for example, by Bryn when discussing miscommunication between himself and the dog trainer. He said: *"It would have been something that I did wrong myself, not [because of bad] advice from the trainer." [P2]* Aled felt the sense of responsibility was imposed on him from *"frequent requests [in work] that I inform people ahead of time what accommodations I'm gonna need." [P1]* For Ffion, she informed the researcher that a sighted person she perceived as a potential friend told her, *"' I can't always hang around with you, I'm not here to look after you'." [P6]* Ffion said that this and other similar experiences made her question how she approached and developed relationships with strangers in potential social situations, choosing instead to avoid initiating them.

Table 2: Table showing blind participants' age, gender, language, level of speech, level of vision, and interaction partner.

Participant	Age	Gender	Speech	Education	Level of vision	Sighted partners
Dafydd	4/5 yrs	Male	Limited	Mainstream	None	Mother and Father
Llian	5 yrs	Female	Yes	Mainstream	None	Teachers, One peer same age
Eira	5 yrs	Female	No	Home	None	Mother and Father
Seren	5 yrs	Female	Yes	Mainstream	Low vision.	Mother and Sister
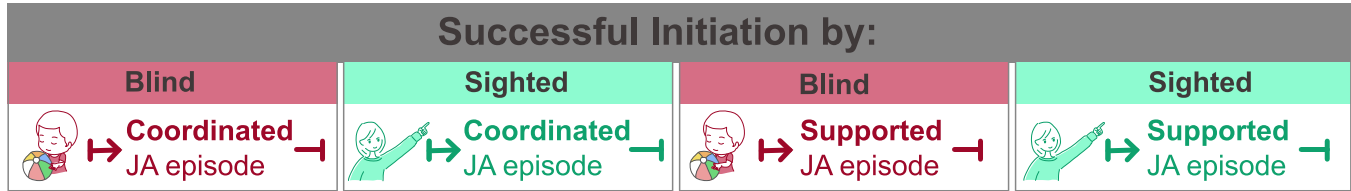
Figure 3: Image showing the variations of initiation of JA and the two main types of JA for both blind and sighted individuals. From this point forward, we focus on initiating JA episodes.

4.5.5 Voids leading to lack of JA opportunities. As mentioned above, the interviewer began the interviews by asking participants for a definition of JA and a positive experience that the individual could remember experiencing. The authors chose to place this theme at the end of the section because what the researcher found appears to encapsulate the overarching theme of the interviews. The observation is that BVI participants have fewer opportunities to initiate JA and, therefore, struggled to explore the concept in a way the questions expected. It is important to note that some participants, for example, Ceri, got the definition right immediately: *"I don't know, like there's a bird in the sky and some person says 'oh, look at that bird', and someone else goes 'okay'." [P3]* However, other participants needed time to think and often gave a correct scenario without the interaction. For example, when discussing building Lego with his grandchild, Emyr described the environment rather than the interaction. *"[I like to keep] the floor clear so I don't trip over something, you know, and all the different pieces all laid out. I got my system to do these kind of things." [P5]*

As the conversations with participants continued, it became clear that positive JA opportunities did not happen. Ffion expressed regret about not having the opportunities. She recalled a childhood memory where *"they had a rabbit that they were going to use in a drama piece that we'd sort of written, and I remember this lady got to hold the rabbit, and I was [dressed up as] a parsnip. I remember feeling the injustice of it, you know, and sort of thinking: Why? Why wasn't I trusted enough to hold the rabbit?"[P6]* Individuals' perceptions of how the lack of events occurred differed. For example, Fion perceived the lack of opportunity as placed on her, whereas Bryn expressed it as a choice: *"I don't tend to really get involved in that way if I don't know the person because the way I feel is, how can I explain to them what I want? They don't know me. They don't know what I want to do". [P2]* The result, however, was the same. Namely, overall, participants struggled to correctly describe and identify positive JA opportunities to engage with their sighted peers and opportunities to initiate and maintain JA in groups was a big problem. These are important concepts to consider when characterising successful

initiations of JA for BVI and sighted people. We derive from this analysis that we need to focus on the initiation of JA episodes instead of the JA episodes themselves. Our discovery from the interview study shows that initiation of JA itself is a problem for BVI people. Without the initiation of JA there is no JA to explore. This is how we continue the focus of our investigation.

4.6 Initiation as a key element of JA

The interview study highlighted that initiation of JA is problematic for BVI people, which in turn has psychosocial implications. In particular, the participants indicated that group settings exacerbated this problem. The themes emphasized that the ability to initiate or respond to attempts to start JA is essential for successful interactions to take place. These observations led to our second study focusing not on the JA episode as a whole, but on the attempts to initiate JA.

5 OBSERVATION STUDY

5.1 Aims

In a second study, we addressed RQ2. We wanted to understand and characterise mechanisms and indicators within social interactions that imply attempts to initiate and maintain JA episodes. Through observations and analysis of interactions of activities conducive of JA between BVI children and sighted adults in home and school settings, and using codes created during the analysis, we sought to distinguish when and how these JA attempts succeeded or failed.

5.2 Participants

We recruited participants via word of mouth, local newsletters and through local schools. Interested parents of BVI children then got in touch to enrol themselves and their children in the study. Child participants were five years old (one participant was four years old in one of the analysed videos). We analysed video clips of four BVI children, three girls and one boy. Of the four participants, three were congenitally blind with no light perception, and one child had

Table 3: Table showing Level 2 codes within the JA activities.

Code Number	Code Name	Code Definition
1	Verbal engagement with JA object or person	Using sounds and words to indicate engagement with the object or person
2	Orientation toward JA partner	Looking at, turning head, body, reaching for, touching, holding partner
3	Orientation toward JA object	Looking at, turning head, body, reaching for, touching, holding object

a condition that causes progressive sight loss (see Table 2 for more details). In addition to the four BVI children, participants in the videos included the children's parents, teachers, siblings, and peers (five adults, three sighted children). Pseudonym names have been assigned to maintain the anonymity of participants.

5.3 Procedures

Parents and teachers sent the researchers video recordings of BVI children participating in potentially JA rich activities with their sighted friends, parents, siblings or teachers. We received videos ranging from 30 seconds to 12 minutes long. We parsed this number down for this initial analysis in the following ways: We selected videos with potential JA interactions. For example, where there were sustained attempts at triangulation between the BVI child, their sighted companion(s) and the object or objects in question. We removed video clips of what appeared to be shared attention but not JA. For example, where the sighted person observed rather than engaged whilst the blind child completed an activity, such as playing the piano or singing. The resulting videos comprised a total of 14 videos between 10.9 seconds and 11 minutes 47 seconds long, totalling 56 minutes long. Of those videos, 25 minutes of data was coded for the analysis of the initiation of JA activities.

5.4 Data analysis

We used a combination of inductive and deductive coding to analyse the video data. In addition to Alfaro's (2018) JA categories [3], we used insights from Section 4.5 to guide more deductive coding.

We produce three levels of codes:

(1) **Level 1 - Low-level modalities of interaction**: The primary coder initially employed an inductive coding technique by recording movements of the body, head, and hands made by each participant towards either the JA object(s), the JA partner(s) or something (person or object) not involved in the triangle of activity. This level also included verbal contributions to the activity. For example, whether the participant was asking, answering, prompting, or encouraging the JA partner. A total of 577 codes resulted from this process.

(2) **Level 2 - Higher-level JA locus**: In the second level, we use inductive coding to group the level one codes into three categories containing more generalized behaviours. The three codes focused on speech, orientation towards the JA partner and orientation towards the JA object. The speech category refers to all instances of speech that addressed the JA object. For example, questions, answers, prompts and exclamations. The person category refers to orientation with hands, body, head, and gaze towards the JA partner. Finally, the object category refers to orientation with hands, body, head and gaze towards the JA object. (see Table 3).

(3) **Level 3 - Types of JA initiated**: We used deductive coding, building on and extending Alfaro et al's protocol developed for infants with visual impairments [3], to identify the type of JA episode initiated by the participants in a given video segment. Here we focused on extracting whether the initiation attempts led to two types: coordinated (see 5.4.1) or supported (see 5.4.1) JA. To this coding protocol, we added codes that specified who initiated the successful JA episodes. We also identified unsuccessful attempts to initiate JA episodes (see Section 5.4.1).

5.4.1 Types of Joint Attention Initiated. Our analyses captured how both the BVI children and their sighted peers may successfully and unsuccessfully initiate JA episodes across two types of JA; coordinated and supported, and identified who initiated each episode (see Figure 3). We then present detailed analyses of the mechanisms and indicators involved in two types of the JA initiations, namely, "Successful initiation by BVI child leading to coordinated JA" and "Successful initiation by sighted peer leading to supported JA" below:

- **Coordinated JA:** Both the blind child and their sighted partner are actively involved with the same object or event. The blind child repeatedly acknowledges the sighted partner's participation. The child acknowledges the partner's involvement. The sighted partner's level of activity directly on the object may be minimal because the blind child is more active in balancing attention between the shared object and the social exchange.

- **Supported JA:** The other person's involvement influences the child's activity with the object, but the child does not acknowledge this involvement. The partner's involvement with the object must influence the child's experience of the object or event. The blind child and sighted partner are actively involved with the same object or event, but the child does not acknowledge the partner's participation. Even if, during the activity, the child focuses primarily on the object and not the caregiver, the caregiver's participation influences the child's experience of the event.

- **Successful initiation by BVI child leading to coordinated JA:** The child captures the other's attention via an

Table 4: Table showing Fleiss Kappa inter reliability coding for the three Level 2 codes.

Code number	Code name	Kappa	Percent	Confidence intervals	Strength of agreement
1	Speech	0.729	72.9	0.710-0.749	Good
2	Person	0.480	48	0.460-0.499	Moderate
3	Object	0.719	71.9	0.699-0.738	Good

Table 5: Table summarising total count and the total percentage of occurrences of level 2 codes within the level 3 initiating JA episodes for all analysed videos.

	Speaking about object	Oriented toward person	Oriented toward object
Blind child	19/49	9/49	38/49
Main sighted peer	45/49	31/49	25/49
Total	64/98	40/98	63/98

object or event. This can be through auditory, physical, verbal, vocal, or visual signals. This leads to a successful episode of coordinated JA.

- **Successful initiation by BVI child leading to supported JA:** The child captures the other's attention via an object or event. This can be through auditory, physical, verbal, vocal, or visual signals. This leads to a successful episode of supported JA.
- **Successful initiation by sighted peer leading to coordinated JA:** The sighted partner attempts to capture the child's engagement via an object or event. This can be through auditory, physical, verbal, or vocal signals. This leads to a successful episode of coordinated JA.
- **Successful initiation by sighted peer leading to supported JA:** The sighted partner attempts to capture the child's engagement with an object or event. This can be through auditory, physical, verbal, or vocal signals. This leads to a successful episode of supported JA.
- **Unsuccessful attempt to initiate JA by BVI child:** The child attempts to capture the sighted partner's engagement with an object or event. This can be through auditory, physical, verbal, or vocal signals. The attempt fails if the sighted partner does not acknowledge or engage with the child or the intended shared object or event. The attempt is also coded as a failure if the sighted partner acknowledges the child's attempt to engage but is immediately distracted by their own (already started) train of thought or action or another external event.
- **Unsuccessful attempt to initiate JA by sighted peer:** The sighted partner attempts to capture the child's engagement with an object or event. This can be through auditory, physical, verbal, or vocal signals. The attempt fails if the child does not acknowledge or engage with the sighted partner or the intended shared object or event. The attempt is also coded as a failure if the child acknowledges the sighted partner's attempt to engage but is immediately distracted by their own (already started) train of thought or action or another external event.

We coded videos using NVIVO software. Once coded, the data were extracted from NVIVO in preparation for quantitative analysis. A custom-made Python code was created to continue the analysis in SPSS. Once the data were in SPSS, we drew the data together to continue analysis beyond this paper's scope. Additionally, SPSS was used to analyse the results of an inter-coder reliability test.

5.4.2 Inter-coder Reliability Test. To test the validity of our coding scheme, we recruited three coders to code samples of the videos for the level two codes. The rationale for the inter-coder reliability test to occur at this level of the analysis is as follows: The level one codes were too low-level (for example, lifting head, turning head, lifting hand and so on). The level three codes were considered to potentially be too abstract to interpret for non-experts of the topic (coordinated JA, supported JA, initiation of supported JA by BVI child, and so on). Therefore, coders were asked to code for the mid-level codes (orientation towards partner, orientation towards object and speech). See Table 3 for more information on these codes. The statistical test compared the coding results of the three coders and the main researcher. Although unconventional, this approach was chosen to determine whether the researcher's deep involvement with the coding results in significantly different results to the other coders. Using Fleiss-Kappa, we obtained overall agreement of 0.687. Fleiss-Kappa was used over Cohen's-kappa because we compared the results of more than two coders. Orientation towards the JA partner showed moderate agreement, object and speech good agreement. (See Table 4).

5.5 Results
We identified the folllowing number of episodes:

- 4 episodes of "Successful initiation by BVI child leading to coordinated JA"
- 4 episodes of "Successful initiation by BVI child leading to supported JA"
- 11 episodes of "Successful initiation by sighted peer leading to coordinated JA"
- 23 episodes of "Successful initiation by sighted peer leading to supported JA"

Time	Llian (BVI child)	Megan (Sighted adult)	
0.27	[Object] touches JA object (dog toy) with her face [Object] squeezes JA object	[Person] Orients head and shoulders towards Llian [Object] Touches JA object with right hand	
0.28	[Speech] *"she's so cute I don't want to put her down"* [Object] Strokes JA object	[Person] Orients head and shoulders towards Llian [Object] Touches JA object with right hand	
0.29	[Object] Tucks JA object into lap	[Speech] *"She's so cute isn't she? Does she remind you of your friend's dog a little bit?"* [Person] Orients head and shoulders towards Llian [Object] Touches JA object with right hand	

Figure 4: Successful initiation by BVI child leading to coordinated JA

- 5 episodes of "Unsuccessful attempt to initiate JA by BVI child"
- 2 episodes of "Unsuccessful attempt to initiate JA by sighted peer"

Table 3 shows that overall, everyone speaks about the object most frequently, orients towards the object the second most frequently and orients towards the person the least. The BVI children orient toward the object the most, and the sighted peers speak about the object the most frequently. To determine if there is a significant difference in the frequency that the BVI children and sighted peers performed in each code, we conducted a CHI Square test of independence. The results of Cramer's V indicate that vision does affect how often you speak ($\Phi_c = 0.571, p < 0.001$), how often you orient towards the person ($\Phi_c = 0.424, p < 0.001$), and how often you orient towards the object ($\Phi_c = 0.255, p = 0.002$). We conclude from this initial data analysis that both the BVI and sighted participants oriented less towards each other than they spoke or oriented toward the JA object. This, in turn, indicates an alternative dynamic to that found in a typical JA triad [89].

Below, we illustrate representative snippets from the video data for two examples of the level three codes, each with a transcript and description of the level 2 codes (Speech, Person, Object) as detailed in Table 3. We then reflect on all of the videos relevant to the JA triads and the interactive tasks they were involved in.

5.6 Analysis of Successful Cases of Joint Attention Initiation

5.6.1 ***Vignette 1: Successful initiation by BVI child leading to coordinated JA.*** In this video section (see Figure 4), Llian sits beside her teaching assistant, Megan and opposite her SEN coordinator, Gwawr, behind the camera. Megan has just finished brushing Llian's hands to wake them up, ready to read some braille. Megan uses a new technique to teach Llian braille by using a fidget poppit toy cut into a block of six to replicate the braille set-up (JA object 2). As Megan is organising the table to prepare for the task, Llian plays with a toy dog (JA object 1). Llian is aware that the task is about to start. Llian squeezes the dog, which whines and tells Megan that she doesn't want to put the dog down. Megan responds. They then spend the next eight seconds focusing on the dog instead of learning braille.

When considering the JA locus, we can observe that Llian and Megan orient themselves towards the JA object. Llian orients herself with her hands and Megan with her head and gaze. They both speak

about the object. However, only Megan orients herself towards Llian. When exploring further within the low-level modalities of interaction, Llian stays facing forward with her head lowered most of the time. She lifts her head just before she speaks and then lowers it towards the object again. On the other hand, Megan turns her head to face Llian and the object and remains in that position.

We observe via the image and image descriptions that the child, Llian, has captured the adult, Megan's, attention by speaking to her about the JA object. Llian's physical and verbal reference to the object has prompted Megan to turn her attention from the poppit to the dog and respond verbally to Llian whilst physically orienting her head towards Llian and the dog. This event is coordinated over supported JA because Llian has acknowledged Megan's role in the exchange about the object, and Megan has returned that acknowledgement.

Across all analysed videos, we identified only four incidents of successful initiation by the BVI child leading to coordinated JA. Of those four, three were made by Llian and one by Seren. Llian and Seren both have speech, whereas Dafydd and Eira do not. Llian and Seren used speech to initiate the JA episode in these cases. When considering the locus of the JA as speech, orientation towards JA partner and orientation toward JA object, we found that Llian and Seren spoke about the event to their partner in all instances. Seren did not orient toward an object because she explored a non-physical concept (the Little Red Riding Hood story).

Finally, neither child oriented toward the partner at all, except that on one occasion, Llian oriented toward Gwawr to pass her an object (a coin given as change in a maths game). In comparison, sighted

partners used verbal responses every time. In three of the four episodes, they oriented toward their BVI partner, and in three of the four, they oriented toward the JA object. We can conclude that in the cases observed, the BVI child orients toward their partner much less than they speak about or orient toward the object.

5.6.2 ***Vignette 2: Successful initiation by sighted peer leading to supported JA.*** In this video (see Figure 5), Dafydd plays a game with his sighted peer, Cerys. Cerys is sitting cross-legged on a sofa with a box on her lap. It is a big box with a large hole cut out of the back and two smaller holes cut out of the front. Daffydd is on the sofa next to Cerys. Throughout the video, he stands, sits, jumps up and down and lies down on the sofa. The game involves Cerys selecting two soft toys from a box on the coffee table in front of the sofa. Dafydd chooses a toy based on the sound. Cerys then puts his

Time	Dafydd (BVI child)	Cerys (Sighted adult)	
1.36	[Object] Squeezes the toy. The toy chirps	[Speech] *"Now, can Dafydd put the duck on the box please"* [Person] Looking at Dafydd	
1.37	[Object] Begins the action of returning the bird to Cerys	[Person] Looking at Dafydd	
1.38	[Object] Seeks the box with the bird in his hand. Remains in a prone position	[Speech] *"So, we can get another bird"* [Person] Looking at Dafydd	
1.39	[Object] Attempts to put the bird back in through the hole in the box	[Speech] *"On the box, on the box"* [Person] Looking at Dafydd	
1.40	[Object] Finds the top of the box	[Speech] *"Thank you"* [Person] Looking at Dafydd	
1.41	[Object] Rests the bird on top of the box, keeping it in his hand	[Person] Cerys is looking at Dafydd [Object] Lifts her hand towards the bird	
1.42	[Object] Feels Cerys' hand brush across his before she takes the bird	[Person] Cerys is looking at Dafydd [Object] Touches Dafydd's hand before she takes the bird	

Figure 5: Successful initiation by sighted peer leading to supported JA

choice in the box through the back, and Dafydd finds the holes in the box and pulls the toy through. Dafydd then has some time with the toy, a bird, listening to it chirp and Cerys sing. Cerys then asks Dafydd to return the bird to the box so that he can choose a new one. In this section of the video, Cerys has just finished singing and has asked Dafydd to return the bird to the box so that they can choose a new one; Dafydd obliges.

When considering the locus of the JA (level 2 codes), we observe that Cerys orients herself towards her JA partner whilst asking him about the object. Dafydd pauses in his movements as Cerys speaks. He responds by engaging in the JA episode through a physical response to her request. When exploring a bit deeper, we see that within the modalities of interaction (level 1 codes), the only indication Dafydd gives that he is engaged with Cerys and the toy is that he responds to her request by moving his arm.

As we see via the image and image descriptions, the adult, Cerys, has captured the child, Dafydd's attention by speaking to him about the JA object. Cerys requested an action from Dafydd, and Dafydd has responded to this request. The successful initiation of JA into supported JA is seen here because although Dafydd does not acknowledge Cerys verbally or through body orientation, he physically responds to her words by engaging with the object.

5.6.3 Coordinated JA. Our analysis indicates that BVI children relied heavily on speech and touch to initiate the coordinated JA triad. Their sighted companion compensated for the child's lack of vision by increasing touch and speech frequency. The sighted person did not appear to mute their physical cues. However, they adapted them by turning a pointing gesture into a tap, a nod or a smile into an affirmative statement even though other than the tap none of these are accessible to the blind child. There were times when the gestures made by the sighted person were not reinforced with speech, which led to confusion on the part of the child and a temporary breakdown in the JA triad.

For coordinated JA, Llian and Seren used speech and touch to indicate the object they wanted to draw attention to. We observed

that coordinated JA only occurred with the two children who implemented speech. This observation suggests that for a BVI child to experience a more coordinated interaction, they need to demonstrate clearly to their partner what they want to convey and what information they want to receive.

5.6.4 Supported JA. The children who did not communicate verbally did not seem able to successfully initiate the JA triad like those with speech did. Instead, they relied on verbal prompts and physical contact from their sighted partner. The incidents of successful initiation by the BVI child leading to supported JA were all made by Dafydd. Dafydd spoke very little, responding to questions but not initiating conversations. Instead, he appeared to use his hands to indicate what he wanted his partner to attend to. Dafydd would choose an object through touch or verbal response to a question, and the sighted person had to notice this preference and respond. Eira did not communicate verbally or appear to use affirmative gestures indicating her JA intent. She responded to verbal and physical prompts and successfully supported JA experiences, but if her partner did not sustain any input, she quickly returned to solitary play. This observation suggests that for the BVI child who does not verbally communicate, successful JA relies on the sighted partner. It appears that the lack of speech paired with no vision degrades the JA triad further.

5.6.5 Unsuccessful JA. The unsuccessful attempts at JA made by the BVI child appeared to fail because the sighted person was paying attention to someone or something else at that time and did not respond to the verbal or physical cue that the child gave. We identified two incidents where the sighted peer attempted to initiate a JA episode but failed. We observed these failures when the BVI child physically removed their hand from the object (Daffydd) or started talking about something else (Llian). This finding implies that the sighted partner was more likely to initiate a successful JA event than the BVI child.

We observed that the dominant initiation of JA episodes was by the sighted partner, who tended to begin interactions with a verbal

prompt, instruction, or question to the BVI child. Speech from all parties, except by Dafydd and Eira, dominated these episodes. The sighted partners oriented towards their JA partner, but the BVI children did not. We observed that all parties oriented towards the JA object more than they oriented towards each other.

6 DISCUSSION

In this paper, we aimed to further our understanding of the mechanisms used to successfully initiate joint attention (JA) episodes between people with low to no vision (BVI people) and their sighted peers. Such understanding is key to filling the gap in our knowledge of designing technologies that foster inclusive JA opportunities between BVI and sighted people. More successful JA episodes are likely to lead to better social interaction and inclusion felt by BVI and sighted people. Indeed, the literature tells us that JA in BVI people experiences delays during early development [13] and that this delay affects later life outcomes for many of these individuals [69]; yet, comparably, we know little about day-to-day JA between BVI and sighted people. Here, we interviewed BVI adults and analysed videos of BVI children in potentially JA-rich environments with sighted peers. In this section, we discuss our findings in the context of whether what we have learned can be applied when considering the design of assistive technology.

In RQ1, we asked: *How do BVI individuals perceive barriers to engaging in JA with sighted peers throughout their childhood and adult experiences?* We found that the main barriers to JA with sighted people, as perceived by our BVI participants, are those where JA occurs in a group setting and where proximity is lost. Our participants found that they could not comfortably initiate JA due to a breakdown of a feedback loop. They expressed a sense of void in attempting to start these interactions, with the responsibility on them to establish and maintain that connection.

In RQ2, we asked: *How should we understand and characterise mechanisms and indicators of successful initiation of JA between BVI and sighted people?* Our findings expand on previous research that identified categories of JA by exploring the initiation of these categories by the BVI child and their sighted JA partner in more detail. We broke down these mechanisms into three main categories: whether the person orients towards their JA partner, the JA object, or indicates engagement through speech. We determine that for the successful initiation of JA, both sighted and BVI people rely on speech. For the BVI child, when speech is not an accessible tool, and there is a lack of vision, the success of initiating and maintaining a coordinated JA experience is greatly reduced. This implies that useful and appropriate intervention at this point would give the BVI child more opportunities for autonomy in these circumstances.

6.1 Perceived barriers to JA and how these barriers help with understanding how to code for JA analysis

Our findings from the interviews highlighted a void, or an absence, of opportunities, signals, and responses on the part of the sighted and the BVI person for social contexts that would lead to successful JA episodes between two sighted people. This void is likely a consequence of the lack of sensory information required both to initiate JA and to respond to JA attempts by the interaction partner.

Similar to the difficulties of the sighted interaction partner in the videos, this void may have also made it difficult for us as coders to identify successful and unsuccessful attempts to initiate JA in the BVI children's videos.

This highlights the importance of collaborating very closely with our BVI companions [90] to learn how they experience the world around them and engage in JA episodes when interacting with other BVI people. For example, one participant expressed how wonderful the world can be without sight and expressed a feeling of regret that her sighted friends would not sometimes close their eyes and experience the sensory richness and delights she and her BVI friends did when in different environments. To her, it often feels as if her sighted friends chose to over-describe the visual world she was missing, but this visual information masked the input of all the other senses. In other words, her multi-sensorial world seemed to get reduced to a mono-sensorial one in her sighted friends for a sense she was missing. From such observations, it becomes clear that sighted and BVI people need to work together to find alternative ways to approach JA so that positive social experiences happen more by taking other senses more strongly into account. In other words, a system is needed that allows cross-modal coordination between interaction partners.

Our interviews focused on BVI people's experiences with sighted people because we live more and more in a sighted world [23]. BVI people enter mainstream classrooms [65] and attend mainstream colleges and universities [10]; they work in offices filled with sighted people who may have yet to work with a BVI person [94]. Yet, we found that the overarching experience of BVI people was that sighted people showed little motivation to slow down in groups or try to experience their world without sight [43].

Previous research has brought valuable understanding to our knowledge of BVI children's early development and how JA experiences develop in infancy. Our work expands this line of research by showing how adults interpret their JA experiences so that we can confidently identify what BVI children and their sighted peers do to initiate JA episodes successfully, what happens when JA episodes break down and what to expect between JA partners with the introduction of physical distance. One of the key points drawn from the interviews is that participants struggle with initiating joint attention for several reasons. This difficulty led us to focus on this aspect of joint attention in the video analysis to understand how BVI and sighted people successfully initiate JA and what happens when they fail so that future research can consider ways to help.

6.2 Strategies for successful and unsuccessful JA

From discussions in the interviews and observations in the videos, it emerges that proximity is a key ingredient for successful JA. As a result, proximity featured in two major themes that ran through the interviews: a) BVI people prefer sitting next to their JA partner, and b) groups and physical distance between interaction partners beyond personal space cause JA opportunities to disappear. Coincidentally, or perhaps not, the videos of the BVI children successfully engaging in JA all involved them being within touching distance of their sighted peers. The only child that sat further away was the child who had residual vision. This proximity aspect harkens us back to Figure 1, which shows that the JA triad is complete when

vision is present, and both people can coordinate their senses [7], especially when objects are within reach. The triad can be maintained with little or no vision because although the two people no longer receive all the visual information, they can still communicate through speech and touch, especially when the object of JA is within reach. However, the information line between the JA partners is already disrupted at this point. This disruption comes from the sighted person not being able to "read" the BVI child's JA cues (e.g. eye gaze as a feedback mechanism is missing). Furthermore, the BVI child cannot "read" the sighted person's physical cues (such as eye gaze, facial expressions, or posture), other than through touch, or sound. Moreover, without vision, information about the JA object is also reduced in the BVI child. The modal communication that has failed becomes replaced with cross-modal communication. There is preliminary evidence for crossmodal communication between interaction partners that audio-visual stimuli can improve performance in a tapping task [63] and its potential to support coordination between BVI and sighted peers in the workplace [61, 62] and in schools [65]. This highlights the need for a more in-depth exploration of crossmodal communication between sighted and BVI people. Sighted peers attempted to mend potential breakdowns when initiating JA by tapping, touching, and physically and verbally guiding their BVI companion. This compensatory behaviour is more successful when coordinated than supported by JA. A reason for this could be that for coordinated JA, the sighted person receives recognisable feedback from the BVI person of engagement.

One of the most apparent signs of coordinated JA being successful was the BVI child's verbal response to the sighted peer's initiation of JA. In supported JA, the sighted peer had to rely only on physical cues. Furthermore, when a JA triad involves distance and out-of-reach objects, and vision is absent, the JA triangle degrades the most. The sighted person only has language as a tool to communicate their intent, so their vocabulary must be useful. As such, the BVI person relies heavily on the sighted person to provide that verbal feedback. Furthermore, the BVI person may give no cues that they would like to engage in JA about a particular event or object because they have no way of knowing that the event is occurring in the first place. This aspect was highlighted in the interviews by examples in which a flurry of activity was happening around a BVI person, nothing of which could be interpreted by the BVI person. Still, as they didn't know what was happening, they could not react and thus might have appeared uninterested. As a consequence, the sighted people around may feel uncertain and hesitant about how to behave because of the lack of signals sent by the BVI person, or indeed their lack of acknowledgement of the situation unfolding around them [48, 55]. Combined, this is a perfect storm that an intervention could ease, that we could design for if we reflect on the lessons gleaned here. Research shows that joint attention relies on vision, vocabulary and touch to succeed. Without vision, JA relies on sound (predominately carefully chosen language) and touch [26, 36]. We go beyond this by characterising the void produced by the lack of vision and touch for JA both from a signalling and a feedback perspective. We also highlight that public settings such as schools and office environments particularly pronounce this void. A rich vocabulary becomes crucial here, and assistive technology might help reduce this void.

6.3 Design Insights

Our findings enabled us to propose suggestions that could be useful when designing for technology that aims to support initiation in a disrupted JA triad. How such assistive technology might look in order to reduce the JA void will be the next step for designers to explore. From both the literature review and the analyses presented in this paper, it becomes clear that any attempt to create such technology should consider the following points:

a) Manage the tension between supporting orientation feedback and avoiding masking and sensory overload; In the videos, we observed that the BVI child seemed almost exclusively oriented towards the JA object and not toward their JA partner as one would see in sighted people [41]. It is important, however, that technology does not interrupt crossmodal communication strategies that the BVI person has already adopted to help themselves navigate their social environment and tries to replace them with orienting behaviours toward the sighted partner, especially as so much natural behaviour is suppressed or masked already [5, 28, 72]. Therefore, technology that supports JA would need to consider that whilst orienting toward the potential JA partner is a valuable tool for the initiation of JA, it needs to avoid disrupting any functional skills, residual vision, or unnecessarily overwhelming other senses, such as hearing [71, 73].

b) Consider reach in the absence of touch and vision; The video participants were all sitting close to each other, and touch was a frequent modality. Interview participants reported struggling when their JA partner moved further away but did not talk about touch as a medium. As we age, touch works in some circumstances but not others and varies across cultures and individuals [17, 92]. Furthermore, in some cultures, the social circle expands as the child ages and touch frequency is reduced [26, 36]. However, rather than adding more audio signals that might overload working memory, consider providing information in a different mode [83] to enable the individual to receive more information about their social environment without overloading their working memory [73]. Designers, therefore, may consider conveying information about proximity or the presence of the JA partner to support initiation of JA episodes [47], whilst minimising cognitive load, for example, using a wider range of sensory modalities [71].

c) Consider form factors and signalling, whether technology should be wearable, embedded in the environment, or both; The interview findings informed us that navigating JA in groups was difficult for the BVI person and that turn-taking and initiation of JA were difficult due to the lack of social cues, such as picking up on eye gaze, facial expressions or gestures. When designing for JA in groups, embedding sensors in the environment that can read potential social cues around presence or proximity can be used to send a signal to a wearable device that feeds back information to the wearer [19, 76]. This could support the filtration of information and aid calibration [9, 24, 37, 39]. The device in the environment could relay cues that the BVI person cannot pick up on because they are happening at an inaccessible pace [64]. It could also pick up environmental information such as who is there, where they are, and whom they are oriented towards [40, 60]. This technology, positioned around a classroom or meeting room, could be a useful and practical solution for both the BVI person and the sighted potential JA partners when

considering turn-taking. The BVI person could be equipped with wearable technology, with the possibility that the other people in the JA triad (or circle) would wear an accompaniment. Feedback, such as light signalling, from the device to the sighted JA partners to indicate crossmodal (non-verbal) intent could be a valuable aspect of a wearable device.

d) Consider explicit support for both the BVI person and the JA sighted partners; The BVI person may not be receiving the visual cues from the sighted person, making it difficult for them to initiate JA episodes successfully [74, 79]. However, the lack of JA opportunities may also be caused by the sighted person not recognising the signals sent from the BVI partner to initiate JA, e.g., typical visual cues such as orientation or facial expressions [75]. Therefore, due to the interchange of the processes underlying successful JA episodes, it is necessary to consider BVI's and sighted people's needs to reduce the void. Any attempt that focuses solely on translating visual information into haptic or audio information for the BVI person is likely to fail as it will not be able to address the void for the sighted person related to the lack of visual feedback (primarily orienting and eye gaze) by the BVI person. We recommend that the design of JA technology should include modes of nonverbal feedback to the sighted JA partner to suggest that their BVI JA partner may want to initiate an interaction with them. This would allow designers to consider the technology not only as an assistive technology for BVI people but as a bridge-building tool for inclusive adaptation of all parties involved in the triad.

6.4 Limitations and future work

The presented work was a first step towards understanding phenomena and behaviour around JA interaction. As such, our findings are constrained by a number of limitations. Our interviews only explored the viewpoint of BVI adults discussing JA with sighted adults. We are conducting a further study to explore the viewpoints of BVI people interacting with other BVI people and sighted people interacting with BVI people. This further research will provide a rounded analysis of the whole JA triad. Our observation study focused on the initiation of JA because we identified it as a key first step in the process of JA. However, further research needs to explore how JA is maintained, how it breaks down, how it is re-initiated, and how it ends. Additionally, the BVI children were all recorded in JA-rich scenarios with parents, teachers, siblings and friends. Future work should explore JA experiences for BVI children and adults in different scenarios, including groups. The next data collection steps are to vary the children's ages and explore new scenarios that include known and unknown peers in less structured situations, such as in the playground or during unstructured play. Future observations could also involve adults in workplace settings or social gatherings with friends, family and strangers.

7 CONCLUSION

We interviewed BVI adults and analysed videos of BVI children and sighted adults engaging in activities conducive to JA. We moved beyond a sighted person's interpretation of JA to understand how BVI people interpret their experiences so that we can continue to confidently identify what BVI children and their sighted peers need to initiate JA episodes successfully and what happens when

they break down through factors such as physical distance. We found that lack of JA feedback is perceived as voids that block engagement, exacerbated in group settings, with an emphasis on oneself to fill those voids. We demonstrated the negative impact of a lack of vision and a lack of touch on JA, a frequent occurrence in public environments such as schools and employment settings. Our video analysis anchored the absence of the person element within typical JA triads. We suggest, therefore, a potential for technology to foster alternative dynamics between BVI and sighted people. We identify a need for more research in this area so that future design considers how BVI people and their sighted JA partners adapt to new technology whilst avoiding overwhelming other senses and masking natural behaviours so that BVI people can have more joint attention opportunities in their social lives.

REFERENCES

[1] Lauren B. Adamson, Roger Bakeman, and Deborah F. Deckner. 2004. The development of symbol-infused joint engagement. *Child Development* 75, 4 (2004), 1171–1187. https://doi.org/10.1111/j.1467-8624.2004.00732.x

[2] John R Aiello and Tyra De Carlo Aiello. 1974. The development of personal space: Proxemic behavior of children 6 through 16. *Human ecology* 2 (1974), 177–189.

[3] Andrea Urqueta Alfaro, Valerie S Morash, Daisy Lei, and Deborah Orel-Bixler. 2018. Joint engagement in infants and its relationship to their visual impairment measurements. *Infant Behavior and Development* 50 (2018), 311–323. https://doi.org/10.1016/j.infbeh.2017.05.010

[4] Mohammed Ibrahim Alhojailan. 2012. Thematic analysis: a critical review of its process and evaluation. In *WEI international European academic conference proceedings*. West East Institute, Zagreb, Croatia, 8–21. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=0c66700a0f4b4a0626f87a3692d4f34e599c4d0e

[5] Mohammad Anwar, Donni Prakosha, Supriyadi Supriyadi, and Intan Nastiti. 2022. Blindism Behavior of Visually Impaired Children That Hinder The Learning Process In School. In *Proceedings of the 5th International Conference on Learning Innovation and Quality Education* (Surakarta, Indonesia) *(ICLIQE '21)*. Association for Computing Machinery, New York, NY, USA, Article 65, 4 pages. https://doi.org/10.1145/3516875.3516953

[6] Dimitris Argiropoulos. 2020. The Blind Child During Early Childhood (0-3 Years) and the Educational Intervention. *Psychological Counseling and Psychotherapy* 13 (2020), 68–86. https://doi.org/10.26565/2410-1249-2020-13-08

[7] Lucas Battich, Merle Fairhurst, and Ophelia Deroy. 2020. Coordinating attention requires coordinated senses. *Psychonomic bulletin & review* 27 (2020), 1126–1138.

[8] Andrew P Bayliss, Emily Murphy, Claire K Naughtin, Ada Kritikos, Leonhard Schilbach, and Stefanie I Becker. 2013. "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General* 142, 1 (2013), 76.

[9] Lope Ben Porquis, Sara Finocchietti, Giorgio Zini, Giulia Cappagli, Monica Gori, and Gabriel Baud-Bovy. 2017. ABBI: A wearable device for improving spatial cognition in visually-impaired children. In *2017 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. 1–4. https://doi.org/10.1109/BIOCAS.2017.8325128

[10] Cynthia L. Bennett, Jane E, Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173650

[11] Alexy Bhowmick and Shyamanta M. Hazarika. 2017. An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends. *Journal on Multimodal User Interfaces* 11, 2 (2017), 149–172. https://doi.org/10.1007/s12193-016-0235-6

[12] Arkadiusz Bialek, Małgorzata Stępień-Nycz, and Marta Bialecka-Pikul. 2013. Responding to Joint Attention Predicts Joint Action. 10th International Gesture Workshop (GW) and the 3rd Gesture and Speech in Interaction (GESPIN) conference.

[13] Ann E Bigelow. 1986. The development of reaching in blind children. *British Journal of Developmental Psychology* 4, 4 (1986), 355–366. https://doi.org/10.1111/j.2044-835X.1986.tb01031.x

[14] Ann E Bigelow. 1995. The effect of blindness on the early development of the self. , 327–347 pages. https://doi.org/10.1016/S0166-4115(05)80018-4

[15] Ann E Bigelow. 2003. The development of joint attention in blind infants. *Development and psychopathology* 15, 2 (2003), 259–275.

[16] Ann E. Bigelow, Kim MacLean, and Jane Proctor. 2004. The role of joint attention in the development of infants' play with objects. *Developmental Science* 7, 5 (2004), 518–526. https://doi.org/10.1111/j.1467-7687.2004.00375.x

[17] Maria Botero. 2016. Tactless scientists: Ignoring touch in the study of joint attention. *Philosophical Psychology* 29, 8 (2016), 1200–1214.

[18] Monique Botha, Jacqueline Hanlon, and Gemma Louise Williams. 2023. Does language matter? Identity-first versus person-first language use in autism research: A response to Vivanti. *Journal of autism and developmental disorders* 53, 2 (2023), 870–878.

[19] Yassine Bouteraa. 2021. Design and development of a wearable assistive device integrating a fuzzy decision support system for blind and visually impaired people. *Micromachines* 12, 9 (2021), 1082.

[20] Stacy M. Branham and Shaun K. Kane. 2015. Collaborative Accessibility: How Blind and Sighted Companions Co-Create Accessible Home Spaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2373–2382. https://doi.org/10.1145/2702123.2702511

[21] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[22] Ingar Brink. 2004. Joint attention, triangulation and radical interpretation: A problem and its solution. *Dialectica* 58, 2 (2004), 179–206.

[23] Dorothy Burlingham. 1979. To be blind in a sighted world. *The Psychoanalytic study of the child* 34, 1 (1979), 5–30.

[24] Héctor Caltenco, Charlotte Magnusson, Bitte Rydeman, Sara Finocchietti, Giulia Cappagli, Elena Cocchi, Lope Ben Porquis, Gabriel Baud-Bovy, and Monica Gori. 2018. Co-designing wearable technology together with visually impaired children. In *Wearable Technologies: Concepts, Methodologies, Tools, and Applications*. IGI Global, 314–332.

[25] Hilary D Cass, Patricia M Sonksen, and Helen R McConachie. 1994. Developmental setback in severe visual impairment. *Archives of Disease in Childhood* 70, 3 (1994), 192–196.

[26] Maricel Cigales, Tiffany Field, Ziarat Hossain, Martha Pelaez-Nogueras, and Jacob Gewirtz. 1996. Touch among children at nursery school. *Early child development and care* 126, 1 (1996), 101–110.

[27] Katia Miguel Colus, Amorim de Souza, et al. 2019. Establishment of Joint Attention in a Baby with Severe Visual Impairment/O Estabelecimento da Atencao Conjunta em um Bebe com Deficiencia Visual Severa. *Psicologia: Teoria e Pesquisa* 35 (2019), NA–NA.

[28] Anna Cook, Jane Ogden, and Naomi Winstone. 2018. Friendship motivations, challenges and the role of masking for girls with autism in contrasting school settings. *European Journal of Special Needs Education* 33, 3 (2018), 302–315. https://doi.org/10.1080/08856257.2017.1312797

[29] Luigi F Cuturi, Elena Aggius-Vella, Claudio Campus, Alberto Parmiggiani, and Monica Gori. 2016. From science to technology: Orientation and mobility in blind children and adults. *Neuroscience & Biobehavioral Reviews* 71 (2016), 240–251.

[30] N. Dale and A. Salt. 2007. Early support developmental journal for children with visual impairment: the case for a new developmental framework for early intervention. *Child: Care, Health and Development* 33, 6 (2007), 684–690. https://doi.org/10.1111/j.1365-2214.2007.00798.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2214.2007.00798.x

[31] Maitraye Das, Darren Gergle, and Anne Marie Piper. 2019. "It doesn't win you friends" Understanding Accessibility in Collaborative Writing for People with Vision Impairments. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.

[32] Sheila Degotardi. 2017. Joint attention in infant-toddler early childhood programs: Its dynamics and potential for collaborative learning. *Contemporary Issues in Early Childhood* 18, 4 (2017), 409–421.

[33] Dana S Dunn and Erin E Andrews. 2015. Person-first and identity-first language: Developing psychologists' cultural competence using disability language. *American Psychologist* 70, 3 (2015), 255.

[34] J Faria, S Lopes, H Fernandes, P Martins, and J Barroso. 2010. Electronic white cane for blind people navigation assistance. In *2010 World Automation Congress*. 1–7.

[35] Anika Fiebich and Shaun Gallagher. 2013. Joint attention in joint action. *Philosophical Psychology* 26, 4 (2013), 571–587. https://doi.org/10.1080/09515089.2012.690176

[36] Tiffany Field, Barbara Soliday, David Lasko, Nini Gonzalez, and Chad Valdeon. 1994. Touching in infant, toddler, and preschool nurseries. *Early Child Development and Care* 98, 1 (1994), 113–120.

[37] Sara Finocchietti, Giulia Cappagli, Gabriel Baud-Bovy, Charlotte Magnusson, Hector Caltenco, Graham Wilson, Stephen Brewster, Ann-Kathrin Rogge, Brigitte Roeder, Elena Cocchi, Elisabetta Capris, Paola Campana, Carla Gilio, and Monica Gori. 2015. ABBI, a New Technology for Sensory-motor Rehabilitation of Visual Impaired People. International Conference on Enabling Access for Persons with Visual Impairment.

[38] Patrick Flink. 2021. Person-first & identity-first language: Supporting students with disabilities on campus. *Community College Journal of Research and Practice* 45, 2 (2021), 79–85.

[39] Euan Freeman, Graham Wilson, Stephen Brewster, Gabriel Baud-Bovy, Charlotte Magnusson, and Hector Caltenco. 2017. Audible Beacons and Wearables in Schools: Helping Young Visually Impaired Children Play and Move Independently. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4146–4157. https://doi.org/10.1145/3025453.3025518

[40] J Antonio Garcia-Macias, Alberto G Ramos, Rogelio Hasimoto-Beltran, and Saul E Pomares Hernandez. 2019. Uasisi: A modular and adaptable wearable system to assist the visually impaired. *Procedia Computer Science* 151 (2019), 425–430.

[41] Matthias S Gobel, Miles RA Tufft, and Daniel C Richardson. 2018. Social beliefs and visual attention: how the social relevance of a cue influences spatial orienting. *Cognitive science* 42 (2018), 161–185.

[42] Monica Gori, Giulia Cappagli, Alessia Tonelli, Gabriel Baud-Bovy, and Sara Finocchietti. 2016. Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. *Neuroscience & Biobehavioral Reviews* 69 (2016), 79–88.

[43] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11 (2002), 411–446.

[44] Marion A Hersh and Michael A Johnson. 2008. On modelling assistive technology systems–Part I: Modelling framework. *Technology and disability* 20, 3 (2008), 193–215.

[45] Masako Hirotani, Manuela Stets, Tricia Striano, and Angela D Friederici. 2009. Joint attention helps infants learn new words: event-related potential evidence. *Neuroreport* 20, 6 (2009), 600–605.

[46] Hanen Jabnoun, Faouzi Benzarti, and Hamid Amiri. 2017. A new method for text detection and recognition in indoor scene for assisting blind people. In *Ninth International Conference on Machine Vision (ICMV 2016)*, Vol. 10341. SPIE, 414–418. https://doi.org/10.1117/12.2268399

[47] Carey Jewitt and Kerstin Leder Mackley. 2019. Methodological dialogues across multimodality and sensory ethnography: digital touch communication. *Qualitative Research* 19, 1 (2019), 90–110.

[48] Adam Kendon. 2010. *Spacing and Orientation in Co-present Interaction*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–15. https://doi.org/10.1007/978-3-642-12397-9_1

[49] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. 2016. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 20, 4 (2016), 442–462.

[50] Markus Lang, Manfred Hintermair, and Klaus Sarimski. 2017. Social-emotional competences in very young visually impaired children. *British Journal of Visual Impairment* 35, 1 (2017), 29–43.

[51] Sooyeon Lee, Madison Reddie, Chun-Hua Tsai, Jordan Beck, Mary Beth Rosson, and John M. Carroll. 2020. The Emerging Professional Practice of Remote Sighted Assistance for People with Visual Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376591

[52] O Levtzion-Korach, A Tennenbaum, R Schnitzer, and A Ornoy. 2000. Early motor development of blind children. *Journal of Paediatrics and Child Health* 36, 3 (2000), 226–229. https://doi.org/10.1046/j.1440-1754.2000.00501.x

[53] London School of Paediatrics 2023. *Developmental Station*. London School of Paediatrics. Retrieved May 2023 from https://mrcpch.paediatrics.co.uk/development/

[54] Jessica Markus, Peter Mundy, Michael Morales, Christine EF Delgado, and Marygrace Yale. 2000. Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development* 9, 3 (2000), 302–315.

[55] Paul Marshall, Yvonne Rogers, and Nadia Pantidi. 2011. Using F-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (Hangzhou, China) *(CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 445–454. https://doi.org/10.1145/1958824.1958893

[56] Linda M McAlpine and Chris L Moore. 1995. The development of social understanding in children with visual impairments. *Journal of Visual Impairment & Blindness* 89, 4 (1995), 349–358.

[57] Daniel Mendes, Sofia Reis, João Guerreiro, and Hugo Nicolau. 2020. Collaborative tabletops for blind people: The effect of auditory design on workspace awareness. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS (2020), 1–19.

[58] Mohamed Dhiaeddine Messaoudi, Bob-Antoine J Menelas, and Hamid Mcheick. 2022. Review of Navigation Assistive Tools and Technologies for the Visually

Impaired. *Sensors* 22, 20 (2022), 7888.

[59] Oussama Metatla, Sandra Bardot, Clare Cullen, Marcos Serrano, and Christophe Jouffrais. 2020. Robots for Inclusive Play: Co-designing an Educational Game With Visually Impaired and sighted Children. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376270

[60] Oussama Metatla, Nick Bryan-Kinns, and Tony Stockman. 2018. "I Hear You": Understanding Awareness Information Exchange in an Audio-only Workspace. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174120

[61] Oussama Metatla, Nick Bryan-Kinns, and Tony Stockman. 2012. Supporting cross-modal collaboration in the workplace. In *The 26th BCS Conference on Human Computer Interaction 26*. 109–118.

[62] Oussama Metatla, Nick Bryan-Kinns, Tony Stockman, and Fiore Martin. 2013. Activity Theory as a Tool for Identifying Design Patterns in Cross-Modal Collaborative Interaction. In *Human-Computer Interaction – INTERACT 2013*, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 232–240.

[63] Oussama Metatla, Nuno N. Correia, Fiore Martin, Nick Bryan-Kinns, and Tony Stockman. 2016. Tap the ShapeTones: Exploring the Effects of Crossmodal Congruence in an Audio-Visual Interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1055–1066. https://doi.org/10.1145/2858036.2858456

[64] Oussama Metatla and Clare Cullen. 2018. "Bursting the Assistance Bubble": Designing Inclusive Technology with Children with Mixed Visual Abilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173920

[65] Oussama Metatla, Alison Oldfield, Taimur Ahmed, Antonis Vafeas, and Sunny Miglani. 2019. Voice User Interfaces in Schools: Co-designing for Inclusion with Visually-Impaired and Sighted Pupils. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300608

[66] Cecily Morrison, Ed Cutrell, Martin Grayson, Geert Roumen, Rita Faia Marques, Anja Thieme, Alex Taylor, and Abigail Sellen. 2021. PeopleLens. *Interactions* 28, 3 (2021), 10–13.

[67] Cecily Morrison, Edward Cutrell, Martin Grayson, Anja Thieme, Alex Taylor, Geert Roumen, Camilla Longden, Sebastian Tschiatschek, Rita Faia Marques, and Abigail Sellen. 2021. Social Sensemaking with AI: Designing an Open-ended AI Experience with a Blind Child. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 396, 14 pages. https://doi.org/10.1145/3411764.3445290

[68] Yukie Nagai, Koh Hosoda, Akio Morita, and Minoru Asada. 2003. Emergence of joint attention based on visual attention and self learning. In *Int. Symposium on Adaptive Motion of Animals and Machines*.

[69] Royal National Institute of Blind People. 2021. *Key information and statistics on sight loss in the UK*. Royal National Institute of Blind People. Retrieved April 2023 from https://www.rnib.org.uk/professionals/health-social-care-education-professionals/knowledge-and-research-hub/key-information-and-statistics-on-sight-loss-in-the-uk

[70] Eugenio Parise, Allison Cleveland, Angela Costabile, and Tricia Striano. 2007. Influence of vocal cues on learning about objects in joint attention contexts. *Infant Behavior and Development* 30, 2 (2007), 380–384. https://doi.org/10.1016/j.infbeh.2006.10.006

[71] Dianne TV Pawluk, Richard J Adams, and Ryo Kitada. 2015. Designing haptic assistive technology for individuals who are blind or visually impaired. *IEEE transactions on haptics* 8, 3 (2015), 258–278.

[72] R. Peter Hobson and Anthony Lee. 2010. Reversible autism among congenitally blind children? A controlled follow-up study. *Journal of Child Psychology and Psychiatry* 51, 11 (nov 2010), 1235–1241. https://doi.org/10.1111/j.1469-7610.2010.02274.x

[73] Caroline Pigeon, Tong Li, Fabien Moreau, Gilbert Pradel, and Claude Marin-Lamellet. 2019. Cognitive load of walking in people who are blind: Subjective and objective measures for assessment. *Gait & posture* 67 (2019), 43–49.

[74] Shi Qiu, Pengcheng An, Jun Hu, Ting Han, and Matthias Rauterberg. 2020. Understanding visually impaired people's experiences of social signal perception in face-to-face communication. *Universal Access in the Information Society* 19 (2020), 873–890.

[75] Shi Qiu, Jun Hu, Ting Han, Hirotaka Osawa, and Matthias Rauterberg. 2020. Social glasses: simulating interactive gaze for visually impaired people in face-to-face communication. *International Journal of Human–Computer Interaction* 36, 9 (2020), 839–855.

[76] Ali Jasim Ramadhan. 2018. Wearable smart system for visually impaired people. *Sensors* 18, 3 (2018), 843.

[77] Ingvill Rasmussen, Anja Amundrud, and Sten Ludvigsen. 2019. Establishing and maintaining joint attention in classroom dialogues: Digital technology, microblogging and ground rules. In *The Routledge international handbook of research on dialogic education*. Routledge, 411–424.

[78] Sally J Rogers and Carol B Puchalski. 1984. Social characteristics of visually impaired infants' play. *Topics in Early Childhood Special Education* 3, 4 (1984), 52–56.

[79] Lynsey K Romo, Cimmiaron Alvarez, and Melissa J Taussig. 2023. An examination of visually impaired individuals' communicative negotiation of face threats. *Journal of Social and Personal Relationships* 40, 1 (2023), 152–173.

[80] Mary Christine Ross. 2017. *Promoting joint attention in children with visual impairment: Proposing an intervention using modified strategies from joint attention symbolic play engagement regulation (JASPER)*. Ph. D. Dissertation. The Ohio State University.

[81] Federico Rossano, Malinda Carpenter, and Michael Tomasello. 2012. One-Year-Old Infants Follow Others' Voice Direction. *Psychological Science* 23, 11 (nov 2012), 1298–1302. https://doi.org/10.1177/0956797612450032

[82] Nuzhah Gooda Sahib, Tony Stockman, Anastasios Tombros, and Oussama Metatla. 2013. Participatory Design with Blind Users: A Scenario-Based Approach. In *Human-Computer Interaction – INTERACT 2013*, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 685–701.

[83] Shatha N. Samman, Kay M. Stanney, Joseph Dalton, Ali M. Ahmad, Clint Bowers, and Valerie Sims. 2004. Multimodal Interaction: Multi-Capacity Processing Beyond 7+/-2. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 48, 3 (2004), 386–390. https://doi.org/10.1177/154193120404800324 arXiv:https://doi.org/10.1177/154193120404800324

[84] Jaime Sánchez and Nelson Baloian. 2006. Issues in Implementing Awareness in Collaborative Software for Blind People. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Joachim Klaus, Wolfgang L. Zagler, and Arthur I. Karshmer (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1318–1325.

[85] Suraj Singh Senjam. 2019. Assistive Technology for People with Visual Loss. *Delhi Journal of Ophthalmology* 30 (2019), 7–12. Issue 2. https://doi.org/10.7869/djo.496

[86] Suraj Singh Senjam, Allen Foster, Covadonga Bascaran, Praveen Vashist, and Vivek Gupta. 2020. Assistive technology for students with visual disability in schools for the blind in Delhi. *Disability and rehabilitation. Assistive technology* 15, 6 (2020), 663–669. https://doi.org/10.1080/17483107.2019.1604829

[87] Kristen Shinohara. 2006. Designing assistive technology for blind users. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) *(Assets '06)*. Association for Computing Machinery, New York, NY, USA, 293–294. https://doi.org/10.1145/1168987.1169062

[88] Garriy Shteynberg, Jacob B Hirsh, Adam D Galinsky, and Andrew P Knight. 2014. Shared attention increases mood infusion. *Journal of Experimental Psychology: General* 143, 1 (2014), 123.

[89] Barbora Skarabela, Shanley E.M. Allen, and Thomas C. Scott-Phillips. 2013. Joint attention helps explain why children omit new referents. *Journal of Pragmatics* 56, 1 (2013), 5–14. https://doi.org/10.1016/j.pragma.2012.08.003

[90] Katta Spiel, Kathrin Gerling, Cynthia L. Bennett, Emeline Brulé, Rua M. Williams, Jennifer Rode, and Jennifer Mankoff. 2020. Nothing About Us Without Us: Investigating the Role of Critical Disability Studies in HCI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3375150

[91] Emilia Strid and Asta Cekaite. 2022. Calibrating joint attention and affective stances in young children's peer interactions. *Journal of Pragmatics* 198 (2022), 29–42. https://doi.org/10.1016/j.pragma.2022.05.016

[92] Juulia T Suvilehto, Lauri Nummenmaa, Tokiko Harada, Robin IM Dunbar, Riitta Hari, Robert Turner, Norihiro Sadato, and Ryo Kitada. 2019. Cross-cultural similarity in relationship-specific social touching. *Proceedings of the Royal Society B* 286, 1901 (2019), 20190467.

[93] Valerie Tadić, Linda Pring, and Naomi Dale. 2009. Attentional processes in young children with congenital visual impairment. *British Journal of Developmental Psychology* 27, 2 (2009), 311–330.

[94] Anja Thieme, Cynthia L. Bennett, Cecily Morrison, Edward Cutrell, and Alex S. Taylor. 2018. "I can do everything but see!" – How People with Vision Impairments Negotiate their Abilities in Social Contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14.

https://doi.org/10.1145/3173574.3173777

[95] Michael Tomasello. 1999. *The cultural origins of human cognition.* Harvard university press.

[96] Michael Tomasello. 2003. *Constructing a language: A usage-based theory of language acquisition.* Harvard university press.

[97] Muhib Ullah, Shah Khusro, Mumtaz Khan, Iftikhar Alam, Inayat Khan, and Badam Niazi. 2022. Smartphone-Based Cognitive Assistance of Blind People in Room Recognition and Awareness. *Mobile Information Systems* 2022 (2022).

[98] UNGeneva 2024. *Disability Inclusive Language Guidelines.* UNGeneva. Retrieved 12 February 2024 from https://www.ungeneva.org/en/about/accessibility/disability-inclusive-language

[99] Prabhat Verma, Raghuraj Singh, and Avinash Kumar Singh. 2012. A framework for the next generation screen readers for visually impaired. *International Journal of Computer Applications* 47 (2012), 31–38. Issue 10.

[100] François Vital-Durand, Valérie Mazoyer, Françoise Koenig, and Serge Portalier. 1998. Visual impairment: specialized education for infants and rehabilitation for adults. (1998).

[101] Herman A. Witkin, Judith Birnbaum, Salvatore Lomonaco, Suzanne Lehr, and Judith L. Herman. 1968. Cognitive Patterning in Congenitally Totally Blind Children. *Child Development* 39, 3 (1968), 767–786. http://www.jstor.org/stable/1126982