International Journal of Research & Method in Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cwse20

Near future practices of interaction analysis: technology-mediated trends in educational research

Jacob Davidsen & Rolf Steier

To cite this article: Jacob Davidsen & Rolf Steier (01 Oct 2024): Near future practices of interaction analysis: technology-mediated trends in educational research, International Journal of Research & Method in Education, DOI: 10.1080/1743727X.2024.2410306

To link to this article: https://doi.org/10.1080/1743727X.2024.2410306

INTERNATIONAL JOURNAL OF RESEARCH & METHOD IN EDUCATION
https://doi.org/10.1080/1743727X.2024.2410306

Near future practices of interaction analysis: technology-mediated trends in educational research

Jacob Davidsen [a] and Rolf Steier [b]

[a]Department of Communication and Psychology, Aalborg University, Aalborg, Denmark; [b]Faculty of Education and International Studies, Oslo Metropolitan University, Oslo, Norway

ABSTRACT
Methodological advancements for the study of learning processes are both shaped by and drivers of technology developments. Interaction Analysis (IA), as a core methodological approach over the past three decades is reflective of this relationship and can be understood by examining moves from analogue video tapes, to digital media, computational and algorithmic technologies, and through recent advances in sensory data and perceptual recordings. It is vital though, that we reflect on the epistemological and ontological implications of these changes as the relationship between researcher, technology, and learning event is changing. In this paper, we identify two techno-methodological trends in education research: computational and immersive approaches, and we discuss what they mean for IA. Computational approaches draw on algorithmic and computational support to identify, capture, and render patterns of learner behaviour found in a variety of interactional data. Immersive approaches on the other hand, invite researchers directly into learning events through technologies like 360° video and virtual reality (360VR). We identify and discuss six implications for these trends and suggest more broadly that researchers should embrace a reflective and entangled relationship between method and technology in their educational research.

ARTICLE HISTORY
Received 26 August 2022
Accepted 27 August 2024

KEYWORDS
Interaction analysis; computer-supported approaches; immersive approaches; educational research; methodology; immersive virtual reality; video-based research

Introduction

Video has long played a crucial role in researching learning in naturally occurring interaction and several research traditions are built on the affordances of video recording technologies (Derry *et al.* 2010, DeLiema *et al.* 2021). Interaction Analysis (IA), as presented by Jordan and Henderson (1995) in particular, has roots in educational research and continues to be one of the most frequently cited methods papers in the learning sciences field. They articulate the 'core commitments' of IA while highlighting the use of video recordings for researching talk, bodily conduct and the use of artefacts in learning processes. Many of the foundations and assumptions of IA are grounded in Conversation Analysis, ethnomethodology, and constructivist theories of learning. In IA, video data allow researchers to study how learning is taking place among participants on a moment-to-moment basis – instead of relying on the participants' reflections of what happened in post-interviews or questionnaires. The focus on the interactional work of the learners has been foregrounded repeatedly since the publication in 1995 (Lindwall and Ivarsson 2011, Furberg *et al.* 2013, Danish and Saleh 2014,

*CONTACT Jacob Davidsen jdavidsen@ikp.aau.dk Department of Communication and Psychology, Aalborg University, Rendsburggade 14, Aalborg, Denmark

Shapiro *et al.* 2017, Bernhard *et al.* 2019) and has served as an argument for studying learning when and where it happens. Of course, looking backwards, much historical work in educational research also took place before Jordan and Henderson's particular framing of IA, (e.g. Erickson (2007) and Mehan (1979)). Jordan and Henderson's paper is a distillation of years of doing IA across different labs, and they point to existing practices and research using video for studying human sociality (e.g. Goodwin (1981) and Heath (1986)). Further, IA is clearly drawing on the socio-cultural tradition of Vygotsky (1978) and Lave and Wenger's (1991) influential work on legitimate peripheral participation. We recognize that video-based approaches to learning interaction are varied but here we focus on the core commitments and assumptions of this specific articulation of IA as it remains such a significant text in the field. Yet since the 1990s, many conditions have changed when it comes to using video data for researching learning processes; recording technologies have become cheaper, smaller, and more mobile; the corpus of data collected has grown to tremendous volumes; and new ways of analysing the recorded data have emerged with the introduction of computational multimodal signal processing of video data (McIlvenny and Davidsen 2017). Additionally, formats of video have expanded from 2D film to 360° video providing new opportunities for analysing learner behaviour with video data. Overall, the past three decades have changed the technological circumstances of Interaction Analytical practices quite dramatically (Steier *et al.* 2019, Davidsen *et al.* 2023), and yet surprisingly little work has been done to revisit and reflect on the core commitments in light of more recent developments in technology.

In reference to the state of the art video technologies, Hall (2000) demonstrated that choices around data recordings are embedded with theoretical stances. The same generative question he posed then is relevant to revisit now:

> By considering how new combinations of text, sound, image, and interactivity can be used to report research findings about learning and teaching, we raise an interesting methodological question. How should we expect new technical media for recording human action, along with a growing suite of tools for their analysis, to shape the specific cultural practices of cognitive and educational researchers? (p. 649)

In reflecting on this question, we see evidence of the ground shifting around IA in the form of new practices and tools referenced in IA publications and in discussions among peers (Steier *et al.* 2019). These changes, which we will explore below, are influenced by developments in technology and demands for more data-driven design and evaluation of pedagogical arrangements in all areas of education.

In this paper, we are informed by speculative (Ross 2017) and future-oriented approaches in which we build on adjacent emerging research practices to anticipate what they will mean for IA. Such approaches can help us develop new methodological practices and allow for 'unintended consequences and emergent properties of technologies in use' (Ross 2017 p. 215). Generative questions and such speculation are useful here because they allow us to make visible and extend the often hidden and unspoken commitments of a method and associated orientations to technology. The goal is not to redefine IA nor is it to pose alternative video-based methods. Rather, we seek to anticipate emerging practices of IA that can adopt or extend its core commitments. The main audience for this paper is thus researchers who use and are interested in IA and related methods, as well as technology-methodology entanglements. This work is necessary to bring focus to the foundational assumptions of IA, and to bring consistency in addressing future applications of IA with novel technologies. A key part of this argument is that researchers should not let technology dictate their methodological choices, but instead pursue and recognize an entangled relationship between method, analysis, and technology. That makes the assumptions and core commitments essential to have in mind when discussing the future practices of IA in educational research. The question driving the paper is *What are the near-future practices of interaction analysis?*

This text is not a traditional review, but a 'position paper' from within the field of educational research using IA to study learning processes. We draw on recent literature from an expanded corpus of video-based techniques and methods in educational research. We are also informed by our own experiences teaching IA, participation experiences in IA and educational sciences communities, as well as a variety of project contexts in which IA methods have been applied. Through our identification and engagement with two techno-methodological trends, we aim to raise questions that can guide the future development of interactional analytic approaches.

First, we revisit the core commitments of IA. We then identify and explore two emerging techno-methodological trends in educational research: Computational and Immersive approaches. Finally, based on these trends and in relation to the core commitments of IA, we discuss six implications for the future of IA methods.

Core commitments and assumptions of interaction analysis

Jordan and Henderson present IA as an interdisciplinary method that supports the collection, analysis and dissemination of learning events using video data. Though IA, as noted by Hall and Stevens, is 'neither a singular tradition nor a prescribed, unitary set of methods' (2015, p. 100), a few core commitments are especially important in the context of identifying near-future practices (Hall and Stevens 2015). A central commitment of this approach is that video recordings give researchers the opportunity and flexibility to revisit scenes, to replay and pause sequences of interaction, and to share these sequences with others (Jordan and Henderson 1995, Derry *et al.* 2010, Heath *et al.* 2010). Repeated viewings also allow researchers to iteratively construct interpretations of these learning events by focusing on different aspects and elements that may be missed or invisible at the moment. This is often a practice linked to a specific research laboratory, a practice which is rarely documented in published studies.

Jordan and Henderson offer another commitment on the status of knowledge in IA: 'One basic underlying assumption in Interaction Analysis is that knowledge and action are fundamentally social in origin, organization, and use, and are situated in particular social and material ecologies' (Jordan and Henderson 1995, p. 41). This is an explicitly relational epistemology and ontology which says something about what is actually being documented in video data – that knowing and being in the world can be seen emerging in activity with others. In the years surrounding the emergence of IA, a shared interest in 'learning in interaction' emerged, emphasising that learning is found in the ways participants interact with each other and with the material ecologies they produce and reproduce in a shared activity (Goodwin and Goodwin 1996, Goodwin 2000). By adopting IA as a method, one is thus also implicitly adopting the stance that social interaction between participants is a (the) site for human learning. Learning *is* the interaction (rather than merely evidenced by it) and being in the world depends on others. This perspective contrasts with more cognitively oriented epistemologies in which interactions tend to be treated as reflections of individual cognitive structures and learning (diSessa *et al.* 2015, Hall and Stevens 2015). Jordan and Henderson argued that the interest in social interactional aspects of learning has 'methodological consequences' (p. 41) and critiqued protocol and survey interview data for reducing the complexity of the learning process. Video plays a unique role in this regard as it makes visible the relationship between tools and signs and their socially emergent meanings (Krange and Ludvigsen 2009). In short, this first assumption links the choice of IA to the view of knowing and being in the world as social.

Extending these ideas, a related commitment is articulated as: 'Analytic work, then, draws, at least in part, on our experience and expertise as competent members of ongoing social systems and functioning communities of practice' (Jordan and Henderson 1995, p. 41). IA is by no means a neutral method and researchers' experiential and 'human' intuition and expertise are central to the analysis. Interactional data must be presented in such a way that this expertise can become relevant. In other words, when data material is abstracted, it becomes more difficult for researchers to place themselves in the shoes of participants or make relevant their own everyday understandings of human interaction. For this reason, analytic work is typically grounded in video and transcripts rather than synthesized representations. The original text presents classic sequential text-based transcripts and different multimodal transcripts. The interest in transcription formats and procedures has occupied researchers ever since, including experiments with comic transcripts (Norris 2004), multimodal transcripts (Mondada 2007) and chronotopic transcripts (McIlvenny 2014). The first analysis of the interaction is actually performed by the participants themselves in the moment of activity (i.e. in the video material); participants show their understanding of the situation through their multimodal interactions, which can guide the researcher's analysis of the activity. This type of perspective-taking work by researchers should not be pure speculation, but is still a valid means of interpreting video data in IA. In this respect, IA diverges from other video-based traditions such as CA where internal states of participants are explicitly not objects of inquiry (DeLiema *et al.* 2021). In IA, intentionality and other thinking processes can be discussed through direct reference to perceivable features of the videotape (Jordan and Henderson 1995). With the coupling of the theoretical and methodological perspectives, IA is not a narrow procedural sequence, but a practice that has a strong theoretical and methodological alignment with a way of doing research on learning and social interaction.

Finally, IA is also viewed as a collaborative process emphasising that analysis gains validity when groups of researchers work together to compare and co-develop interpretations of events. In Nivala (2012) an international group of researchers collaboratively applied IA on their audio-visual data using networked technologies emphasising the importance of such data sessions. Perhaps more importantly there is also an opportunity to invite participants into the research process through shared reflection around video material (Shapiro *et al.* 2017).

Based on the core commitments of IA, Jordan and Henderson developed what they refer to as foci for analysis. The foci are not meant as a strict set of guidelines to follow when doing interaction analysis, but as ways of approaching the video of the event. The named foci are: the structure of events, the temporal organization of activity, turn-taking, participation structures, trouble and repair, and artifacts and documents. For researchers using IA the foci have provided a lens for inspecting video data – for example sequentiality and turn-taking (Silseth 2012), the use of artifacts (Rystedt and Lindwall 2004) and participations structures (White 2018). Together, they give a sense of the relevant frames for social interaction inherent to the method.

To lay the groundwork for the arguments that follow, we summarize these commitments and assumptions of IA: (1) The affordances of video shape much of the analytic work in IA; (2) Learning is fundamentally social, rooted in interaction, and IA is based on explicitly relational epistemology rather an individual achievement which happens to occur in a social context; (3) Members' perspectives and researcher intuition are seen as valid lenses for interpreting data; (4) IA is a collaborative activity developed by researchers working together. Next, we turn to discussing two larger methodological trends and what they might mean for IA.

Identifying the computational and immersive trends

Drawing on recent literature and observations of the field of video-based learning research, we identify two trends as a basis for our speculative analysis. The two trends are not necessarily presently visible in IA studies but are viewed as possible near-future trajectories to be reflected on. This work is thus not pointing towards a coherent version 2.0 of IA, but to an expanded horizon.

On the one hand, we see more computational approaches to studying learning and learning processes more generally (Worsley *et al.* 2021) – distributing much of the analytical endeavour to algorithms and Artificial Intelligence (AI). Especially with the growth of AI and the introduction of Big Data in educational research there is a growing demand for developing technology that can 'clean,' systematize and analyse learning behaviour in ways that humans may not be capable of. We note that even during the time writing this paper, the possibilities with and public discourse around AI are moving so rapidly that it is difficult to keep up. However, we now see for example, how body posture and movement can be detected and analysed using computer vision techniques (Cao *et al.* 2019) supporting Multimodal Learning Analytics (MMLA) (Martinez-Maldonado *et al.* 2020, Vieira *et al.* 2021, Worsley *et al.* 2021). From a practical point of view, there is also a demand for enhancing automatic analytics of learner behaviour – as the collection of video data (and other types of data) has exploded in the last decade. Researchers are simply able to capture and store more and more interactional material in higher quality than we were previously (Ferguson *et al.* 2019). Similarly, web-based and digitally mediated learning interactions produce a plethora of data from timestamps and clicks to transcripts, text-logs, and eye-gaze records. The emphasis on human perceptual and analytic processes, deemed essential in Jordan and Henderson's articulation are perhaps pushed out of the computational trend and yet we cannot ignore what these developments might mean for analytic practices with IA. Further, as noted by Baker and Hawn (2022) it is crucial to understand (and avoid) algorithmic bias in educational research as the consequences are unknown and unpredictable.

On the other hand, we also see a second divergent or perhaps complimentary trend towards more immersive methods for studying learning interactions (Davidsen *et al.* 2023). These approaches position the researcher in the scene (again) using tools like 360° video and Virtual Reality (VR) technology (360VR). Such approaches represent perhaps the state of the art of what have historically been referred to as video-based and ethnomethodological entryways (Garfinkel 2002) into the analysis of unfolding learning activity (Zahn *et al.* 2021). McIlvenny (2020b) presents the idea of Immersive Qualitative Analytics inspired by immersive theatre and media consumption studies. This emerging idea is fuelled by the introduction of 360° cameras and Head Mounted Displays (HMD) for VR though it might soon extend to other kinds of first-person sensory technologies. With Immersive Qualitative Analytics, researchers are no longer bounded by a square frame, instead a scenographic space can be inhabited individually or together with others – which allows researchers and practitioners to re-activate the interactional scenography again and again (McIlvenny 2020b; Vatanen *et al.* 2022). These ideas raise interesting epistemological questions about presence and perspective-taking in data and provide an important complement to computational approaches in video-based educational research.

The computational trend

The central idea of the computational trend is to explore how features of human social interaction ('foci') can be interpreted in new ways through algorithmic and datafication techniques (Blikstein 2013, Worsley *et al.* 2021). In several areas, such computational techniques and methods are addressing classical interaction analytical interests, like gaze, movement, gesture, turn-taking patterns, pauses, and body posture, but the analytical work is now distributed between researchers and computational agents. With increasing volumes of data, new sensors for data registration, and a focus on quantifying learner interaction, we see different algorithms and software packages that can track verbal and bodily conduct in video recorded data (e.g. open-source software packages like OpenPose, MediaPipe, AlphaPose, FrankMoCap, OpenAI Whisper, etc.). We are focusing on computer-supported techniques and methods that are particularly concerned with human social interaction, (e.g. *not* Social Network Analysis or text mining). We emphasize that though most studies referred to in this section do not explicitly link their work to IA, identifying future practices of IA requires a broad approach to new methods concerned with learning interactions. The computational trend is an umbrella term for these approaches, though we are aware that this is not a label used by the researchers themselves.

The computational trend in educational research includes, for example, Learning Analytics (LA) and Multimodal Learning Analytics (MMLA), etc. In common is an interest in exploring and conceptualising how computers can support the analysis of learning activities based on different types of 'data in high-frequency (video, logs, audio, gestures, biosensors)' (Blikstein 2013, p. 105). LA originally dealt with the digital traces that students generate in online environments (e.g. keystrokes, mouse clicks and text messages), but more recently MMLA has been applied in laboratory settings (and gradually classroom settings) full of multimodal social interaction, which requires new ways of capturing and analysing learner behaviour. Capturing and analysing are overlapping processes, but the former begins with learner interactions in the respective environments and interfaces whereas the latter is rooted in researcher activities with data sets after the interaction. Blikstein and Worsley

(2016, p. 233) presents MMLA as 'a set of techniques employing multiple sources of data (video, logs, text, artifacts, audio, gestures, biosensors) to examine learning in realistic, ecologically valid, social, mixed-media learning environments'. According to Mangaroska *et al.* (2020) MMLA focuses on actual learner behaviour instead of focusing on 'recalled behaviour or subjective impressions as self-reported data does' (p.80). Further, Worsley *et al.* (2021) outline twelve commitments for MMLA which align with some of the core commitments of IA; make the invisible visible, focus on the material ecologies in the social interaction, and present students' interactional work across modalities. This is interesting as it aligns with the IA aim of capturing (on video) learning interactions as they unfold between participants in a face-to-face setting and not in post-tests or interviews. This epistemological stance found in MMLA aligns with Jordan and Henderson's commitment to the situated and material ecologies learners produce and reproduce in their social interaction. Whereas IA traditionally focuses on how the participants understand each other and the learning activity, the computational trend is adding measurements and insights that are not presently available for learners in the practice (though perhaps new methods or tools, like participant dashboards, will make learner interpretations relevant in-the-moment).

A common feature of these computational approaches is the creation of visual outputs (for the researcher, currently) based on computationally intensive analysis of data. Human behaviour is translated into graphs and curves after careful implementation and training of computational systems (Blikstein 2013). As with any technology or algorithm this is of course not a neutral process (Baker and Hawn 2022), and assumptions and theories about what counts as learning, collaboration, and negotiation of meaning are 'coded' into the transformation of the original data. The relation between theory and data has been heavily debated by educational researchers in relation to Big Data (Martinez-Maldonado *et al.* 2021, p. 130) as some fear that LA will make theory unnecessary (Mazzocchi 2015), while others advocate that LA is making theory more relevant than ever in order to interpret the computational output as professionals (Wise and Shaffer 2015).

Gaze has long been in interest in IA as an indicator of attention in interaction, and computational studies using eye tracking equipment have similarly worked with joint visual attention as a measurement for collaboration. Recently Sharma *et al.* (2021) critiqued this use of joint visual attention as a shared 'view' does not necessarily say much about how students are engaged in a collaborative task. In a study by Spikol *et al.* (2018), they counted the number of faces looking at a screen as well as the distance between them as a way of quantifying collaboration. Crescenzi-Lanna (2020) challenge this measure as a 'proxy for collaboration' (p. 1494) rather than a direct indicator. This is an interesting development as the original event is turned into another analytical object, which may be more generalizable, but perhaps less ecologically sensitive. Of course, IA researchers have turned the event recorded on video into transcripts, but the analytical object is still the video and the transcript of the event – not a graph or set of numbers.

Overall, such computational approaches collectively are in a transition and there is an incentive to expand the focus from clicks and keystrokes on online platforms (Martinez-Maldonado *et al.* 2021, p. 127) to multimodal arrangements of learning (Blikstein and Worsley 2016). More broadly, the interest in analysing social interaction is also gaining momentum in studies of embodied and multimodal interaction. For example, Trujillo and Holler (2021) used cameras, microphones and Microsoft Kinects to capture and analyse the timing of body posture related to different types of questions. The layers of data provide unique access to study social interaction beyond talk. Likewise, Nota *et al.* (2021) used the transcription and annotation software ELAN to manually score a dataset which was afterwards analysed using R and other software packages to provide evidence of the relation between facial signals and social actions in multimodal face-to-face interaction (2021).

To summarize, the computational trend is changing how educational researchers collect, systematize and analyse data of learner behaviour (Blikstein and Worsley 2016). The computational trend is informed by computational advancements making it possible to work with big data sets and scrutinize talk and bodily conduct with different AI's and machine learning techniques. As noted earlier the computational trend is rapidly developing and difficult to keep track of, but in the following, we identify three implications of this trend in educational research.

Three implications of the computational trend for IA

Though computational approaches do not necessarily share the same core commitments as IA (for example notions of member perspectives and relational epistemologies are not necessarily present in such approaches), we can still see that technological and methodological advancements will surely inform what IA looks like in the near future. Here we identify 3 specific implications of computational approaches on IA methods. We recognize that these are speculative, but it is still crucial that we consider these implications now in the changing research landscape rather than retroactively.

1. Computational techniques (will) assist and expand researchers' early analytic work in identifying hotspots, sorting, and reducing video data for IA leading to new researcher-data relationships.

The development in recording technologies has dramatically expanded the possibility of collecting and archiving big video data sets, but there is a need to develop and implement new computational techniques for identifying interactional hotspots and for sorting and reducing video data for IA. First, it is impossible for researchers to review and systematize large data sets by hand due limitations in time and energy. Second, from an ethical perspective and as stated by Erickson (1985), researchers are obliged to work with the data we have collected as we have 'disturbed' the practice and participants and their contribution needs to be respected. Aided by visual image processing algorithms and machine learning techniques a researcher will be able to identify hotspots in the video data: e.g. in classroom data one could identify all the scenes where the teacher touches the interactive whiteboard, or one could single out all scenes from collaborative activities where the students use certain objects or point with their index finger at each other. Furthermore, machine learning techniques such as deep learning and computer vision can be applied to the video data to automatically classify and sort it. For example, researchers could train a machine learning model to identify different types of classroom activities or student interactions based on visual cues in the video data. This can help to reduce the amount of manual effort required to analyse the data and can enable researchers to identify important patterns and trends that may not be immediately apparent from manual analysis. A recent exploration of such analytic techniques is found in Fonteles *et al.* (2024), with a multimodal timeline tool that advances IA in mixed-reality learning by integrating machine learning for data visualization, enabling a shift from qualitative to mixed-methods approaches.

It is clear that the technology is re-shaping the analytic process and as noted by Worsley *et al.* (2021) algorithmic labels of human behaviour are not who they are as people and computational methods require a deep sensitivity of the context and human social practice. One issue that we as a field will need to consider involves the blurred boundary between the supposedly neutral organization of interactional data (e.g. historically, placing tapes in filing cabinets) and the more interpretive modes of sorting and selection. If a core commitment of IA is that it depends on researchers' (human) ways of knowing and making sense of social systems, at what point is this epistemological premise violated with these computational techniques?

2. Computational techniques (and AI) will radically transform transcription practices.

Different transcription services are already producing transcripts of talk in 'clear' environments with one or two persons, however, more complex settings like classrooms or makerspaces will produce audio-data with more 'noise/disturbance', which poses some difficulties for automated transcription (Umair *et al.* 2022). Still, the rough AI-produced transcripts require careful review and revision from the researchers afterwards. Perhaps new models of AI-based transcription can cater to the needs of IA kinds of transcription. This includes organising the sequential organization of turns, overlaps between participants and intonation. The maturity of AI transcription is not yet suitable for Interaction Analytic studies of learning processes, but there is promising software in the making, like GailBot (Umair *et al.* 2022). Gailbot is tailored towards producing AI-based transcription for Conversation Analysis following the Jeffersonian transcript annotation system (Jefferson 2004). Furthermore, it is important to remember that IA is treating the process of transcription as a crucial analytical step – this is where analytical findings are first developed as the researcher is becoming familiarized with the data. As aspects of transcription work do become automated, it is important to consider whether certain kinds of insights may be lost or how to maintain or enhance the analytic steps inherent to transcription (Ayass 2015). Engaging with the raw data prior to creating a representation of it (i.e. transcript) is very different from engaging with the transcript as a proxy for the data as an initial engagement. Contextual nuances such as tone, accent, attention, background noise or movement, etc may be missed. This has implications for what the researcher chooses to pay attention to and what kinds of interactional patterns become highlighted. On a productive note, AI-based transcription should quickly allow flexibility in changes to what is included and excluded from the transcript (e.g. add pauses, add overlaps, remove intonation, translations, etc). Therefore some of the manual and time-consuming work for researchers could eventually be minimized. On the whole, new transcription routines that preserve transparency and flexibility will need to be considered and developed systematically *by* educational researchers.

3. Systematic pattern recognition and visualizations of large audio-video (and multimodal) datasets will be used alongside more traditional Interaction Analysis techniques challenging (or complimenting) researcher interpretation.

Just as artificial intelligence techniques have exploded in terms of impact on the analysis of texts (Hassani *et al.* 2020) and images (Wang *et al.* 2019) we suspect that there will be new and unexpected consequences as digital technology's role shifts from the organization and visualization of (video) data to actually performing the interpretive analytic work. We know that AI can now perform what we typically think of as 'creative' work, such as storytelling (Sharples and Pérez y Pérez 2022), so the next question becomes what will computers contribute when they *see* things in interaction that we as humans do not? It's not unreasonable to think that in the near future we could ask our computer to look at a recording of a classroom for example and ask, 'what is the relationship between turn-taking patterns in student talk and concept use?' or 'How do movement patterns relate to resource use?' It's important that we anticipate what this means for the relationship between the researcher and the learning interaction event. Today, we often think of the video, perhaps together with the transcript, as being analysed by the researcher in a relatively unmediated way. New kinds of representations, interfaces and controls come between that data-researcher relationship, serving as a kind of proxy embedded with other assumptions and distortions. These layers of mediation can lead to researchers asking new kinds of questions which are more suited to this kind of analysis. A hypothetical illustrative example might involve a researcher interested in classroom movement patterns. If the analysis of such movement shifts from being rooted 'in the video' to AI generated map-like representations with learners depicted as dots or avatars – then notions of embodiment, bodily expression, and bodily identity could easily become obscured or lost. Critical reflection will be required to ensure that such additional interpretive layers are consistent with the epistemological stance that learning unfolds through social interaction, or in this hypothetical case, is embodied. In addition, it is of vital importance that AI is used in an ethical way and that IA researchers inform the algorithms.

These three implications are emerging from a computational trend in education research which does not necessarily share the same relational epistemology and ontology as the version of IA articulated by Jordan and Henderson. As noted above, Worsley *et al.*, (2021) do see a mutual interest between IA and MMLA in the social and material ecologies of learners. Perhaps the biggest tension though, when looking across these implications has to do with when in the research process the computational support comes into play, and the extent to which the computational support distances the relationship between the researcher and the actual learning interaction. For example, large scale data sorting techniques can probably be adapted to IA in ways that are consistent with its core commitments as long as the sorting is transparent, preserves opportunities for perspective-taking in the subsequent analysis. Transcription work poses some challenges though as there is the possibility that some kinds of analytic insight might be lost if researchers are removed from the transcription process. When it comes to pattern recognition techniques of interactional data, there are serious tensions that emerge. It is difficult to see how abstracted and automated pattern recognition can reliably maintain member perspectives and such processes would need to be explicitly designed to highlight the relational epistemology of learning processes to be consistent with the IA method.

The immersive trend

Immersive approaches to analysing interaction are about inhabiting, sensing and experiencing the learning event (McIlvenny 2018, Davidsen et al 2023, McIlvenny and Davidsen 2023) – thereby getting access to how social interaction unfolds between the bodies and material ecologies in practice. Instead of loading interpretive work into visual graphs or statistics, immersive approaches are explicitly staying with the event by zooming in on the interactional work of the participants and through empathising with the participants in new and valid ways. In relation to immersive approach for analysing interaction, it is important to clarify the concept of immersion, which has been used both to characterize a human sense of being in some alternate reality related to 'presence' (Nilsson *et al.* 2016), as well as a technology's capacity to mediate the alternate experience (Oh *et al.* 2018). We adopt the former perspective and consider immersion to involve a sense of being in the scene or to be experiencing the learning interaction as though one were really there. Immersion in itself is not inherently technology dependent and researchers performing IA in collaborative group settings often experience a kind of immersive engagement with the learning event or with the learning activity (Jordan and Henderson 1995). What then distinguishes IA from this immersive trend? First, there is not a binary distinction between IA with flat video recordings and these immersive developments. We have been on a gradual trajectory moving from single tripod-mounted cameras to multi-camera and microphone arrangements, to mobile, wearable, and first-person (Pink 2015) recordings, to 360° video, and finally to the immersive experience of 360VR. Recordings of learning interactions have been becoming more immersive and it is important that we step back to see what assumptions might be challenged and what new practices emerge. Second, the key yet perhaps subtle distinction we wish to raise is that the shift from analysing a static video on a flat screen is qualitatively different from placing oneself in a constellation of digital media that actively recreates aspects of the original learning interaction (McIlvenny 2020a; Davidsen *et al.* 2023). There is a powerful sense of being transported to an alternate place and time that is difficult to characterize in text (see Vatanen *et al.* (2022) for different examples). The key distinction is in the research frame; in immersive approaches technology and data arrangements are intentionally designed to recreate a sense of presence in the event rather than as documentation for reflection and post-event analytic work. We view this movement towards immersive forms of analysis as a trend with implications for IA, which can influence and enhance the core commitments and assumptions of IA. Influenced by recent technological developments like 360VR, emerging practices of Immersive Qualitative Analytics (McIlvenny 2020b, Davidsen *et al.* 2023) and Collaborative Immersive Qualitative Analysis (Davidsen and McIlvenny 2022) emphasize human entryways for dealing with new interactional material. Such data and interpretation are characterized by Zahn *et al.* (2021, p. 665) through the potential to be 'holistic and environmentally sensitive (…) enable(ing) a holistic immersion in a collaborative learning situation even after the capture of the event'. Thus, we return to and reemphasize the core commitment of IA that *learning is interaction.* Much of this thinking is not new here – the DIVER-project at Stanford in the beginning of the century had similar aspirations and used panoramic cameras to capture learning events for later analysis in flexible ways (Pea *et al.* 2004). Similarly, Ferguson *et al.* (2019) described how a process of working with video data involving immersive techniques allowed familiarization with the data and learning event.

While video historically has provided access to study learning processes in detail, it is also clear that a camera in a specific position is only allowing one to see exactly what is captured in that specific 2D frame. Of course, researchers have also used mobile cameras, where the cameraperson is making decisions about what to record as the participants go about doing what they are doing (Mondada *et al.* 2024). Also, researchers have used complex recording setups to generate composite videos showing simultaneous activities in a classroom (and many other settings), but these composite videos are often difficult to make sense of for fellow researchers (at least it will take some time to get acquainted with the specific composition of the videos). Previous projects have set out to tackle this problem by developing software and methods for dealing with video, sometimes through a curatorial process of data reduction e.g. DIVER (Pea *et al.* 2004). More recently, software for enhanced transcription practices, such as DOTE (McIlvenny *et al.* 2022), and curational and analytic software like DOTEbase (McIlvenny *et al.* 2024) have been developed to tackle large audio-visual datasets. We also see new genres of representation for IA being developed, e.g. turn charts (Shapiro *et al.* 2024). An additional challenge is that it can be difficult for teachers (or other educational professionals) to engage with this complex composite video if these representations are not part of their normal practices. With the introduction of 360° cameras and 360VR, it is now possible to move beyond the flat 2D representation of video. While 360° video can be played using a traditional media player it is also clear that this way of working does not really take advantage of this new format. Here McIlvenny (2018 2020b) points towards the need for more immersive ways of working with 360° video using 360VR (see Vatanen *et al.* (2022) for a number of examples). VR is traditionally used for facilitating learning processes and not as an immersive research environment enabled by 360° video and HMDs (Davidsen *et al.* 2023). VR in many ways supports a higher level of accessibility and flexibility for learners, as they can visit remote places or experience dangerous situations in programmed and simulated worlds (Pirker and Dengel 2021). That is interesting to study as a learning activity in itself, but for research, the combination of 360° recordings from a classroom (or other educational settings) and the immersive power of VR creates new possibilities for the study of learner interaction. The devotion to exploring new technologies for researching learning is well-established through the core commitments of IA and for us, it is important to critically and constructively engage in this conversation. Video tape was written extensively about as a technology with particular affordances in the 1990s, yet we have not seen similar reflections in more recent work despite the fact that 'video' is not even really the same technology today.

Data collection techniques will evolve as the immersive trend develops, but in general one could consider the move as involving a change from documentation for subsequent analysis towards documentation for subsequent inhabiting. Choices of, for example, camera placement, have implications for how the event is later experienced. Here, researchers need to think carefully about how the activity can be captured and later inhabited in a 'holistic and environmentally sensitive way' (Zahn *et al.* 2021, p. 65). A camera placed on a table will later 'feel like' one embodying an object on the table and a head-mounted camera may recreate first-person immersion. This is not to say that more cameras necessarily are better, but that one should consider research questions closely in relation to the recording setup. Davidsen and McIlvenny (2022) presented an example of an immersive analysis from architecture education with 6 students working in pairs or individually in different places across a workshop area – each of them occupied with different tasks. The 360° cameras were positioned in strategic places to allow researchers to obtain different perspectives in the room in the subsequent analysis. In the immersive research environment, the researchers can switch between different camera positions in a flexible and sensible way and thereby inhabit the scene from different positions to inspect the way the group worked together. In one situation, it is clear that one of the pairs encounters trouble as their digital model does not align with the physical model being assembled using foam pieces. Without any apparent talk or contact, another student from across the room notices the problem and approaches. But why? How are such moments of trouble made visible to other students (or teachers)? Through an interactive process of inhabiting the different 360° camera positions, and feeling present in the room, the researchers started to notice the subtle cues that become relevant for the participants in identifying breakdowns. Working in the immersive research environment allowed researchers to intuitively transition between different cameras and audio sources while making sense of the interaction.

Overall, the immersive trend stands out as emerging and speculative, but it is also clear that this trend is an important contrast or complement to the computational trend. In the following, we raise three implications of the immersive trend for IA.

Three implications of the immersive trend for IA

Experiencing a 360VR provides a dramatically different perceptual experience than watching a flat video recording due to the level of immersion. There are different aspects of immersion which might rely, for example, on sensory perceptual features of the environment but also narrative or socio-relational orientation to the recorded event (Enyedy and Yoon 2021). This sense of being in the learning event has important ontological and methodological implications for IA work and the point we wish to expand on is that new analytic resources become available when one inhabits the scene.

1. Experiential aspects of interpretation such as empathy and first-person perspectives will be strengthened and foregrounded.

One implication of the immersive trend involves the status of empathy and perspective-taking in relation to the analytic work. VR has been referred to as the 'ultimate empathy machine' and experimental research has now demonstrated that VR may increase certain measures of empathy and perspective-taking of participants (Herrera *et al.* 2018). These features involve either seeing the world through the eyes of others or else involve the level of sensory engagement afforded by (I)VR. VR has also been shown to elicit stronger emotional responses than traditional film (Ding *et al.* 2018). For example, Curran *et al.* (2022) showed that 360VR is enhancing the feeling of presence in a practice for the VR participants and that the students experienced a higher level of agency as they can freely look at what they want in the 360° sphere. Further, programmed virtual experiences have repeatedly showed that immersion foster participants emphasising and perspective-taking in VR (Shin 2018, Young *et al.* 2022). This indicates that across VR activities emphasising and perspective-taking can take place whether you are with an avatar or recordings of real persons. At the same time, it is reasonable to be sceptical about precisely how powerful these empathetic responses to immersive environments are. The existing research tends to be based on post-experience surveys (Shin 2018, Young *et al.* 2022) which don't capture the nuance of what this kind of empathy might look like or how perspective-taking manifests in the interaction. Our point is not to conclusively say that the immersive trend directly increases the empathy of researchers towards participants, but rather to call attention to the possibility of empathetic responses influencing researchers' analysis of interactional data, which follows the core assumption of IA about the centrality of the researchers' experiential and 'human' intuition and expertise for the analysis.

Empathy gains relevance as you are immersed into the scene, but also because you are able to interact with the video data (and subjects) in a flexible and intuitive way. As noted by Pink (2017) researchers should carefully consider how video can be used for 'empathetic and negotiable forms of imaginative encounters with other people's experience' (2017, p. 379, for example in relation to marginalized voices). This line of reasoning about video as a technology for empathy is present in Harris' (2016) ethno-cinema and discussed by Krishnamoorthy *et al.* (2021), where video is also used as a means of getting inside in other people's lives, and not solely as means for capturing and analysing the sequential structure of a learning event.

These properties invite us to revisit commitments and assumptions about how IA should treat the internal states and intentionality of participants. Earlier we noted that IA may differ from Conversation Analysis in that one may infer kinds of intentionality and thought of participants in the video recording if one is able to point to evidence 'on the tape'. But what does it mean if 'the tape' reliably mediates empathising with participants or adopting their perspectives in new ways? This is an epistemological question as much as it is a methodological one. It also suggests a risk that a sense of immersion could give researchers a misguided confidence in the validity of the interpretation of the event. After all, feeling as though one is in the scene is definitely not the same as being a participant in the actual interaction.

In light of this immersive trend in educational research, new IA procedures need to be developed to document and make transparent where these kinds of perspective-taking responses come from. Experiencing and empathising can of course happen using traditional video, but with the immersive approach, it seems to be further ignited.

2. The embodied agency of the researcher is strongly emphasized when performing IA in immersive settings.

There is also an embodied component to this immersive analytic work. Being present in the scene of the learning event allows researchers to rely on bodily understandings to perform the analysis. This might be as simple as making choices about orientation and analytic attention based on bodily movement (e.g. turning one's head to shift attention rather than dragging a cursor on a screen, or moving forward rather than zooming in). Of course, you are not taking part in the actual recorded event, but immersive approaches allow a more subtle way of engaging with the participants ways of establishing and renewing an embodied participation framework. The possibility of being able to reposition yourself in the scene can provide access to multiple perspectives and to how participation frameworks are negotiated between the participants. Changing a camera view or moving to a new location in the immersive scene leads to the construction of an alternative first-person perspective in-the-scene by the researcher. The point is that the body of the researcher, through orientation, movement, attention, and sense of presence are lifted up in the analytic process. However, some people can get motion-sickness wearing an HMD, and of course, accessibility to technology is also an issue that must be considered carefully in relation to the immersive trend.

Most of the attention presently seems to be on visual representations of interaction, and to some extent auditory (e.g. location-based audio recordings merged with multiple immersive video sources). We would be naive to not anticipate that other forms of perceptual or experiential and embodied data such as touch (Price *et al.* 2021) or perhaps physiological or emotional recordings could become commonplace. As bodily aspects of experience become documented in interaction, the work of doing interaction analysis will also become more embodied.

3. Collaboration opportunities for doing IA will expand.

Collaborative virtual environments will radically alter what we think of as the IA research lab. One of the core commitments of IA includes the collaborative and social aspect of doing analysis together with other researchers in a lab setting, but this work has tended to be purely co-present or else asynchronous with researchers taking/sending analysis or interactions back and forth in dialogue (Nivala *et al.* 2012). Immersive techniques will expand the possibilities for collaborative analysis as researchers remotely can already now co-inhabit the 360° video data in collaborative 360VR environments (Davidsen and McIlvenny 2022, Davidsen et al 2023).

This immersive trend will not only invite opportunities for perspective-taking with the participants recorded on video, but also with co-analysts that may interpret and analyse the interaction from another stance (McIlvenny 2020a; Vatanen *et al.* 2022). This will also allow for more sophisticated ways of working together with practitioners from practice – as both researchers and practitioners can inhabit the learning event again, together. Practitioners unfamiliar with interaction analysis techniques (e.g. transcript conventions or approaches to 'hot spots') may find a more welcoming entryway to providing analytic through immersive environments. We have seen a parallel in architectural design where architects use VR models to share ideas with clients who may be unfamiliar with reading and interpreting the conventions of architectural drawings (Pierroux *et al.* 2019). Architects are skilled in using drawings to imagine possible spaces while clients may need more support.

Collaborative IA work in an historical sense can be thought of in terms of as guided by switches in attention between the video recording and the outside analytic conversation. Researchers sitting around a shared screen in one moment attend to a video of an interaction and then in the next move to a discussion between peers to jointly interpret the interaction. Such attentional shifts preserve a distance between the data and the analysis. And yet collaborative immersive techniques may allow them to occupy a kind of simultaneous dual presence in the scene, experiencing and analysing the event together. We are excited to see what these new forms of collaborative IA work might look like, perhaps in immersive online analysis sessions, or purely distance-based IA labs. Physically co-present analysis activities may, however, be better served by preserving non-immersive forms of analysis to take advantage of the affordances of co-presence. New opportunities for collaborative work in immersive spaces should not be seen as replacing current practices but as a window of opportunities for IA researchers. There is an evident need for better structures for collaborative IA work, which is expressed in detail in a recent symposium (Love *et al.* 2024), where groups of researchers described how they performed 'Caring relations across Interaction Analysis labs'.

To summarize, interpreting any recorded learning event is in many ways an imaginative act. Making sense of the interactional moves of participants requires that researchers place themselves in an alternative time and place, in the setting of the participants, and that they adopt hypothetical points of view (Pierroux *et al.* 2019). Similarly, when we co-analyse interactional data with the participants, they are imagining the conditions of their previous experiences to gain reflective insight into their own learning processes. With the immersive trend, this imaginative work is actively facilitated by tools and perceptual resources mediating such spatial and temporal work.

Conclusion

To summarize our overall argument in this paper, we began with the observation of two technology-mediated trends in educational research that may inform future versions of video-based interaction analysis. The first, what we refer to as the computational trend, can be described as the application of computational techniques to data of interactional phenomena using a plethora of different data streams (Blikstein 2013). These are phenomena and learning mechanisms which have long been of central interest to IA researchers such as language use, movement patterns and the use of artefacts. The new computational approaches developed in adjacent fields like learning analytics, may not be grounded in the same (relational) epistemological and ontological premises of IA. Yet specific techniques involving the handling of large amounts of data, the transcription and representation of learning events, and the sorting and recognition of patterns and hotspots may have significant implications in the coming years for IA practices. Further, within the field of MMLA, there is a historical interest in understanding learning as it happens in practice (Worsley *et al.* 2021) – how students use talk and material ecologies to make sense and learn – which aligns with the epistemic and ontological attitude of IA. While there are overlaps between the epistemic and ontological attitudes it is also clear that the process towards understanding learning in interaction can be quite different.

We also see a second trend, the immersive trend, leading towards richer, perceptual, and presence-oriented documentation of learning events, which increases researchers' sense of being in the learning event. Implications for the future of IA from this trend include more emphasis on perspective-taking and empathy, an increasing relevance of the embodied practices of the researcher, and new opportunities for collaborative forms of analysis. This immersive trend more broadly also has implications for the relationship of the researcher to the learning event and needs to be carefully considered. For example, if the sense of presence in a learning event is strengthened and becomes more personal, does this have implications for reliability of interpretations across researchers? We suspect that new practices for data transparency will need to be developed, beyond linking to or attaching raw data. And at a practical level, will the navigation between a growing number of multimodal and interactional data require its own set of skills and practices that we are not anticipating? Perhaps such skills will naturally be appropriated from everyday (non-research) experiences with new technologies. The immersive trend is perhaps not as defined and mature compared to the computational trend but just as significant for future studies of learning.

Together, these are two technology driven trends which will inform new IA practices and hopefully new IA technological infrastructures for studying learning. We recognize that many of these implications we have identified are in a sense speculative as we don't know precisely what the future will bring, but our larger goal is to lift up these questions so that they are taken seriously by researchers and practitioners of IA and so that we are intentional in developing the future of interaction analysis. Both the computational and immersive trends have significant epistemological and ontological legacies, which must be treated carefully by the researchers involved and we hope that this paper can stimulate this conversation. Within the former trend, there is a positivistic emphasis on measurable and countable aspects of learning interactions, whereas the latter trend emphasizes the perceptual and experiential aspects, perhaps at the expense of generalizability.

With the computational techniques, patterns can be identified that may be more generalizable and that may be difficult or impossible to interpret by relying on human perception of the analyst. Wise *et al.* (2021) argue that a core challenge for computer-supported analytics is to map clicks and strokes to behaviour and finally to learning constructs. This indicates that there is a link between digital traces, behaviour and learning construct – for computational techniques this is important as machines can detect digital input and to some extent behaviour. For immersive research the entryway to analysing and conceptualising learning constructs is of a different nature – here the situated activity is analysed in its totality – the immersive experience of the whole. The immersive trend builds on some of the core premises of earlier articulations of IA by creating opportunities for deep and human analysis of learning interactions. Immersing oneself in these interactions changes the researcher-data relationship in ways that must be further explored. As noted by Sinclair (2023, p. 378) 'with researchers going beyond passive spectator roles, automation, or relying on data-based abstractions from an educational situation' the prospect of situated research is re-established.

While the synergies between the computational and immersive approaches can impact the future of IA, researchers need to be aware of the differences in implicit assumptions about what constitutes learning. It would be naïve to integrate both approaches in IA without carefully considering when, how and on what basis they can complement each other. Educational researchers need to be deliberate and reflective about the ways we approach technology and method: is our research driven by technological possibilities or are we methodologically orthodox? Both the computational and immersive trends are driven by dedicated work in the entangled relationship between technology and method – and we urge educational researchers to take part in the development of both technology and method for researching learning.

Disclosure statement

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

ORCID

Jacob Davidsen http://orcid.org/0000-0002-5240-9452
Rolf Steier http://orcid.org/0000-0001-9809-3169

References

Ayass, R., 2015. Doing data: The status of transcripts in conversation analysis. *Discourse studies*, 17 (5), 505–528.

Baker, R.S., and Hawn, A., 2022. Algorithmic bias in education. *International journal of artificial intelligence in education*, 32 (4), 1052–1092.

Bernhard, J., *et al.*, 2019. Practical epistemic cognition in a design project-engineering students developing epistemic fluency. *Ieee transactions on education*, 62 (3), 216–225.

Blikstein, P., 2013. Multimodal learning analytics. *Proceedings of the third international conference on learning analytics and knowledge*, 102–106.

Blikstein, P., and Worsley, M., 2016. Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. *Journal of learning analytics*, 3 (2), 220–238.

Cao, Z., *et al.*, 2019. Openpose: realtime multi-person 2D pose estimation using part affinity fields. *IEEE transactions on pattern analysis and machine intelligence* 43: 172–186.

Crescenzi-Lanna, L. 2020. Multimodal Learning Analytics research with young children: A systematic review. *British Journal of Educational Technology* 51 (5): 1485–1504. https://doi.org/10.1111/bjet.12959

Curran, V., *et al.*, 2022. A phenomenological study of the use of 360° virtual reality (VR) video in pediatric and neonatal resuscitation training. *Health and technology*, 12 (1), 151–159.

Danish, J.A., and Saleh, A., 2014. Examining How activity shapes students' interactions while creating representations in early elementary science. *International journal of science education*, 36 (14), 2314–2334.

Davidsen, J., and McIlvenny, P., 2022. Towards collaborative immersive qualitative analysis, edited by Armin Weinberger, W. Chen, D. Hernández-Leo, and B. Chen, eds. *CSCL2022 conference proceedings*, 304–307. ISLS.

Davidsen, J., McIlvenny, P., and Ryberg, T., 2023. Researching interactional and volumetric scenographies – immersive qualitative digital research, edited by P. Jandric, A. MacKenzie, and J. Knox, eds. *Postdigital research*, 119–136. Cham: Springer.

DeLiema, D., *et al.*, 2021. Methodological innovations at the intersection of video-based educational research traditions: reflections on relevance, data selection, and phenomena of interest. *International journal of research & method in education*, 46 (0), 19–36.

Derry, S., *et al.*, 2010. Conducting video research in the learning sciences: guidance on selection. *Analysis, technology, and ethics. journal of the learning sciences*, 19 (1), 3–53.

Ding, N., Zhou, W., and Fung, A.Y.H., 2018. Emotional effect of cinematic VR compared with traditional 2D film. *Telematics and informatics*, 35 (6), 1572–1579.

diSessa, A.A., Levin, M., and Brown, N.J.S., 2015. *Knowledge and interaction: A synthetic agenda for the learning sciences*. New York: Routledge.

Enyedy, N., and Yoon, S., 2021. Immersive environments: learning in augmented + virtual reality. In: U. Cress, C. Rosé, A. F. Wise, and J. Oshima, eds. *International handbook of computer-supported collaborative learning*. Springer International Publishing, 389–405.

Erickson, F. 1985. *Qualitative Methods in Research on Teaching. Occasional Paper No. 81*. http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno = ED263203.

Erickson, F. 2007. Ways of seeing video: Toward a phenomenology of viewing minimally edited footage. In: R. Goldman, R. Pea, B. Barron, and S. Derry, eds. *Video research in the learning sciences*, 145–155. Lawrence Erlbaum Associates.

Ferguson, J., Aranda, G., Tytler, R., and Gorur, R., 2019. Purposeful selection from rich data sets. In: L. Xu, G. Aranda, W. Widjaja, and D. Clarke, eds. *Video-based research in education: cross-disciplinary perspectives*, 124–139. Abingdon: Taylor & Francis Group.

Fonteles, J., *et al.*, 2024. A first step in using machine learning methods to enhance interaction analysis for embodied learning environments. In: A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, and I. I. Bittencourt, eds. *Artificial intelligence in education*. Cham: Springer Nature Switzerland, 3–16.

Furberg, A., Kluge, A., and Ludvigsen, S., 2013. Student sensemaking with science diagrams in a computer-based setting. *International journal of computer-supported collaborative learning*, 8 (1), 41–64.

Garfinkel, H., 2002. *Ethnomethodology's program: working out Durkheim's aphorism*. Lanham: Rowman & Littlefield.

Goodwin, C., 1981. *Conversational organization: interaction between speakers and hearers*. London: Academic Press.

Goodwin, C., 2000. Action and embodiment within situated human interaction. *Journal of pragmatics*, 32 (10), 1489–1522.

Goodwin, C., and Goodwin, M.H., 1996. Seeing as a situated activity: formulating planes. In: Y. Engeström, and D. Middleton, eds. *Cognition and communication at work*. Cambridge: Cambridge University Press, 61–95.

Hall, R., 2000. Videorecording as theory. In: A. E. Kelly, and R. Lesh, eds. *Handbook of research design in mathematics and science education*. Mahwah: Lawrence Erlbaum Associates, 647–664.

Hall, R., and Stevens, R., 2015. Interaction analysis approaches to knowledge in use. In: A. A. diSessa, M. Levin, and N. J. S. Brown, eds. *Knowledge and interaction a synthetic agenda for the learning sciences*. Oxford: Routledge, 72–108.

Harris, A.M., 2016. *Video as method*. Oxford: Oxford University Press.

Hassani, H., *et al.*, 2020. Text mining in big data analytics. *Big data and cognitive computing*, 4 (1), Article 1.

Heath, C., 1986. *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.

Heath, C., Hindmarsh, J., and Luff, P., 2010. *Video in qualitative research: analysing social interaction in everyday life*. London: SAGE.

Herrera, F., *et al.*, 2018. Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLoS One*, 13 (10), e0204494.

Jefferson, G., 2004. Glossary of transcript symbols with an introduction. In: G. H. Lerner, ed. *Conversation analysis: studies from the first generation*. Cambridge: John Benjamins Publishing Company, 13–34.

Jordan, B., and Henderson, A., 1995. Interaction analysis: foundations and practice. *The journal of the learning sciences*, 4 (1), 39–103.

Krange, I., and Ludvigsen, S., 2009. The historical and situated nature design experiments—implications for data analysis. *Journal of computer assisted learning*, 25 (3), 268–279.

Krishnamoorthy, R., *et al.*, 2021. Learning to center relational ontologies: desettling interaction analysis methods. In: E. de Vries, Y. Hod, and J. Ahn, eds. *Proceedings of the 15th international conference of the learning sciences—ICLS 2021*. Bochum, Germany: International Society of the Learning Sciences, 851–858.

Lave, J., and Wenger, E., 1991. *Situated learning: legitimate peripheral participation*. Cambridge Univ Pr.

Lindwall, O., and Ivarsson, J., 2011. Differences that make a difference. In: S. Ludvigsen, A. Lund, I. Rasmussen, and R. Säljö, eds. *Learning across sites: new tools, infrastructures and practices*. Routledge, 364–380.

Love, C., *et al.* 2024. Caring relations across interaction analysis labs. *Proceedings of the 18th international conference of the learning sciences-icls* 2024. 18th International Conference of the Learning Sciences (ICLS) 2024. https://repository.isls.org/handle/1/10839.

Mangaroska, K., *et al.*, 2020. Multimodal learning analytics to inform learning design: lessons learned from computing education. *Journal of learning analytics*, 7 (3), 79–97.

Martinez-Maldonado, R., *et al.* 2020. From data to insights: a layered storytelling approach for multimodal learning analytics. *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–15). Association for Computing Machinery.

Martinez-Maldonado, R., *et al.*, 2021. What Do You mean by collaboration analytics? A conceptual model. *Journal of learning analytics*, 8 (1), 126–153.

Mazzocchi, F., 2015. Could Big data be the end of theory in science? A few remarks on the epistemology of data-driven science. *EMBO reports*, 16 (10), 1250–1255.

McIlvenny, P., 2014. Vélomobile formations-in-action: biking and talking together. *Space & culture*, 17 (2), 137–156.

McIlvenny, P., 2018. Inhabiting spatial video And audio data: towards A scenographic turn In The analysis Of social interaction. *social interaction*. *Video-Based studies of human sociality*, 2 (1). https://doi.org/10.7146/si.v2i1.110409

McIlvenny, P., 2020a. New technology And tools To enhance collaborative video analysis In live 'data sessions. *QuiViRR: qualitative video research reports*, 1, a0001–a0001.

McIlvenny, P., 2020b. The future of 'video'in video-based qualitative research is not 'dumb'flat pixels! exploring volumetric performance capture and immersive performative replay. *Qualitative research*, 20 (6), 800–818.

McIlvenny, P., *et al.* 2022. *DOTE: distributed open transcription environment* [Computer software]. Github. www.dote.aau.dk.

McIlvenny, P., and Davidsen, J., 2017. A Big video manifesto: Re-sensing video and audio. *Nordicom information*, 39 (2), 15–21.

McIlvenny, P., and Davidsen, J., 2023. Beyond video: using practice-based VolCap analysis to understand analytical practices volumetrically, edited by P. Haddington, ed. *Methodological explorations in and for EMCA: emerging directions for the study of social order*, 221–224. Routledge.

McIlvenny, P., Davidsen, J.G., and Stein, A., 2024. *DOTEbase: software tools for qualitative analysis*. Aalborg: Github.

Mehan, H. (1979). *Learning lessons: social organization in the classroom.* Cambridge: Harvard University Press.

Mondada, L., 2007. Commentary: transcript variations and the indexicality of transcribing practices. *Discourse studies*, 9 (6), 809–821.

Mondada, L., Monteiro, D.T., and Tekin, B.S., 2024. Collaboratively videoing mobile activities. *Visual studies*, 39, 267–290.

Nilsson, N.C., Nordahl, R., and Serafin, S., 2016. Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human technology*, 12 (2), 108–134.

Nivala, M., *et al.*, 2012. Interactive visual tools as triggers of collaborative reasoning in entry-level pathology. *International journal of computer-supported collaborative learning*, 7 (4), 499–518.

Norris, S., 2004. *Analyzing multimodal interaction: A methodological framework*. Routledge.

Nota, N., Trujillo, J.P., and Holler, J., 2021. Facial signals and social actions in multimodal face-to-face interaction. *Brain sciences*, 11 (8): 1–39.

Oh, C.S., Bailenson, J.N., and Welch, G.F., 2018. A systematic review of social presence: definition, antecedents, and implications. *Frontiers in robotics and AI*, 5, 114.

Pea, R., *et al.*, 2004. The diver project: interactive digital video repurposing. *IEEE multimedia*, 11 (1), 54–61.

Pierroux, P., Steier, R., and Sauge, B., 2019. Imagining, designing and exhibiting architecture in the digital landscape. In: Å. Mäkitalo, T. E. Nicewonger, and M. Elam, eds. *Designs for experimentation and inquiry*, 87–109. London: Routledge.

Pink, S., 2015. Going forward through the world: thinking theoretically about first person perspective digital ethnography. *Integrative psychological and behavioral science*, 49 (2), 239–252.

Pink, S., *et al.*, 2017. Empathetic technologies: digital materiality and video ethnography. *Visual studies*, 32 (4), 371–381.

Pirker, J., and Dengel, A., 2021. The potential of 360° virtual reality videos and real VR for education—A literature review. *IEEE computer graphics and applications*, 41 (4), 76–89.

Price, S., Jewitt, C., and Yiannoutsou, N., 2021. Conceptualising touch in VR. *Virtual reality*, 25 (3), 863–877.

Ross, J., 2017. Speculative method in digital education research. *Learning, media and technology*, 42 (2), 214–229.

Rystedt, H., and Lindwall, O., 2004. The interactive construction of learning foci in simulation-based learning environments: a case study of an anaesthesia course. *Psychology journal*, 2 (2), 168–188.

Shapiro, B.R., *et al.*, 2024. "Turn charts for interaction analysis: Visually mapping the conversation floor." *Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024* 43–50. https://doi.org/10.22318/icls2024.311167

Shapiro, B.R., Hall, R.P., and Owens, D.A., 2017. Developing & using interaction geography in a museum. *International journal of computer-supported collaborative learning*, 12 (4), 377–399.

Sharma, K., *et al.*, 2021. Challenging joint visual attention as a proxy for collaborative performance, edited by C. E. Hmelo-Silver, B. De Wever, and J. Oshima, eds. *Proceedings of the 14th international conference on computer-supported collaborative learning—CSCL 2021*, 91–98. Bochum, Germany: International Society of the Learning Sciences.

Sharples, M., and Pérez y Pérez, R., 2022. *Story machines: How computers have become creative writers*. Oxon: Routledge.

Shin, D., 2018. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in human behavior*, 78, 64–73.

Silseth, K., 2012. The multivoicedness of game play: exploring the unfolding of a student's learning trajectory in a gaming context at school. *International journal of computer-supported collaborative learning*, 7 (1), 63–84.

Sinclair, C., 2023. Afterword: A study of growth. In: P. Jandrić, A. MacKenzie, and J. Knox, eds. *Constructing postdigital research: method and emancipation*. Cham: Springer Nature Switzerland, 375–384.

Spikol, D., *et al.*, 2018. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning.*Journal of Computer Assisted Learning* 34 (4): 366–377. https://doi.org/10.1111/jcal.12263

Steier, R., *et al.* 2019. Tools and methods for '4e analysis': new lenses for analyzing interaction in CSCL. *A wide lens: combining embodied, enactive, extended, and embedded learning in Collaborative Settings*. The Computer Supported Collaborative Learning (CSCL) conference 2019, Lyon.

Trujillo, J.P., and Holler, J., 2021. The kinematics of social action: visual signals provide cues for what interlocutors do in conversation. *Brain sciences*, 11 (8), Article 8.

Umair, M., *et al.*, 2022. Gailbot: an automatic transcription system for conversation analysis. *Dialogue & discourse*, 13 (1), 63–95.

Vatanen, A., *et al.* 2022. *Experiences in collecting 360° video data and collaborating remotely in virtual reality*. *3*.

Vieira, F., *et al.*, 2021. A learning analytics framework to analyze corporal postures in students presentations. *Sensors*, 21 (4), 1525.

Vygotsky, L., 1978. *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, R., *et al.*, 2019. Examining destination images from travel blogs: a big data analytical approach using latent dirichlet allocation. *Asia pacific journal of tourism research*, 24 (11), 1092–1107.

White, T., 2018. Connecting levels of activity with classroom network technology. *International journal of computer-supported collaborative learning*, 13 (1), 93–122.

Wise, A.F., Knight, S., and Shum, S.B., 2021. Collaborative learning analytics. In: U. Cress, C. Rosé, A. F. Wise, and J. Oshima, eds. *International handbook of computer-supported collaborative learning*. Cham: Springer International Publishing, 425–443.

Wise, A.F., and Shaffer, D.W., 2015. Why theory matters more than ever in the age of big data. *Journal of learning analytics*, 2 (2), 5–13.

Worsley, M., Martinez-Maldonado, R., and D'Angelo, C., 2021. A new era in multimodal learning analytics: twelve core commitments to ground and grow MMLA. *Journal of learning analytics*, 8 (3), 10–27.

Young, G.W., O'Dwyer, N., and Smolic, A., 2022. Exploring virtual reality for quality immersive empathy building experiences. *Behaviour & information technology*, 41 (0), 3415–3431.

Zahn, C., Ruf, A., and Goldman, R., 2021. Video data collection and video analyses in CSCL research. In: U. Cress, C. Rosé, A. F. Wise, and J. Oshima, eds. *International handbook of computer-supported collaborative learning*. Cham: Springer International Publishing, 643–660.