Designing for Joint Attention and Co-presence Across Parallel Realities

Rolf Steier, University of Oslo, rolf.steier@iped.uio.no

Abstract: Most studies of joint attention and co-presence as mechanisms for learning are built on the premise that collaborators are engaged in a shared perceptual experience. This paper identifies new ways that visitors negotiate joint attention and co-presence while making meaning inside and outside of a multisensory virtual reality museum exhibition. The study is based on a large-scale design-based research project in which an interdisciplinary team designed and built an immersive architecture exhibition. Informed by perspectives on embodied interaction and meaning making, interaction analysis methods were applied to video data of pairs of visitors as they took turns exploring and guiding each other through the experience. Findings suggest that in situations where gaze and gesture cannot be relied on, visitors develop creative means to develop shared understandings. The study contributes to conceptions of presence in meaning making and the design of collaborative learning interactions across perceptual experiences.

The distinction between face-to-face learning, and online or digitally mediated collaborative learning is increasingly becoming blurred. New situations and technologies are challenging what it means to be present in a particular place, and participants can work together while engaged in a variety of perceptual experiences; across spaces, times, and resources. For example, classroom configurations may be built on a combination of face-to-face participants interacting with those in other places through web and video, or else built on blended combinations of physical and digital infrastructure (Enyedy, Danish, & DeLiema, 2015). However, we have little understanding of how learners negotiate joint attention and co-presence when they know that they are not sharing the same visual and sensory fields. Such situations have important implications for how we design learning activities that enable collaborators to coordinate and develop intersubjectivity. In this study, the design of an immersive multisensory architecture exhibition is presented as a setting to investigate the ways that visitors negotiate joint attention and co-presence while interacting from within and across the virtual scenes and physical installation. The following research question is addressed*: how do visitors negotiate joint attention and co-presence across parallel sensory experiences to develop shared meanings?*

Perspectives on joint attention and co-presence

Joint attention, and the related phenomenon of co-presence, are core mechanisms for collaborative learning, and as such have a long tradition as objects of study in the learning sciences. Joint attention is a social phenomenon that is achieved when two people are attending to the same feature of their shared environment, *and* this attention is coordinated through mutual awareness (Tomasello, 1995). Sociocultural and interactional perspectives approach learning as a process of meaning making in which co-participants construct joint interpretations of their context (Greeno, 2006, Suthers, 2006). Accordingly, joint attention plays an important role in meaning making by supporting such learners in establishing common premises, by facilitating referential practices, and by allowing for mutual elaboration and coordinated action. Joint attention is thus a relevant feature in all socially situated learning contexts from early childhood onward (Tomasello, 1995). In small group contexts in particular, the establishment of shared understandings and frames are crucial for productive learning and problem solving (Stahl, 2013; Mercer & Wegerif, 1999).

If joint attention involves people focusing together on a particular feature of the environment, co-presence can be understood as occurring when people develop a shared understanding of that environment itself. Co-presence is negotiated by participants to establish shared temporal, spatial, or conceptual references among which joint attention may be focused. In the classic methodological article on interaction analysis, Jordan and Henderson (1995) note that, "the physical co-presence of persons is always managed by socially recognized (although often unstated) expectations regarding occupancy of space, interaction with others, use of objects and resources, display of physical presence, and voice" (p72). These shared expectations in turn structure activity.

We typically think of joint attention and co-presence as being in alignment. That is, persons physically in the same shared space have access to the same resources and references. Accordingly, many studies in the learning sciences have investigated the ways that joint attention is achieved in face-to-face interaction. Such studies have investigated joint attention by considering the gaze or bodily orientation of small groups of collaborators in relation to mediating resources (e.g. Barron, 2000; Goodwin, 2007). For example Barron (2000)

studied small groups of students engaged in problem solving tasks by focusing on their gaze and orientation to a shared workbook, and found they were less successful when they failed or struggled to achieve joint attention. When collaborators are physically oriented towards the same artifact, and the artifact is also present in participant talk, this can be an indicator of joint attention. New developments in eye tracking technology allow systematic evaluation of joint attention during problem solving activities by recording precisely when visual attention moves in and out of alignment (Schneider et al., 2018). Deictic gestures, or the use of tools such as laser pointers or flashlights can also facilitate joint attention among learners in educational contexts while also being used as an indicator for researchers of precisely how attention shifts within and among visual features of the environment. In the museum context, Povis & Crowley (2015) found that flashlights mediate joint attention on objects and artifacts by visitors and thus support family learning conversations.

However, when collaborators are not in the same shared space, other means that don't rely on bodily action must be developed to establish common premises. Notably, technology mediated interactions create such contexts. Dillenbourg & Traum (2006) summarize this distinction clearly:

> For example, face-to-face communication differs from telephone communication by including constraints of visability and visual co-presence. In face-to-face communication, pointing and gaze can be used as a complementary channel to express some information (such as the fact that the communicators are talking about the same object), whereas in telephone conversations, one must rely solely on the audio channel to try to coordinate such information (e.g., by asking questions and giving descriptions of what one is looking at). (p125)

For cases when collaborators do not share a visual field, affordances for joint attention can be intentionally introduced into learning resources. Stahl (2013) presents the design of an online platform for collaborative geometry education called Virtual Math Teams. Students using this platform interact solely through a computer interface, but are able to achieve joint attention through features such as the shared control and manipulation of mathematical objects or a text-based chat room. Just as with physically co-present collaborators, digital co-presence can be achieved through multiple modes (here through text and artifact manipulation). Importantly, participants in this platform may not be in the same physical space, but are interacting through the same interface with access to visual information that is largely the same. Here the establishment of co-presence around (mathematical) objects thus depends in part on the assumption by collaborators that they are each seeing the same things. Yet how do learners respond when this assumption is challenged? The following architecture museum exhibition presents a setting for investigating this question.

Background: Design of a multisensory immersive architecture exhibition

Meaning making in architecture can be understood as a multisensory process in which individuals or groups attend to sensory features including sounds, smells, touch, as well as the visuo-spatial surroundings in order to develop interpretations of the place (Pallasmaa, 2012). In the social context, the production of meaning making through collective interpretation thus relies on making such aspects relevant through notions of joint attention and co-presence. The multisensory experience of actually being in architecture presents unique challenges to museums that often instead must rely on abstracted representations such as drawings and models. The potential for virtual reality and other immersive environments to highlight such experiences serves as a backdrop for this study.

Figure 1a,b,c. The architecture exhibition including left to right: the screen, the platform, and the VR view.

This study was conducted through a large-scale design-based research project based on a collaboration between an interdisciplinary team of museum curators, university researchers, a renowned architecture firm, and technology and sound designers. The aims of the project were multifaceted reflecting the different research interests of the project partners including curatorial, architectural, and educational agendas. After developing, testing, and iterating a series of small-scale prototypes, the design process culminated in a museum exhibition called *The Forest in the House: Exploring Parallel Realities* installed in a large pavilion in a national architecture museum. The design was informed by the architecture firm's interest in the relationship between nature and architecture, and in exploring bodily and sensory aspects of moving through such spaces. The design involved laser-scanning a site in a coastal forest, constructing a digital model of the scan, and translating that model into architectural form such that, for example, trees became columns and canopy became ceiling.

Visitors to the exhibition were greeted by a four-by-six-meter platform containing steps and columns and surrounded by an array of speakers (Figure 1b). A large screen to the side displayed the view of visitors from inside the VR scenes (Figure 1a). Visitors were invited to participate in pairs with one partner wearing the VR headset, hand, and feet sensors, as well as open-ear headphones, while the other partner served as a discussion partner and safety guide. The virtual reality component consisted of two scenes: the natural landscape, and a house inspired by the natural forms and placed on the same site (Figure 1c). Visitors wearing the VR headset could walk around, up and down steps, and explore the space because the physical structure corresponded to the virtual scene. They could switch between the two scenes, and the soundscape also changed depending on the natural or architectural scene. The pairs could move around the space together switching roles after 10 or 15 minutes. The learning aims of the design included supporting visitors in sensing and sharing the relationships between nature and architecture, as well as body, space, and movement.

Consideration for visitors' experiences of joint attention and co-presence was purposely built into the design of the exhibition. First, the exhibition was designed to support only one visitor at a time within VR. This was partially due to technical constraints, but the role of the partner/ guide was introduced to support a social experience typical of museum visits and also to provide interactional data to be studied by the research team. Additionally, the large projector screen displayed a feed from the visitor's perspective inside the VR. These views were of course not identical as the VR experience affords looking in all directions producing a full sense of immersion, while the screen was cropped to a rectangular image. This screen allowed the partners, along with others in the room, to view some of what the VR-user was seeing. Similarly, though the speaker array provided a shared soundscape in the space, the visitor in VR also wore headphones which produced an additional layer of location-based sound. Opportunities for co-presence thus depended on complex interplay between parallel physical, visual, auditory, and social experiences.

Methods

Prior to the public opening of the exhibition, 17 pairs of visitors were recruited to participate in the study. To account for diverse museum demographics of interest to the museum, participants included a variety of ages from children to retirees, and levels of expertise including architects, students, and non-experts. Participants were filmed while they explored the installation together and as they switched between the two roles. The VR views were also recorded to document the visual experience of the headset-wearing visitor. Afterwards, they were separated for a short post-visit survey and then they rejoined for post-visit group interview. This study focuses on the negotiation of joint attention and co-presence and thus draws on the interactional video data for analysis.

Following multimodal interaction analysis methods (Hall & Stevens, 2015; Jordan & Henderson, 1995) video data were first transcribed and then iteratively viewed and sorted identifying sequences in which joint attention and co-presence became relevant. These sequences were coded first based on participants' different use of senses in talk and activity, examined for patterns and grouped into emergent interactional themes. Exemplary sequences were then identified, reviewed, and analyzed by the author as well as in a research group setting. To aid in the analysis of these sequences, video data from the cameras were combined with recordings from the VR scenes to allow comparison between the two visitor partners' perspectives. In the section below, particularly rich sequences are presented in order to illustrate as well as theorize about the interactional meaning making processes. The goal is not to present an exhaustive generalizable typology of joint attention and co-presence, but to expand our understanding of these mechanisms in light of technological and social-interactional developments.

Data and analysis

This particular exhibition design, by inviting a social experience between two visitors in 'parallel realities', was intentionally designed to explore issues of presence and co-presence in order to investigate the potential of VR as a social experience in museums. Accordingly seemingly simple everyday questions such as "Where are you?"

take on interesting and ambiguous meanings depending on who is asking and who is answering the question. An answer to the question could include a physical location on the exhibition platform, a distinction between being in the natural or architectural VR scenes, a particular location within one such scene, or even the imagined site upon which this design was based. In investigating the core research question about the negotiation of joint attention and co-presence, four interactional themes were identified and are presented below. Note that pseudonyms are used and that talk has been translated from the original language.

Inhabiting

One interactional theme that emerged involved visitors imagining what it would be like to actually live in the virtual house. Visitors following this pattern had conversations about what their lives would be like in this place, and the activities they would perform. This theme is exemplified by the sequence below in which two men plan an imagined meal in the place. They have switched roles a few minutes prior, so Tom having already been in the VR is guiding, while Martin is now in the VR.

Tom: It could be comfortable here, right?

Martin: I suppose so

Tom: Yes.

Martin: Especially with the weather here

Tom: You can pick up a bottle of white wine.

Martin: Yeah.

Tom: Champagne maybe, isn't that what one drinks?

Martin: Some shrimp and a baguette

In this sequence, Tom extends the discussion from being primarily descriptive to entering some future possibility noting that "it could be comfortable". This turn opens up a shared co-presence not in the present reality but in an alternate temporality and version of the place. Martin agrees mentioning the weather, which was a common point of discussion among visitors as the virtual scene was built around a blue sky with strong sunlight. The pair then further extends this hypothetical situation from merely *being* here to *doing* something, to eating and drinking a rather luxurious meal. It was common for visitors to view the scene as a kind of vacation home with the rather abstract depiction of an uninhabited house becoming a place to be lived in. This negotiation of co-presence can be understood as a form of collective imagining in which the pair responds to the hypothetical possibility of being in this house as though it were real (Nemirovsky et al., 2012; Steier & Kersting, 2019). In this case, the imagining takes the form of 'inhabiting' as the pair enacts what it might be like to live in this house. Though they are both physically located in the museum exhibition, and only one is present in the VR scene, they establish a co-presence in this alternate place in spacetime. Interestingly, Tom may be relying on his previous experience from a few minutes prior as the VR participant to consider this hypothetical, while Martin is able to extend his current perceptual experience. This sequence demonstrates an important temporal orientation to co-presence (Pierroux, Steier, & Sauge, 2019).

Guided looking

Another interactional pattern that emerged can be referred to as guided looking. Though the partner wearing the VR headset had control of the view, by moving around the space and turning their head, frequently the guide would direct the activity by asking the partner to move or look at something in particular. As we enter the sequence below, the pair is largely silent while Adam gazes around the virtual nature scene and Julie follows along on the screen. The sounds of waves on the coast are audible from the speaker array. Julie then gives instruction to Adam.

Julie: Look behind. Also look at little further to the left. No no stop. Also press and go into the house again. ((Figures 2a & 2b))

Adam: Yes. ((Triggers button to switch to architecture scene))

Julie: Do you see the post? ((Figure 2c))

Adam: Oh yes.

Julie: Which is a tree.

Adam: Which is a tree, yes I see it. ((Adam switches back and forth between nature and architecture scenes - Figures 2b and 2c))

Figure 2a,b,c. Guided looking. Left to right: the visitors, the nature scene, and the architecture scene.

Julie breaks the silence by asking Adam to look in a particular direction. This is not a trivial act for several reasons. First Adam would not be able to see a pointing gesture, and also cannot see Julie's bodily orientation. Second, the object that Julie is trying to draw attention to is not currently visible to Adam. Julie has noticed a relationship between the column standing in front of her and the tree that stands before Adam (visible on the large screen). She cannot say "look at the column" because there is no column currently visible to Adam. In an attempt to achieve joint attention on this object, she directs him gradually to look behind, adjusting his direction until he is facing the correct way, telling him to "stop."

Next, she asks Adam to switch VR scenes by pressing the button on the hand sensor. Finally, after about 7 seconds, she confirms their joint attention by asking "Do you see the post?" Of course, although they are both physically standing in front of the same column, Adam is looking at a virtual object and Julie is looking at a physical object. Adam then realizes the relationship that Julie is trying to communicate. As Julie states that the column is also a tree, Adam switches back and forth between the scenes several times confirming the visual relationship between these two objects. At the same time, Julie turns back and forth between the screen behind her and the column in front.

There are several aspects to this short sequence. First, discovering and articulating this nature-architecture relationship (the post "which is a tree") is precisely the kind of learning experience intended by the exhibition designers. This is a collective interpretation that builds on the perceptual experiences of both partners. Second, this sequence exemplifies a negotiation of joint attention and co-presence in which the participants are in different perceptual fields. Julie only notices the architectural relationship by turning between the screen behind her and the platform in front of her. To make this relationship clearly visible to herself, and of course to her partner, she invites Adam to hold a steady gaze in a particular direction so that the tree is visible on the screen. Once they coordinate their orientation in this direction, they are able to build shared co-presence around the tree/ column relationship. The scene switching by Adam and the head turning by Julie are two distinct versions of the same act of comparison, which is itself an important means for interpretation in art and architecture museums (Knutson & Crowley, 2010; Steier, Pierroux, & Krange, 2015). In one sense, this sequence resembles a familiar pattern noted by Heath, vom Lehn, & Osborne (2005) in an early study of computer-based exhibits typically designed for a single user. They found that groups of visitors would have one visitor take control of the museum interactive (perhaps a mouse or touch screen) while another visitor stood to the side giving directions. The division of labor may support collaborative meaning making through different roles around a shared exhibit. Here, Julie serves this directing role while Adam literally embodies the interface by moving and turning to perform the interaction. Through acts of comparison and verbal guidance they develop co-presence in and across the scenes.

Touch

Above I raised the notion of the interesting ambiguity raised by the question "Where are you?" in this particular exhibition. Depending on the context this same question can contribute to different forms of co-presence. One instance of such an interaction plays out below as a two children use touch to co-navigate. Mindy is in the VR but has lost track of her partner, Lucas.

Mindy: Where are you? ((Figure 3a - Reaches out with left hand to find Lucas))

Lucas: I`m here. I`m standing right next to you ((Figure 3b - Reaches with right hand to tap her left shoulder))

Mindy: Okay, we can keep going.

Lucas: Okay. Where?

Mindy: I actually see almost only white now. ((3 second pause))

Mindy: Where are you?

Lucas: I`m still here ((Figure 3c - Steps in front and taps her left shoulder))

Figure 3a,b,c. Sequence of co-orientation through touch.

The use of touch as a navigational tool is not unique to VR or immersive experiences but does shed light on how people might occupy multiple spacetimes simultaneously. Mindy asks "Where are you?" and Lucas responds that he is right next to her while tapping her shoulder. The location of his voice, the verbal description of "right next to you," along with the touch on the shoulder collectively serve as orientation mediators. Co-presence in this situation is not in fact negotiated within the VR scene but within the physical reality of the installation. From the perspective of Mindy however, she may feel as though she is moving around a strange virtual scene guided by an invisible disembodied voice. Assured that they are together, Mindy says she is ready to "keep going" and continue exploring the scene. After a few seconds, she becomes disoriented though, seeing mostly white. Lucas again steps in front, taps her other shoulder this time, confirming that he is "still here".

Whereas in the previous excerpts, joint attention and co-presence were negotiated in relation to features within the virtual scene, in this excerpt Mindy and Lucas are working in relation to their own bodies in space. Though standing physically next to each other in what we might consider to be a face-to-face context, the lack of shared vision means that co-presence is disrupted. They do not have a shared awareness of each other's presence. Using a simple touch on the shoulder, they are then able to proceed with the exhibition.

Sensory narratives

Co-presence can also emerge through the use of elaborate verbal description. This interactional theme could be considered an alternative version of the guided looking described above. Instead of the outside partner giving instruction to the VR participant - the participant leads by describing what they are experiencing and translating visual and sensory phenomena into language. In the excerpt below Karl narrates his view from the VR kitchen.

Karl: Table and kitchen in front of me.

Selma: Yes.

Karl: I have also stairs up to the villa, up to the left. There is so much light here as well, there is a kind of trellis that both holds the light, keeps the warmth out, but lets in the light you might say? Around the whole villa

Selma: Yes. ((Figure 4. Laughs. Looks back at the screen))

Karl: And a strip of a garden, where do you want me to go now?

Selma: Down

Karl: Oh well

Selma: One small step forward ((Guiding Karl down the stairs))

Karl provides a rich description of the scene, noting the features of the house, his orientation to those features as well as more atmospheric conditions such as light and warmth. Selma actually does not speak much, but she confirms her presence through short utterances of "yes" which also let Karl know she is present. As Karl speaks, Selma turns back and forth between a shared orientation with her partner and the large screen behind them (Figure 4). She thus has the opportunity to align some details with those described by Karl.

Figure 4a,b. Karl narrates while Selma glances back at the screen.

Karl is in fact aware of the screen in the pavilion, having seen it prior to putting on the VR headset. The narrative performance should not be interpreted as only describing a scene which is mutually understood to be completely invisible to Selma. Actually, the interactional dynamic is reflective of typical museum conversations in which a group interacts with an artifact by first describing it to each other before elaborating on the interpretations and responses to the work. The descriptive act contributes to joint attention on the object (in this case, the features in the virtual scene) in order for a collective interpretation to develop (Povis & Crowley, 2015). As their joint attention shifts between these different features, co-presence emerges within the scene.

Then, Karl asks "where do you want me to go now?" suggesting that Selma guides the physical movement on the platform, even as Karl narrates. Selma leads Karl down the stairs, demonstrating a shift in co-presence from the virtual scene to the physical structure itself. They shift from relying on Karl to initiate attention to the virtual, to Selma initiating attention to the physical as she is able to see the real stairs that are hidden from Karl. Their negotiated co-presence is consequently best understood as a dynamic and unfolding process.

Discussion and design reflections

Looking back, this study opened with the research question: *how do visitors negotiate joint attention and co-presence across parallel sensory experiences to develop shared meanings?* The above four themes illustrate how the visitors develop multiple means to coordinate their activity. Due to the constraints of the exhibition configuration, these means do not rely on gesture and gaze as shared visible features of interaction. Instead the visitors use *touch* to align their bodies in space; they use *sensory narratives* to communicate perceptual experiences and lay a foundation for interpretive acts; they use *guided looking* to coordinate their attention using each other's vision as resources; and they engage in forms of *inhabiting*, performing imaginative practices to establish co-presence outside of the present reality. In general terms, one might be tempted to reduce the latter three patterns to simply being instances of talk similar to the telephone example described by Dillenbourg & Traum (2006). However, though these themes certainly all rely on talk to overcome the absence of a shared perceptual field, they are in fact nuanced practices with particular social, spatial, and temporal aspects.

Hall & Stevens (2015) notes that 'facing formations' (dynamic configurations of bodies and place - Kendon, 1990) are an important unit of analysis for studying knowledge in use. They specify interest in "[...] facing-formation systems that form joint attention for doing things that are specific to the conceptual practices we are studying" (p93). The above sequences include such systems oriented towards the conceptual practices of interpreting architecture and visiting an interactive exhibition. *The Forest in the House* is perhaps a unique case for investigating joint attention and co-presence as it is highly contextualized in a particular museum and particular technological and social constraints. However this uniqueness has provided an opportunity to explore phenomena that lie at the intersection between face-to-face and digital interactions. A pair of visitors could be standing side by side yet having very different perceptual experiences, and we might consider such arrangements as new facing formations. Interestingly, this formation included two visitors, yet three bodies (2 physical, 1 digital).

This general type of unbalanced or parallel perceptual experience is actually becoming more common as new digital learning infrastructures blend physical and digital experiences (e.g. Enyedy et al., 2015). The ways that participants work around the perceptual constraints of such arrangements to establish joint attention and co-presence becomes crucial for meaning making and for designing learning situations. Stahl (2013) showed how particular features can be designed into online learning environments to support co-presence such as chatrooms, shared objects, and visible cursors. In this study, the designed roles of visitor and guide provided important structure to the collaborative activity by inviting visitors to develop intersubjective understandings of their differing body-space-movement experiences—the type of architectural learning conversation intended by the

museum. Similarly, the turn-taking supported an iterative, cumulative experience for visitors which reveals a temporal orientation to co-presence as, for example, a guide draws on their memory of the VR scene. The other design choice related to the establishment of co-presence included the large screen projecting the VR scene. This type of display may in fact be commonplace in VR-museum installations and other public learning spaces. However, the above analysis reveals the process in which such a screen mediates co-presence. Note that the screen itself is never truly an object of joint-visual-attention as only one partner at a time has visual access to it. However, it was shown to be an important resource for the guide-role in particular in navigating across the digital and physical spaces. We may recognize that joint attention and co-presence mediate collaborative learning, but in order to apply design perspectives to the development of such mediating resources, we need to continue to investigate the new constellations of digital-physical activity and parallel collaborative interaction.

References

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*(4), 403–436.

Dillenbourg, P., & Traum, D. (2006). Sharing solutions: Persistence and grounding in multimodal collaborative problem solving. *The Journal of the Learning Sciences*, *15*(1), 121-151.

Enyedy, N., Danish, J. A., & DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *International Journal of Computer-Supported Collaborative Learning*, *10*(1), 7-34.

Greeno, J. G. (2006). Learning in activity. In K. Sawyer (ed.) *Handbook of the Learning Sciences* (pp. 79–96), Cambridge, MA: Cambridge University Press.

Goodwin, C. (2007) Environmentally Coupled Gestures. In *Gesture and the Dynamic Dimensions of Language*, edited by Duncan, S., Cassell J, & Levy, E. Amsterdam/Philadelphia: John Benjamins. pp. 195-212.

Heath, C., Lehn, D. V., & Osborne, J. (2005). Interaction and interactives: collaboration and participation with computer-based exhibits. *Public Understanding of Science*, *14*(1), 91-101.

Hall, R., & Stevens, R. (2015). Interaction analysis approaches to knowledge in use. In *Knowledge and Interaction* (pp. 88-124). Routledge.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences, 4,* 39–103.

Kendon,A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge, UK: Cambridge University Press.

Knutson, K., & Crowley, K. (2010). Connecting with art: How families talk about art in a museum setting. In Instructional explanations in the disciplines (pp. 189- 206). Springer US.

Mercer, N., & Wegerif, R. (1999). Is "exploratory talk" productive talk? In K. Littleton & P. Light (Eds.), Learning with computers: Analyzing productive interaction. (pp. 79-101). New York, NY: Routledge.

Nemirovsky, R., Rasmussen, C., Sweeney, G., & Wawro, M. (2012). When the classroom floor becomes the complex plane: Addition and multiplication as ways of bodily navigation. *Journal of the Learning Sciences*, *21*(2), 287-323.

Pallasmaa, J. (2012). *The eyes of the skin: Architecture and the senses*. John Wiley & Sons.

Pierroux, P; Steier, R., & Sauge, B. (2019). Imagining, Designing and Exhibiting Architecture in the Digital Landscape, In Åsa Makitalo; Todd E. Nicewong & Mark Elam (ed.), *Designs for Experimentation and Inquiry. Approaching Learning and Knowing in Digital Transformation*. Routledge. Chapter 5.

Povis, K. T., & Crowley, K. (2015). Family learning in object-based museums: The role of joint attention. *Visitor Studies*, *18*(2), 168-182.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *13*(3), 241-261.

Stahl, G. (2013). Translating Euclid: Creating a human-centered mathematics. Morgan & Claypool Publishers.

Steier, R., & Kersting, M. (2019). Metaimagining and embodied conceptions of spacetime. *Cognition and Instruction*, *37*(2), 145-168.

Steier, R., Pierroux, P., & Krange, I. (2015). Embodied interpretation: Gesture, social interaction, and meaning making in a national art museum. *Learning, Culture and Social Interaction*, *7*, 28-42.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 315-337.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P.J. Dunham (Eds.), Joint attention: Its origins and role in development (pp. 103–130). Hillsdale, NJ: Erlbaum.