

Exploring the Utility of Social-Network-Derived Collaborative Opportunity Temperature Readings for Informing Design and Research of Large-Group Immersive Learning Environments

Aditi Mallavarapu¹, Leilah Lyons², Stephen Uzzo³

Abstract

Large-group ($n > 8$) co-located collaboration has not been adequately studied because it demands different conceptual framings than those used to study small-group collaboration, while also posing pragmatic constraints on data collection. Working within these pragmatic constraints, we use video data to devise an indicator of the current possibilities for learner collaboration during large-group co-located interactions. We borrow conceptualizations from proxemics and social network analysis to construct *collaborative opportunity networks*, allowing us to define the concept of *collaborative opportunity temperature* (COT) readings: a “snapshot” of the current configuration of the different social subgroup structures within a large group, indicating emergent opportunities for collaboration (via talk or shared action) due to proximity. Using a case study of two groups of people ($n = 11$, $n = 12$) who interacted with a multi-user museum exhibit, we outline the processes of deriving COT. We show how to quickly detect differences in subgroup configurations that may result from educational interventions and how COT can triangulate with and complement other forms of data (audio transcripts and activity logs) during lengthier analyses. We also outline how COT readings can be used to supply formative feedback on social engagement to learners and be adapted to other learning environments.

Notes for Practice

- Bridging theories of affordance networks, which are network representations that describe the affordances for collaboration, and the complementary theory of proxemics, which describes the physical structure of these networks during collaboration and social network analysis (SNA), we conceptualize *collaborative opportunity networks* (CONs). The CONs use inter-learner proximity to create networks that indicate where collaborative engagement is possible.
- We show that clustering algorithms can be applied to CONs to detect different types of social groupings (in our learning environment, we identified four different component social subgroup structures: singleton, coterie, crowd, and club).
- This article defines a large group’s *collaborative opportunity temperature* (COT) as an instantaneous “reading” of the potential for collaboration, as defined by the current mix of social subgroup structures.
- COT can be ethically used in public settings because it works without storing identity information or tracking individuals, ensuring their anonymity.
- The case study we present serves as a proof of concept that COT can be useful to both designers and researchers to highlight social patterns in large-group immersive learning environments both as a quick way to contrast group behaviours and as a complementary triangulation to data that requires lengthier analysis.

Keywords

Co-located large groups, collaboration, multi-user learning environment, social network analysis, immersive museum exhibit, data-driven formative feedback

Submitted: 03/01/21 — **Accepted:** 19/11/21 — **Published:** 11/03/22

Corresponding author ¹ Email: amalla5@uic.edu Address: Department of Computer Science, University of Illinois at Chicago, Chicago, Illinois, USA; New York Hall of Science, Queens, New York, USA. ORCID ID: <https://orcid.org/0000-0002-2906-1478>

² Email: llyons@nysci.org Address: Department of Computer Science, University of Illinois at Chicago, Chicago, Illinois, USA; New York Hall of Science, Queens, New York, USA. ORCID ID: <https://orcid.org/0000-0003-1404-551X>

³ Email: suzzo@nysci.org Address: New York Hall of Science, Queens, New York, USA. ORCID ID: <https://orcid.org/0000-0002-6508-3535>

1. Introduction

In preparing learners for the future workforce, educational researchers and policy-makers have been drawing attention to the need to support 21st-century skills (National Research Council, 2010). These include intrapersonal skills (e.g., self-regulation and metacognition) and interpersonal skills (e.g., communication and relationship building), which allow learners to flexibly adapt to a complex and constantly changing workplace and to solve non-routine complex problems (National Research Council, 2011). To this end, educational researchers have responded by creating learning environments that allow learners to work together as they are immersed in complex problems. The increasing use of mixed-reality experiences (e.g., Moher, 2008; Malcolm et al., 2008; Dunleavy et al., 2009; Peppler et al., 2010; Tscholl et al., 2013; Slotta et al., 2013; Enyedy et al., 2015; Planey & Lindgren, 2018; Mallavarapu et al., 2019) situates learners in a collaborative experience in the real world with face-to-face interactions.

But our ability to build mixed-reality computer-supported collaborative learning environments has outpaced our ability to efficiently interpret the effect that the designs have on social learning experiences. Traditional observational methods do not scale well for larger groups, especially for learning environments that expose learners to open-ended problems that, lacking prescribed solutions or solution paths, give learners wide latitude in how they choose to organize and coordinate their work. While some design variations are simple to enact (e.g., “Does increasing the frequency of feedback given to the learners support or disrupt collaborative processes?”), the testing of the design variations is constrained by the difficulty of large-scale data collection. Researchers have proposed methods for converting qualitative observations into quantified measures of collaborative learning to support A/B comparisons of designs (Roberts & Lyons, 2017), but practically speaking such techniques take too much time (weeks to months) to realistically inform design cycles.

Some researchers have thus turned toward automated data analytics to respond to the pragmatic challenges of studying co-located collaboration (Chua et al., 2019; Y. Kim et al., 2020). By using “physical learning analytics” per Martinez-Maldonado and colleagues (2018), researchers are investigating micro-level collaborative dynamics. A review of the state of analytics for co-located learning found that most physical learning analytics research has taken place in laboratory or carefully controlled classroom settings, often using bespoke equipment with small groups of learners (Chua et al., 2019). This points to the need to explore practical and sustainable data collection techniques that can be applied to larger groups “in the wild.” An approach that is being used to analyze collaboration at a macro-scale is social network analysis (SNA). The digital traces from mobile network usage (Eagle et al., 2009) and emails (Fisher & Dourish, 2004) have been used to construct network diagrams for larger communities to portray collaboration and other behavioural indicators. However, the use of SNA in education has been limited to online learning environments like MOOCs (Carolan, 2013), mostly due to the readily available large-group interaction data from discussion forums (Dawson et al., 2010) and artifact manipulation logs (Hecking et al., 2014). To use SNA for co-located groups in learning environments, one has to decide how to conceptualize social interactions and select a data collection instrument that can conveniently detect these forms of interactions.

This work seeks to address the gap in existing literature for automated approaches that can characterize large-scale in-person collaborative learning interactions by devising an easy-to-deploy, low-cost, low-touch (i.e., no special instrumentation attached to learners) automated physical learning analytics system. We used a simple, uncalibrated two-camera setup to record how learners moved within our learning environment, a large-scale mixed-reality museum exhibit. Our setting strongly influenced our data collection process, it being common practice in museums to collect easy-to-gather visitor movement data (“timing and tracking” data) to provide rapid feedback on exhibit design variations (Yalowitz & Bronnenkant, 2009). In contrast to physical learning analytics work that has directly studied collaborative mechanisms at a micro-level, in this work we are interested in characterizing how learners’ opportunities for engaging with one another can be impacted at a meso-level by the different types of subgroups learners can form and at a macro-level by how those groups gather and break apart. We used proxemics and features derived from SNA to parameterize what we call the *collaborative opportunity network* (CON).

The CON instantaneously maps which learners might reasonably be expected to be able to engage with which other learners via talk or shared action. To make the network useful for evaluation, we used a clustering algorithm to distill common network substructures (in our case, we detected singletons, coteries, crowds, and clubs). We use the observed collection of substructures to conceptualize *collaborative opportunity temperature* (COT): an instantaneous reading of the potential for collaboration at any given moment. Such information is directly useful to designers of large-scale mixed-reality learning environments in evaluating

how their interaction design decisions affect the opportunities learners have to engage with one another (Lyons & Mallavarapu, 2021). For researchers, such readings can serve as a quantitative triangulation with richer qualitative data, such as observations, interviews, and qualitative coding, or as a data reduction technique by flagging episodes that warrant more in-depth analysis. We envision the readings being used to generate formative feedback to help educators respond in real time to trends in group formation.

A significant advantage of our technique is that it preserves privacy: we do not track individuals, so we retain no identifiable information that could potentially be misused (Paullada et al., 2021). As an institution that serves many immigrant families with undocumented members, our museum is keenly aware that our audience has a deep distrust of institutional promises not to misuse identifiable information. So we committed to designing a system that would throw away video data as soon as each frame was analyzed. While this decision precludes us from performing more sophisticated analyses of how learners build relationships over time, it is also more robust to errors that can result from tracking glitches. Our approach will not work for all large-group in situ collaboration research, but it fills a niche currently unmet by existing approaches by generating simple, more reliable indicators.

We first situate the contribution of our approach (Section 2) in the existing literature, and then we describe our research questions, research setting, and data collection (Section 3). In Section 4, we derive CONs and use SNA to detect common social subgroups in the CONs. Then, in Section 5, we explore how to interpret these subgroups, constructing COT readings and applying them to two cases as a proof of concept of the technique. Section 6 discusses the potential of COT to support further data collection (by informing exit interviews) and analysis (by guiding qualitative analysis). In Section 8, we describe plans for validation and elaboration of shared opportunity networks, and their use for generating formative feedback on social engagement for learners, before concluding.

2. Related Work

Over the past two decades, researchers have used instrumented data from various contexts to analyze how collaborative learning takes place within physical settings, building from structured classroom contexts (with fixed groups) to more chaotic informal contexts like public platforms and workplaces (with fluid groups). Different contexts affect the ease or difficulty of instrumentation and permit different collaborative enactments (e.g., fixed positions of learners in classrooms versus visitors scattered along an exhibit hall). Our work was performed in the context of a large collaborative exhibit in a science museum, so this section will begin by reviewing the history of spatial movement metrics in museums and then discuss the more recent development of metrics that combine spatial movement information with other features relevant to visitor engagement and learning. We then review other methodological areas (collaborative analytics and SNA) that we drew upon to devise our methodology.

2.1 Studying Movement and Collaboration in Museums

Falk and Dierking (2018)'s contextual model of learning frames learning in informal learning environments as a process governed by the cognitive and motivational properties of an individual as well as the sociocultural and physical dimensions of the learning environment. Many of the early metrics that researchers and evaluators developed to study museum learning experiences involved documenting physical aspects of learners' engagement. Historically, museums have relied heavily on the timing and tracking observational studies (Yalowitz & Bronnenkant, 2009) pioneered in the early 19th century (Robinson, 1928), which study how visitors move about exhibitions and where they choose to linger. The purpose of these metrics was to provide tools for museum evaluators and researchers to quickly gauge the impact of different design decisions (e.g., exhibit signage, exhibit layout) on visitors' ability to engage with museum content. Researchers would typically spend several days observing visitors, using rubrics and various sampling strategies (Diamond et al., 2009), and then sum the observation counts to produce numerical values that could be compared across different variations of the exhibit.

Over the years, the timing and tracking approach has been updated to include numerical indexes used to reflect patterns of visitor engagement with exhibits (Serrell, 1997) and to construct formal models of visitor attention (Bitgood, 2006). It has been streamlined to use tracking technologies like point-of-view cameras worn by visitors (Shapiro et al., 2017), mobile devices (Moussouri & Roussos, 2013, 2015; Schautz et al., 2016), RFID (Lanir et al., 2017), and eye-tracking equipment (Eghbal-Azar & Widlok, 2013; Filippini Fantoni et al., 2013; Magnussen et al., 2017). Timing and tracking studies have been used as a loose proxy for learning, with the assumption being that longer linger times equate to greater opportunities for learning (Barriault & Pearson, 2010). Some tracking technologies primarily streamline collecting visitor location data to make traditional timing and tracking computations more seamless (Schautz et al., 2016), with the additional ability to more easily aggregate visitor data across space and time (Lanir et al., 2017). The inclusion of additional information along with location data, however, can further elucidate visitor behaviours. For example, at a minimum, visitor gaze data can be used to determine if visitors really are attending to exhibit elements when they pause at an exhibit (Filippini Fantoni et al., 2013) or whether they can find

relevant signage (Magnussen et al., 2017). Creative forms of representing visitor location data, like the “interaction geography” visualizations of Shapiro and colleagues (2017), have been used to explore how parents and children manage their experiences exploring joint exhibits.

To support stronger inferences regarding learning, evaluators and researchers have augmented location-based metrics with behavioural observations like the degree of engagement (Barriault & Pearson, 2010) or more qualitative analyses of visitor talk (e.g., Allen, 2003). It is worth noting that under these schemes collaboration is not often documented for its own sake—rather, the social interactions of visitors are most often used for their potential to reveal clues about why they might choose to approach a location (Eghbal-Azar & Widlok, 2013) or about the state of learners’ individual understanding and engagement (e.g., Allen, 2003). For example, combining tracking information with qualitatively coded visitor talk and logged activity data can shed light on how visitors jointly make sense of digital exhibits (Roberts & Lyons, 2017) or engage in joint information-seeking (Magnussen et al., 2017).

There have been a few attempts to use visitor movement patterns to document informal social interactions directly. Several researchers have built from “F-formation” theory (Kendon, 2010), which describes how people orient themselves toward objects and one another when interacting, to study museum gallery visitors via observations (Marshall et al., 2011), small groups of nursing students via depth cameras (Martinez-Maldonado et al., 2017), and museum visitor pairs via RFID tags worn by visitors and sensors placed in a gallery (Dim & Kuflik, 2014; Kuflik & Dim, 2013). The focus of F-formations is how people manage their joint attention, which is known to be critical to collaborative learning in general (Barron, 2003) and has been studied in the context of museum exhibit use (Lyons, 2009; Lyons et al., 2015; Povis & Crowley, 2015). There have been some attempts to streamline or automate the detection of small-group collaborative behaviours around exhibits by using algorithms to define group identity (e.g., member, collaborator, connector, joiner, intruder, or shopper) from manually coded observations (Block et al., 2015) or by using hidden Markov models (Tissenbaum et al., 2016) and clustering (Jorion et al., 2020) to translate logged data from touch screens into information about joint engagement. Thus far, however, there have been no automated attempts to study collaborative museum interactions at a larger scale.

One of the barriers to studying large-group collaborative interactions in informal contexts has been the difficulty in instrumentation, since such settings often encompass large spaces with a lot of movement. Recently, Yan and colleagues (2021) used sensors in wristbands to track students and teachers in an open-classroom environment, but wearables are not pragmatic for an exhibit that sees more than 3,000 visitors a day. Camera-based tracking systems are a touchless solution, but glitches that invalidate any analytics that rely on uniquely identifying learners can be a major issue with these systems (Cafaro et al., 2013), for example, analytics that study how the social engagement of individuals evolves over time. Nonetheless, there is an opportunity for exploring how “lossy” camera-based methods could characterize social learning behaviours.

Our work builds directly on the longstanding timing and tracking metrics used in museums, incorporating a social dimension and responding to pragmatic instrumentation challenges by intentionally biasing toward a simple and reliable metric. We call this next-generation timing and tracking method a “timing and tracking, together” approach, operating akin to Kuflik and Dim (2013) and Dim and Kuflik (2014), but for large groups. As emphasized above, timing and tracking metrics do not directly measure visitor learning; they assume that when visitors linger longer in front of an exhibit, the *opportunity* for them to learn increases (Barriault & Pearson, 2010). Similarly, we do not propose to make strong claims about measuring the *quality* of collaboration via our metric—we seek only to document *opportunities* for collaboration, as suggested by learners’ proximity to one another. This simplifying assumption reduces the sophistication of the analyses we can perform, but it allows us to develop a technique that is easier to deploy, scalable for larger group sizes, efficient, and robust against errors. Our analyses can be used to support simple A/B comparisons of alternative designs as traditional timing and tracking methods or can be used to complement other more nuanced forms of analysis.

2.2 Collaborative Analytics

The studies described above that used visitors’ relative positions to infer social engagement are an example of proxemics. The term *proxemics* was coined by Hall (1966) to describe the study of the cultural dimensions of how humans use space. Because spatial relations between people can give insights into their intentions to communicate and interact with each other (Hall, 1966), proxemics has been combined with automated data collection to identify social structures in human behaviour through mobile phone data usage (Eagle et al., 2009), to understand collaboration in the workplace through sociometric badges (T. Kim et al., 2012; Lepri et al., 2012; Wu et al., 2008), and to identify plays made in sports through multimodal sensing (Alldieck et al., 2018; Intille & Bobick, 1995; Thirde et al., 2006; Xu et al., 2004).

The kinds of data collection technologies used can affect how data will be analyzed and the information about collaboration that can be inferred. Historically speaking, collaborative learning has frequently been studied by qualitatively coding dialogue between learners, and so some researchers have begun automating the process of analyzing voice data to detect indicators of collaboration (D’Angelo et al., 2019). But ubiquitous sensors are opening the door to new “physical learning analytics” (Martinez-Maldonado et al., 2018), permitting the study of gestures (Schneider & Blikstein, 2014), facial expressions (D’Mello

et al., 2017), physical activity levels (Echeverria et al., 2019), attention (Chan et al., 2020), and gross motor movements (Cafaro et al., 2013). For small-group learning scenarios, some researchers have looked for ways to turn these physical analytics into markers of collaboration by examining the alignment or synchrony of learners' activity (Cukurova et al., 2018; Reilly et al., 2018; Schneider & Blikstein, 2015). Just recently, there has been a new push for creating a domain of "collaborative analytics" (Schneider et al., 2021). Most researchers use some theoretical understanding about collaborative processes to transform raw data into information about collaboration. The majority of analytics-based studies in educational research have framed collaboration through an intrapersonal lens, gauging an individual's collaborative engagement through student interaction logs (e.g., Janssen et al., 2007; Janssen et al., 2010; Martínez et al., 2011; Suthers et al., 2008; Tissenbaum et al., 2016; Tissenbaum et al., 2017; Wang et al., 2017). We were interested in studying collaboration at a group level, and so we turned toward SNA in order to transform our proxemic data into information about collaborative potential.

2.3 Studying Collaboration through SNA

SNA allows researchers to study collaboration as a process of interplay between individuals and the collective (Dado & Bodemer, 2017; Carolan, 2013), where a social network model is composed of groups of individuals (referred to as nodes) and the relations (referred to as edges) among them. It conceptualizes the structure of the network as enduring patterns of relations among the individuals (Wasserman & Faust, 2012). A key tenet of SNA is that "one's location in a social structure shapes one's opportunities and outcomes" (Carolan, 2013). What "counts" as a relation can be quite varied, from organizational hierarchy connections to communication acts to physical proximity, and the way relations are defined shapes the research questions that can be posed (Carolan, 2013).

SNA has been readily applied to study the collaboration of large groups of learners engaged with e-learning and computer-mediated learning environments, where the relations between individuals are easily defined and captured, usually in the form of acts of digital communication like forum posts (Bakharia & Dawson, 2011; Willging, 2005; De Laat et al., 2007; Rabbany et al., 2014; Puntambekar et al., 2011; Aviv et al., 2003; Dawson et al., 2010; Erlin et al., 2009). In these cases, there is no question about whether or not the learners are interacting; the challenge for researchers is characterizing the form of collaboration these interactions represent. Adapting SNA from digitally mediated learning environments to study collaboration in face-to-face co-located social settings requires researchers to define what kinds of loggable behaviours can be used to define a relation and, in turn, what kinds of claims can be made about collaboration from that data. Some SNA research on e-learning environments has characterized collaboration as a network of dynamically evolving interpersonal connections with the other learners (e.g., Dráždilová et al., 2008; Haythornthwaite, 1999; Mansur et al., 2012; Poquet & Jovanovic, 2020), but given the dual constraints of being unable to reliably identify learners from camera data and to know whether learners in proximity to one another actually interacted with one another, we cannot claim to be able to characterize large-group collaboration itself. Rather, we can use SNA to analyze the relations formed by moment-to-moment proximity to characterize *opportunities* for collaboration.

3. Background and Methods

3.1 Research Questions

This work is the first to our knowledge to explore how to study co-located large-group collaboration in an open-ended learning environment via SNA. To develop our metric and explore its utility, we revisit a pair of cases already studied in depth using log-file analysis of actions performed upon the simulation (Mallavarapu et al., 2019). These cases provide an interesting opportunity to examine how changes in group problem-solving strategies may evidence themselves in social structures, since one case received data-driven dashboard feedback whereas the other did not.

Our approach is to first construct CONs, building on the concept of affordance networks, which are the "functionally bound potentials . . . that can be acted on to realize particular goals" within a learning environment (Barab & Roth, 2006). In other words, a CON describes where and when social interactions are possible, should learners choose to take advantage of that opportunity. Respecting logistical and privacy concerns (Greenberg et al., 2014), we refrain from using dynamic SNAs that track individuals in the space. Rather, we use the CON to derive a large group's moment-to-moment configuration of social structures, an instantaneous reading of the "state of collaboration" that we conceptualize as COT.

Our study explores the following research questions:

1. What kinds of social subgroup structures can be found in CONs in a large-group learning environment?
2. How can the social subgroup structures derived from CONs be used to construct a measure of the potential for collaboration (a COT)?
3. How can COT readings be used to complement qualitative analyses of large-group immersive learning environments?



Figure 1. Connected Worlds exhibit with visitors, showing the biomes and water sources (clockwise) desert, mountain valley, plains, waterfall, jungle, reservoir, and wetlands. The tangible logs are used to direct water toward the biomes.

3.2 Study Site: Connected Worlds

The Connected Worlds exhibit is a 2,500-square-foot gallery that immerses up to 50 visitors in a simulated world. The visitors can interact with four different interconnected biomes (desert, plains, jungle, and wetlands; see Figure 1) by raising their hands in front of the projected screens to plant seeds (gestures detected by Kinect cameras) and can manage the shared resource, water, by moving the foam “logs” (detected by the IR camera) to divert water from three sources (the 60-foot-high waterfall, the mountain valley, and the reservoir). Water flowing into a biome collects as “groundwater,” which is consumed by the simulated plants in the biome. The exhibit is intended to introduce the challenges involved in attaining and sustaining a diverse ecosystem: both the natural challenge of distributing resources and the human challenge of coordinating efforts to achieve that distribution.

Connected Worlds is the first large-scale immersive digital simulation of its kind and offers unique opportunities to conduct research on large-group collaboration during problem solving (Levy-Cohen et al., 2021). It also poses interesting new methodological challenges, since the scale of the exhibit is much bigger than any prior immersive learning environment and because the actions of learners can be observed at multiple grain sizes: individual (micro), small group (meso), and large group (macro).

3.3 Study Design

We use a multiple case study approach (Stake, 2013) to explore the use of our metric to study large-group collaboration, following the lead of Derry and colleagues (1998), who found this to be a useful way to explore how small-group collaboration could unfold. The cases are already well studied (Mallavarapu et al., 2019; Levy-Cohen et al., 2021) and between them hold a number of design elements constant (the simulation settings, the duration of engagement, the activity structure, the instructions provided to participants, the demographics of participants), such that any observed differences in the SNA metrics highlight the potential of this method to reveal differences in the social behaviours. Moreover, the cases were structured so that the feedback intervention was deployed mid-session, so that within-case comparisons can be made in addition to cross-case comparisons, giving us an opportunity for deeper analysis of the two cases under consideration.

3.3.1 Setup and Data Collection

The sessions were recorded using two tripod-mounted iPads equipped with cameras (placed on either side of the entrance to the exhibit space). The video data provided two (uncalibrated) views of the learning environment that were post-processed to yield a unified coordinate system (see Section 4.1.1 for details). The exhibit also recorded the interaction logs that capture the evolving system state of the exhibit every second in real time. The participants were randomly assigned to teams of two or three people to help manage a specific biome within the ecosystem. Three researchers were involved in facilitating the sessions and prompting questions during discussions.

3.3.2 Activity Design

Two groups ($n = 11$, $n = 12$) were given the standard Connected Worlds experience, which is typically run as a 15-minute session, divided into different periods. They were first given a two-minute introduction to Connected Worlds, including how they could divert water using logs, how to plant seeds, and how they could monitor the biomes’ groundwater levels. Each group was given the same goal: to attain and maintain diversity in the biome plant life while sharing the water across the four biomes.

They interacted with the exhibit for five minutes (Phase 1), and then the simulation was paused for a three-minute whole-group reflection discussion, where the facilitator helped them reflect on what their goals were, if they were able to achieve them, and if they had any plans for the next interaction phase.

During reflection, participants in the feedback condition were additionally shown a data visualization presented on a large mobile tablet, computed from the automatically logged system state data. The charts summarize how during Phase 1 water was distributed across the four biomes, and how much water the plants in each biome needed to survive. The facilitator explained what the data and the axes represented but otherwise did not interpret them for the visitors, allowing them to use the data to inform their discussion to the extent they desired.

After the reflection, the groups were given another five-minute interaction phase (Phase 2). For this study, we analyzed the video data from Phase 1 and Phase 2, but a qualitative analysis of the participants' reflection discussions and collaboration-coordination behaviour can be found in Mallavarapu and colleagues (2019).

3.4 Participants

Twenty-three adults, including 11 women and 12 men (aged 22–57 years, $M = 33$), participated in this study for one hour on a Saturday. The adults were a part of a volunteer group that was visiting the museum and had no prior Connected Worlds experience. All chose to consent according to our institutional review board (IRB) procedures. These volunteers were in keeping with the diverse ethnic demographics of the visitors (30% white, 30% Hispanic/Latinx, 40% Asian), and their education level (10% high school, 65% bachelor's, 25% graduate) matched the institution's visitor education demographics as well as general demographics for science centres (Korn, 1995).

We chose to work with this adult group partly out of convenience, because the logistical need to have eight or more consenting people present in order to run just a single session proved to be surprisingly difficult (we are not permitted to video record students on school trips—a major motivation for exploring anonymous automatically computed measures was to comply with this IRB constraint). We further reasoned that adults who were new to the exhibit would be best positioned to reveal any changes in collaboration associated with the feedback intervention: they would have the most need for help understanding the exhibit, they would have the best chance of understanding the data representations, and they would have had real-world experience with collaborative problem solving. (Recall that at its heart this is an exploratory case study intended to gauge the potential of SNA-derived metrics for revealing useful differences in collaborative behaviours—selecting “best-case” conditions maximizes the opportunity to explore this potential.)

4. Procedure: Constructing and Analyzing CONs

4.1 Deriving CONs from Proxemics

We do not claim that individuals in close proximity are necessarily communicating; rather, a *collaborative opportunity* is a circumstance wherein participants are close enough to one another to easily converse and be exposed to the same visual cues, which is known to play a synergistic role in collaborative coordination (Gergle et al., 2007). Proximity networks can be drawn from participants' location data by considering individuals as nodes and drawing edges between nodes when the proxemic distance indicates an opportunity for discussion or shared engagement. These abstract representations indicate potentials for social interactions, so we call them CONs. From a designer's perspective, creating the right opportunities for desired actions often *is* the end goal—while designers cannot make a person pick up a teapot and pour it, they can certainly communicate the possibility of that action via the affordances of the design of the teapot handle (Norman, 1988). We thus position *collaborative opportunities* as affordances for collaboration and CONs as collaborative “affordance networks” (Barab & Roth, 2006), which learners can make use of at their discretion. It is worth noting that this is one major difference between small-group and large-group collaboration: in most small-group collaboration scenarios, the participants will all belong to the same CON for the full duration of their work—CONs are thus not useful for examining collaboration in most small-group settings. (Other affordances, like the position and visibility of technology, may affect a group's ability to collaborate (Lyons & Mallavarapu, 2021), but we are not considering them in this paper; see Section 8.)

Below we describe our three-step process of converting the unstructured video data into computation-amenable features (see Figure 2).

4.1.1 Identifying Nodes

We experimented with three different pose estimation algorithms—fastpose (Zhang et al., 2019), tf-pose (Howard et al., 2017), and OpenPose—to detect the nodes (the positions of visitors), but we achieved the best performance using OpenPose, which performs well dynamically and in less controlled environments (where occlusions are common). We applied OpenPose (Cao et al., 2017) to each camera feed and aligned the feeds using timestamps to build the CONs (see Figure 3). We used the dimensions from an architectural schematic of the exhibit and the known position of the cameras to reverse-engineer the 2D

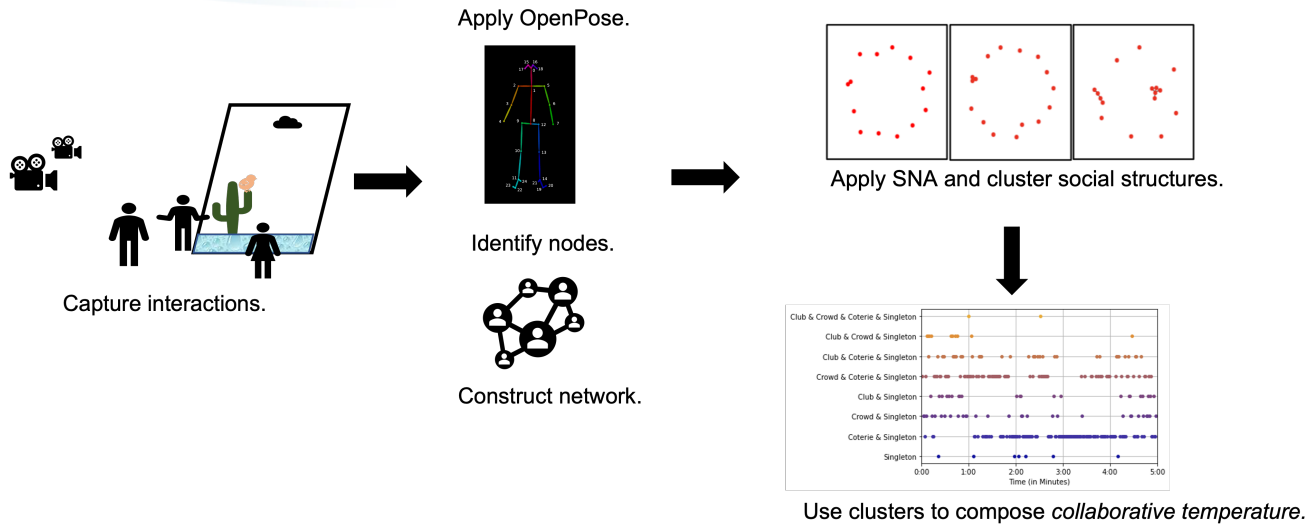


Figure 2. Process employed to convert video data to features amenable to SNA to extract COT.

(*x, y*) image coordinates provided by OpenPose into the 3D Cartesian space of the exhibit. When both cameras detected the same node (i.e., when the same person was present in both feeds), we discarded one node copy.

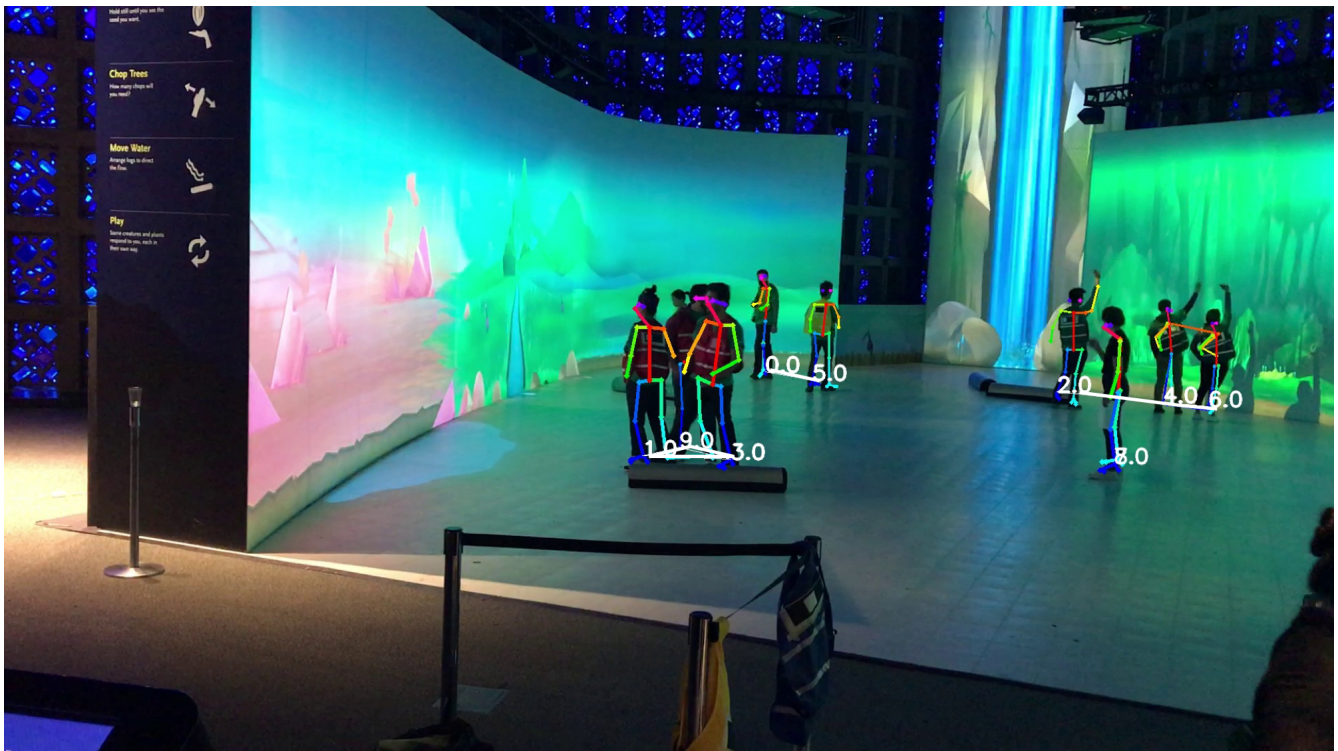


Figure 3. Left camera frame overlaid with the key points and positions of individuals (the numbers indicate the NodeID in that frame). This image shows two triads, one dyad, and two singletons.

The main risks of this approach are (1) interpersonal occlusions and (2) perspective distortion. Interpersonal occlusions were very common in Connected Worlds given its size and the dynamic conditions imposed by the problem. Fortunately, the occlusions also tended to be very short (an average of three frames of video). Because we were not tracking what individuals do (e.g., what ties they form over time, and with whom), we judged that this would not overly skew our data. The perspective distortion from the size of the exhibit made precisely pinpointing the location of people by the waterfall at the far end difficult, but fortunately the error fell within the 2.1-m proxemic boundary we used (see Section 4.1.2 for more details). The algorithm

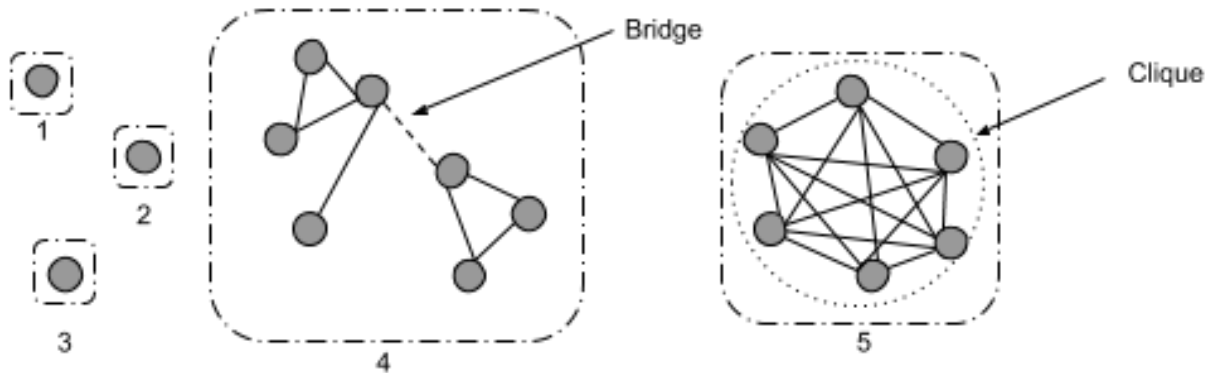


Figure 4. Example subgroup arrangements used to illustrate the SNA measures described in the text.

yielded ~95% accuracy in detecting visitor positions (we manually inspected 2% of the frames—a 1-Hz sampling—for false positives and false negatives). See Section 6.1 for a discussion of how our process would port to other settings.

4.1.2 Constructing CONs

We used the 2.1-m social proxemic distance defined by Hall (1966) to create CONs. We constructed CONs on a frame-by-frame basis and did not attempt to maintain the identity of individuals across frames. Across the 20 minutes of video data we had ~36,000 frames, so we sampled the frames at a 1-Hz frequency, detecting a total of ~1,200 networks across the two groups in the two phases. We used the Python network library NetworkX (Hagberg et al., 2008) to compute the different static network metrics.

4.2 Detecting Social Subgroup Structures in CONs

A property of large-group collaboration that sets it apart from small-group collaboration is that participants are very seldom (if ever) all working together (i.e., all focused on the same task, or in communication with one another). Rather, learners in large groups form transient ad hoc working groups to execute specific tasks or communicate information. At any given moment, we would expect a large group to have a different collaborative composition, with a mix of singletons and small subgroups of various sizes and degrees of connection.

4.2.1 Parameterizing Social Subgroup Structures with SNA Metrics

Since different subgroup arrangements allow for different opportunities for collaboration, we first need to be able to recognize what kinds of subgroups are commonly formed in large-group collaboration scenarios. Subgroups can be defined by both size and structure. We first parameterize the CON, selecting standard SNA measures that capture structural features relevant to collaboration:

1. The *size of the component* is the number of learners in the subgroups formed within the network. It is the size of the working groups. In Figure 4, component 4 is of size seven.
2. A *clique* is a component where every node is connected to every other node in the subgroup. The cliques are potentially tightly coordinated working groups. In immersive settings, they may also indicate conversational groups. In Figure 4, component 5 represents a clique.
3. *Density* is the ratio of ties in the component to the number of possible ties between the nodes in the component, defined as

$$D = \frac{2L}{N(N-1)}, \tag{1}$$

where L = number of edges in the component and N = total number of nodes. The density is another way to examine working groups for their potential for tight coordination. In Figure 4, the density of component 4 is 0.38, while the density of component 5 is one.

4. The *number of bridges* counts the number of unique paths between two connected components, which if removed would split the graph into two components. The number of bridges can indicate connections between two working groups, which can serve as information and resource diffusers. In Figure 4, there is one bridge.

4.2.2 Clustering Components of CONs to Detect Collaboration Structures

The 1,200 CONs yielded $\sim 70,000$ distinct non-singleton components (and $\sim 13,000$ singletons). We excluded singleton components from the clustering process because for single individuals, SNA parameters such as density, clique, and number of bridges are zero. We recorded the number of singletons and used that in later analyses. We applied the KMeans (Lloyd, 1982) algorithm to the parameterized non-singleton components. Our approach differs from methods like spectral clustering (Newman, 2006) (which uses SNA internally to *identify distinct subgroups* within a network) because we rely on proxemics to define the subgroups and apply clustering to *characterize the compositions* of subgroups.

4.3 Defining and Analyzing Collaborative Temperature Readings

Corresponding to the compositions characterized by the clustering process, we define the COT as a snapshot of these compositions. For RQ1, we follow a quasi-experimental setup to compare the frequencies of subgroups for the two cases. To map the instant-by-instant presence or absence of the compositions and their co-occurrence patterns with respect to each other and to observe the evolutions of the COT, we used a qualitative visual analysis method. First, to strengthen the idea of evolution, we used a temporal scatter plot to understand how the different compositions changed over time. The scatter plot defines the dynamic and emergent nature of the social compositions. To further contextualize the results, we triangulated the results from the COT readings using log-file analysis done by other researchers (Mallavarapu et al., 2019).

5. Results

We first describe the emergent social subgroup structures detected from the CON-based clustering and how we use these subgroup compositions to define momentary readings of the COT. We show the potential of the COT to track changes in social behaviours by comparing the identified subgroup compositions over time. We further show the compatibility of COT with other data sources to show how we could use the readings to more deeply analyze social behaviour.

5.1 RQ1: Defining Social Subgroup Structures Found in CONs

5.1.1 Detected Social Structures

Clustering the parameterized descriptions of the non-singleton components of the CONs revealed the most common social subgroup structures of large-group collaboration in our learning environment. We varied k for the clustering algorithm from two to seven and measured the quality of the clusters using the silhouette measure (Rousseeuw, 1987), which led to the selection of an optimal k of three.

We roughly categorize the three identified subgroup structures (obtained from clustering) and the singletons along two dimensions: (1) coarseness, which describes the size of the components—components with lower coarseness have fewer learners, and those with higher coarseness have more learners—and (2) cohesiveness, which is the degree of connectedness of the learners within the components. Components with high cohesion have higher density, more cliques, and fewer bridges (see Figure 5). For ease of discussion, we have labelled these four quadrants as singletons (low coarseness, low cohesiveness), coterie (low coarseness, high cohesiveness), crowds (high coarseness, low cohesiveness), and clubs (high coarseness, high cohesiveness). In our case, the social structures in the coterie quadrant happen to all be dyads (i.e., pairs of learners), and the crowds were mostly chains where every learner had the opportunity to interact with at most two other learners, while the clubs showed opportunities for more intricate connections among the learners (where each learner could have the opportunity to interact with nearly every other learner in the subgroup if they wanted to).

5.2 RQ2: Constructing a COT

Here we explore how we can further use the nature and frequency of the identified social subgroups to compare two cases of a quasi-experimental setup to see if a design intervention affected collaboration opportunities. The goal of this research question is to explore how to produce a social version of a timing and tracking measure, something that can be deployed and interpreted to get quick feedback on the design of large-group interactive learning environments.

5.2.1 Comparing Frequencies of Social Subgroups

We compared the frequencies of the subgroup structures across the two interaction phases of our two cases using the Pearson's chi-squared test. Table 1 shows that there was no significant difference in the frequencies of high-coarseness subgroup structures (crowds and clubs) in Phase 1 versus Phase 2 for the non-feedback condition. The feedback condition, however, showed significant increases in the frequency of both crowds and clubs in Phase 1 versus Phase 2 (see Table 2). This suggests that the

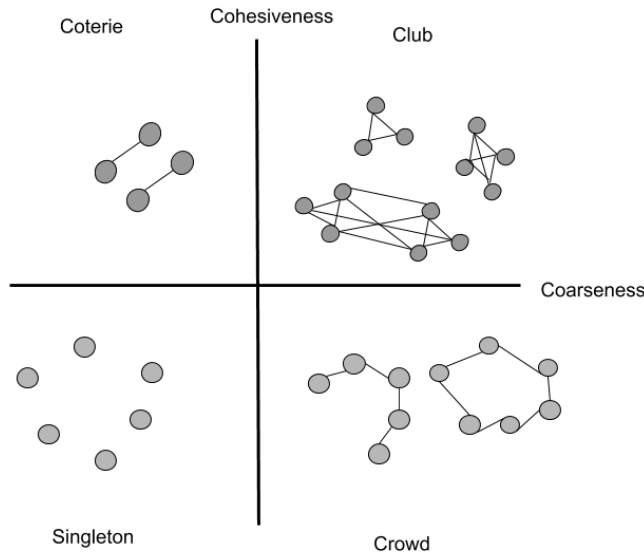


Figure 5. Sociograms representing the four main categories of social structures we were able to detect in our data. The grey dots are the learners, and the black lines represent a potential for shared engagement, as defined by proximity.

educational intervention used in the feedback condition may have encouraged social configurations that would have supported more opportunities for participants to plan and coordinate.

Table 1. Comparison of the Frequencies of High-Coarseness Subgroup Structures (Crowds and Clubs) Seen in Phase 1 versus Phase 2 for the Non-feedback Condition

Social Structure	Phase	Per-Frame Count of Structures Seen			df	N	χ^2	p
		0	1	2				
Club	1	187	108	8	2	706	2.76	0.252
	2	266	132	5				
	Δ	79	24	-3				
Crowd	1	185	105	13	2	706	0.451	0.798
	2	236	149	18				
	Δ	51	44	5				

Table 2. Comparison of the Frequencies of High-Coarseness Subgroup Structures (Crowds and Clubs) Seen in Phase 1 versus Phase 2 for the Feedback Condition (* $p < 0.05$)

Social Structure	Phase	Per-Frame Count of Structures Seen			df	N	χ^2	p
		0	1	2				
Club	1	256	42	0	2	597	9.647	0.005*
	2	229	68	2				
	Δ	-27	26	2				
Crowd	1	219	75	4	2	597	11.286	0.004*
	2	182	108	9				
	Δ	-37	33	5				

There were no significant differences in the frequencies of coterie between Phase 1 and Phase 2 for either the non-feedback or the feedback condition, nor was there a significant difference in the number of singletons in the non-feedback condition, but there was a significant *decrease* in the frequencies of singletons seen for the feedback condition [$\chi^2(8, N = 597) = 54.157, p < 0.001$]. This aligns with the increase of high-coarseness structures in Phase 2 for the feedback condition. One could

hypothesize that the feedback participants, after being exposed to the educational intervention, were less likely to “act alone” and, instead, gathered in subgroups that provided opportunities for coordination.

Examining each meso-level subgroup individually does not reveal much about the full social context (e.g., how these subgroups co-occur and how their frequencies change over time), however. To explore this wider context, what is needed is a moment-to-moment measure of the macro-level *composition* of social groups, which we explore next.

5.2.2 Constructing COT: A Measure of Social Subgroup Compositions

Another approach to characterizing the potential for social engagement in large groups is to characterize the overall composition of social subgroups present. We define a large group’s COT as an instantaneous “reading” of the potential for collaboration at a macro-level, as defined by the current mix of social subgroup structures present. Recall that the social structures found in our data are characterized along two different dimensions, cohesiveness and coarseness, each of which has different implications for the nature of collaboration, so we must be careful to stress that COT cannot be cleanly mapped to a single scale from “low” to “high,” although we have arranged the readings in order from less diverse to more diverse subgroup compositions. We elected to use the term *temperature* to evoke the way thermometers are used to take an instantaneous sample—here, of the current potential for collaboration. This paints a richer picture, supplying the full macro context of co-occurring subgroups, than the summative statistical comparison of the frequency of subgroup structures from Section 5.2.1.

A scatter plot (see Figure 6), taken from the feedback condition, shows on the y-axis the different COT readings, with time in minutes on the x-axis. The dots represent the presence of a specific combination of subgroup structures (i.e., the COT) with respect to time. Despite the high frequency of singleton social structures, this allows us to see that there were seldom any instances where the large group was all divided into singletons; rather, singletons co-exist with other subgroup structures. The most common subgroup composition was singleton and coterie, which aligns with the fact that in the exhibit divisions of labour are expected and needed, but there were periods of time when other compositions dominated. By plotting the COT along a timeline, we can get a better sense of how the temperature changes throughout the activity.

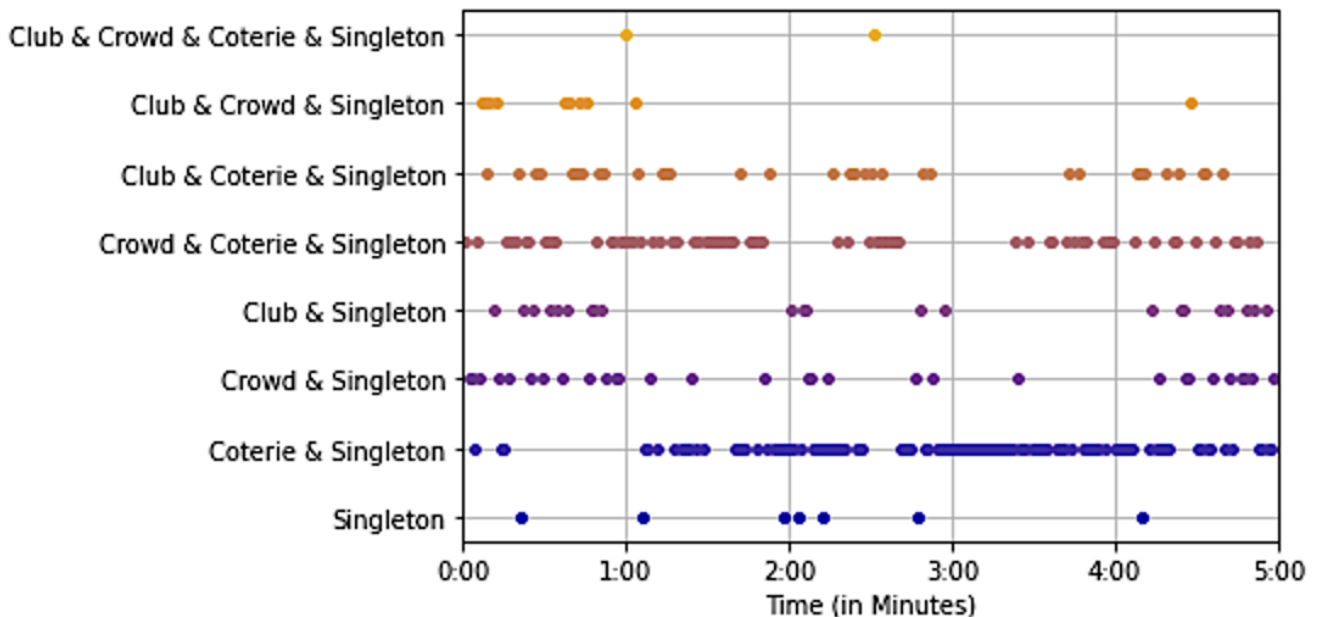


Figure 6. The scatter-plot representation of the feedback condition in Phase 2 (after the intervention) showing the incidence of subgroup compositions over time. Note that the y-axis presents the eight types of subgroup compositions (COTs) seen in our data set. Each dot represents the presence of that particular COT at that instant in time.

5.3 RQ3: Using COT to Complement Qualitative Analyses of Large-Group Dynamics

Our chosen cases (feedback and non-feedback) are useful because we collected a number of other forms of data to complement the COT. On its own, each data source gives an incomplete picture, but triangulating between them reveals a coherent story. For example, an analysis of the log file data shows how the participants in the two cases shared their water across Phases 1 and 2. It reveals that in Phase 1, the groups in both cases were hoarding far more water than they actually needed and weren’t planting much. In Phase 2, the non-feedback group continued hoarding in several biomes and overplanted in the fourth, whereas

the feedback group’s hoarding decreased, and they planted a number of plants proportional to the amount of water, with no overplanting.

While on its own not a clear window into the collaborative dynamics of the two cases, as reported in Mallavarapu and colleagues (2019), this difference in logged outcomes was a direct result of the differences in strategy coordination. The feedback group used the data visualization to inspire a strategy for Phase 2 where they would first split the water among the biomes and then plant using an “each according to its needs” strategy. This plan was enacted by initially coordinating the water distribution and then, after joint agreement (1:40 into the phase), shifting from water distribution to planting activities, and again by attempting to coordinate water distribution at 2:24 (the precise timing is confirmed by both audio records and log files; see Figure 7). By plotting COT over time (see Figure 8), we can see that the non-feedback condition in both phases and the feedback condition in Phase 1 were more or less continuously found to have a singleton and coterie temperature reading (second line from bottom), interleaved with other temperature types. However, in Phase 2, the feedback participants took quite a while (around one minute) to fall back into this pattern of smaller working groups—with the temperatures taken during this time containing a lot of crowds and tightly connected clubs. This coincides with the intense coordination associated with their water distribution plan early in Phase 2, as reported in Mallavarapu and colleagues (2019). Another break in the singleton and coterie temperature reading, and an uptick in the configurations with crowds and clubs, occur at the 2:24 mark, when the visitors began their second water distribution effort. Qualitative dialogue coding sheds further light on this (cf. Levy-Cohen et al., 2021), but we omit this for space.

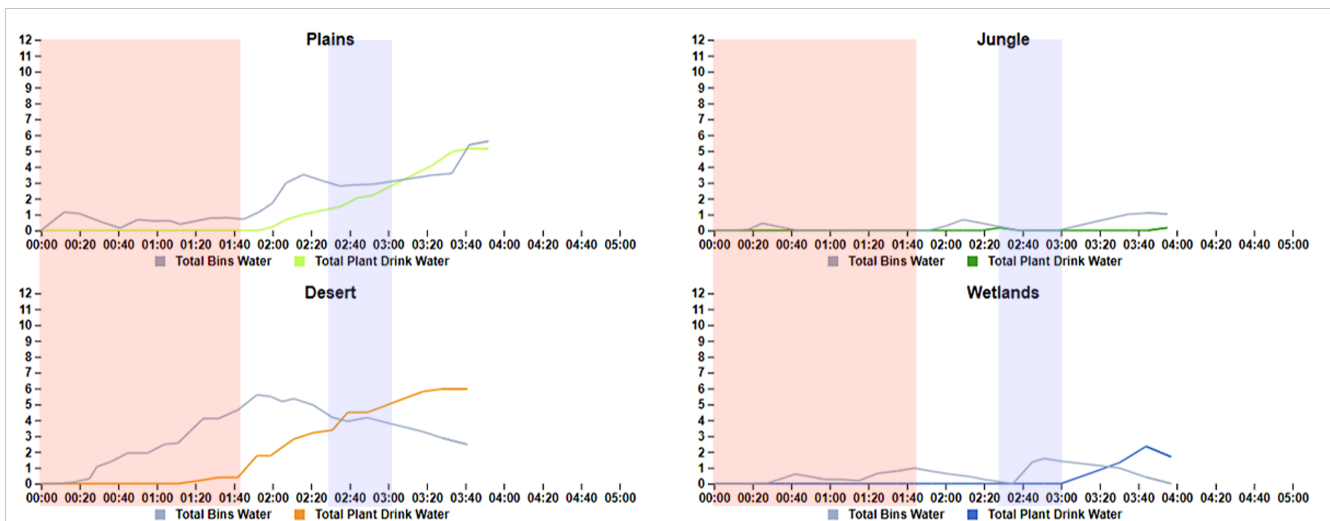


Figure 7. Graphs of water (grey) and plants (light green: plains, dark green: jungle, orange: desert, blue: wetlands) added to the feedback group’s biomes in Phase 2. The red- and blue-shaded regions indicate episodes of shifts in COT that align with the feedback group Phase 2 COT scatter plot.

6. Discussion

We define COT as the composition of social structures present in a large group at any given moment in time. We used COT readings to study the impact of an educational intervention in a large-group immersive learning environment, as a proof of concept of its potential as a quick and easy metric for evaluating the impact that changes in design can have on opportunities for collaboration. By comparing the two cases, we showed that in the case where learners received data-driven feedback about the health of their biomes, the distribution of social subgroups changed, and that by further characterizing those changes as changes in COT readings (over time), we were able to align the measure with other data that could help interpret what forms of collaboration might be taking place. The time series plots of the COT showed how the shifts in social group compositions aligned with the stated strategy of the feedback group, as evidenced by the temporal alignment with plots of the logged activity data. More groups would need to be studied before concluding that the educational intervention was indeed the cause of these shifts, but these results demonstrate the promise of CON and COT for gauging shifts in potential collaborative engagement.

The COT encapsulates multiple hierarchical levels of social behaviour. The CON is constructed from multiple micro-level interactions, which in our case are defined by proximity but which might be defined by other means for other learning environments. The identification of social subgroups within CONs captures the meso-level nuances of the collaboration opportunities in the learning environment. At the macro-level, the COT captures the compositions of these subgroups. Different

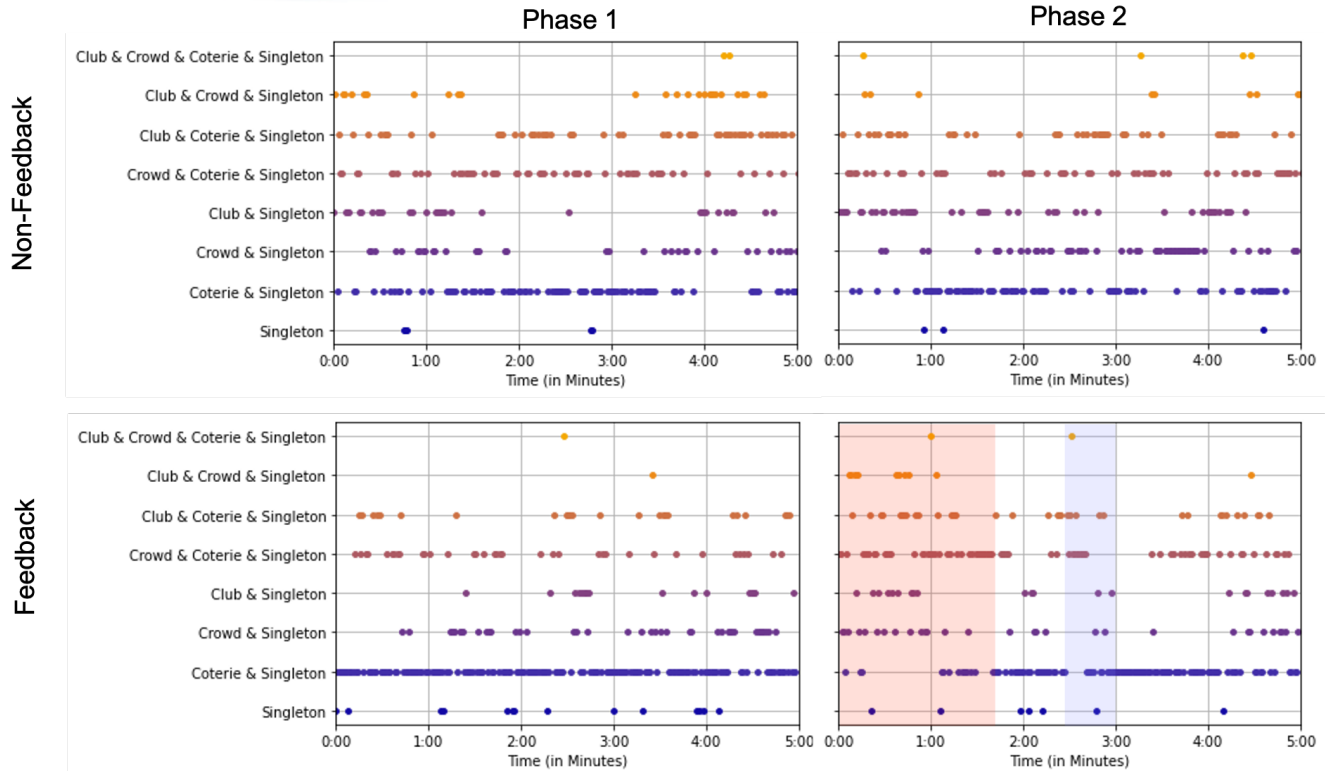


Figure 8. COT readings taken across Phase 1 and Phase 2 of the engagement of the non-feedback and feedback conditions.

affordances are present at each level—at the meso-level, simple comparisons can be made to assess gross changes in the potential for collaboration, which emulates traditional timing and tracking studies, which only make claims about the potential for learning outcomes. At the macro-level, COT series provide affordances to more qualitatively compare the potential impacts of design interventions on collaborative potential and to contextualize the differences through complementary data sources. We next discuss the adaptation of the technique to other settings and its extension to support other uses.

6.1 Generalizability and Extensibility of Collecting COT Proxemics

In environments that would experience lengthier occlusions, we recommend mounting the cameras higher up. If temporarily storing video data is not a problem for some sites (i.e., if it poses no IRB complications), a system that performs real-time tracking like OpenPTrack (Munaro et al., 2014) may produce more valid data, since it is better able to account for occlusions. (Our poor lighting conditions, owing to the dim room and multiple high-lumen projectors, made this colour-based tracking technique unworkable for us, unfortunately.)

6.2 Generalizability of the COT Metrics

Our study held many factors constant and worked with a “best-case” population so that we could vet the potential for COT being useful for designers and explore its potential for researchers. We derived four main types of subgroups (singletons, coterie, crowds, and clubs) via SNA metrics followed by clustering. In our case these types aligned with two descriptive axes, cohesion and coarseness. Each learning environment is likely to have distinct subgroups that vary with the design of the learning environment, the activity design, and the number and types of learners in the large group. The characteristic subgroups discovered for other environments may well promote other social configurations not seen here that may be best characterized by properties other than cohesion and coarseness. We thus propose a procedure of “calibrating” the “collaborative opportunity thermometer” by using clustering algorithms to identify the components of the CONs recorded in the learning environment, controlling for activity design and learner numbers. However, we recommend that researchers use the five suggested SNA metrics to capture the collaborative properties of the social opportunity networks.

6.3 Supporting Exit Interviews with COT

COT readings can be used to inform exit interview questions, allowing designers and researchers to probe learners about what strategies they were using. For example, in our case study, after Phase 2 the researcher could have used semi-structured

interviews to confirm if the detected temperature transitions matched the visitors' perceptions. These questions could be generalized confirmatory probes (e.g., "What proportion of your time did you feel you were working alone?") or more targeted questions seeking rationales (e.g., "It seems that halfway through your session the group conferred with one another quite a bit—do you remember what you were trying to do at that time?"). This would allow the researchers to correlate the visitors' intentions with the automatically detected nuances.

6.4 Using COT to Guide Qualitative Analyses

Qualitative analysis of audio and video recordings of a large group yields a huge amount of data that needs to be coded. Pragmatically speaking, coding and consolidation of results is a time- and effort-intensive task. We argue that COT readings can help researchers focus on specific instances in the large corpus. For example, when we use COT for future data collection, we can use shifts in the readings from larger to tighter groups, or vice versa (like those that occur at 1:40 and 2:24 for the feedback group in Phase 2), to help focus our analysis on what are likely to be critical planning and coordination episodes. In this, we echo other researchers who have successfully used movement metrics to reduce the analysis time of small-group collaboration around a large shared touchscreen (von Zadow & Dachsel, 2017).

7. Limitations

When we began this line of work, we had initially thought to study the evolution of specific sub-communities in large groups over time as in Reda and colleagues (2011), but privacy concerns and technical challenges (changing low-light conditions; the challenges of working within a historical building, which limited our camera placements) led us to pursue low-effort video data collection. We exploited the situation of not being able to track people to instead use SNA to take the COT of the group from moment to moment. The limitation of this is that we lost much of the richness that could be gained by being able to track learners over time, both in terms of what we could distill about their experience (e.g., we would not be able to identify leadership and decision-making behaviours, as has been done successfully with baboons (Strandburg-Peshkin et al., 2015)) and in terms of being able to triangulate data at the individual level. This limitation forces the unit of analysis to be at the level of the group, which can be wholly appropriate for certain research questions about collaboration. However, researchers and evaluators need to understand what questions COT can and cannot be applied to, and the ways in which this is different from how, in cognitive psychology, collaboration has been studied through the lens of individual learners and individual outcomes.

This study is a proof of concept showing two deeply analyzed case studies, highlighting the potential of the SNA-based COT readings for surfacing differences in social behaviours. We acknowledge that the small sample of data is insufficient to claim that the differences between the two cases would represent all groups who might receive either the feedback or the non-feedback condition.

8. Future Work

The conceptualization and use of COT is at its very inception, but it opens up a number of exciting directions. Of course, the first step is to more fully validate the robustness of our conceptual contribution—both our research group and others need to apply COT to other populations and learning environments to see how the approach is affected by differing demographics, activity structures, and physical layouts.

8.1 Interpreting Social Structures in Context

We fully expect that applying our technique to different settings will produce different social structures (i.e., structures beyond singletons, coteries, crowds, and clubs, which might differ along axes other than coarseness and cohesiveness). We hope that the ease of data collection and analysis would allow researchers and evaluators to quickly build up a larger corpus of COT data to permit both within-setting and cross-setting comparisons that would shed light on how learners fluidly gather and break apart while working together in large groups.

We plan to combine our COT data with our qualitative analysis of dialogue once it is completed (Levy-Cohen et al., 2021) to better understand why certain social structures emerged as they did, and what purpose they may have served.

8.2 Incorporating Pose Data into COT

Informal social learning environments have always been a source of legitimate peripheral participation, as defined by Lave and Wenger (1991), which allows learners to learn from observation even when not centrally engaged in the activity. Museums, in particular, see a great deal of learning coming from visitors just watching other visitors (vom Lehn et al., 2001). Future work will explore if pose data can be used to label the roles individuals play in the opportunity networks as either "active engagement" or "passive engagement." We plan to further enhance our definitions of passive engagement actions by detecting the orientation

of learners (i.e., which direction they are facing), information that is available from OpenPose. This would allow us to build on existing theories like F-formations (Kendon, 2010) and distinguish between shared visual grounding opportunities (when visitors are likely attending to the same visual stimulus) and shared conversational engagement opportunities (when visitors are just close enough to converse, but may or may not be looking at the same stimulus), thus developing passive engagement labels with more nuance.

Understanding the ratio of passive to active engagement, and how it is distributed in group formations, is valuable for informing the design of museum exhibits, especially computer-based exhibits. A number of researchers have explored how the size and placement of exhibit components can impact active versus passive shared engagement (Diamond et al., 1995; vom Lehn et al., 2001; Allen & Gutwill, 2004; Heath et al., 2005; Meisner et al., 2007; Lyons, 2009), since exhibit layouts have long been known to be a major factor in individual engagement (Nicks, 2002). COT augmented with active and passive engagement markers could be used to produce “heatmaps” that indicate not just where individuals gather in a learning environment (e.g., Martinez-Maldonado et al., 2018) but which types of groupings form and where the active and passive members within those groupings gather, allowing designers to fine-tune where to place ancillary information like signage, how-to videos, or parallel engagement opportunities that can entice more-passive members to become part of the activity. Which leads us to the next area for future work: incorporating objects into CONs.

8.3 Combining Object Locations with CONs

CONs could benefit by including the object that is driving the social structure as a part of the network (much like the “object-driven sociality” defined by Engeström (2005)). The location, size, and other properties of physical interactives have been shown to influence how groups of learners collectively engage with large-scale immersive learning experiences (Lyons & Mallavarapu, 2021). These object-based affordance networks (Barab & Roth, 2006), when combined spatially with CONs, could be used to generate new temperatures, like collaborative *occupancy* temperature, that could speak to the degree to which groups engage with interactive technology in immersive environments. Some of the social structures we could identify may be the opposite of “collaborative,” since visitors have been shown to position their bodies to prevent other visitors from accessing interactive exhibit components (vom Lehn et al., 2001), but this is extremely useful information for both designers and researchers.

8.4 COT for Formative Feedback

There is a larger trend within learning analytics to allow learners themselves to directly benefit from information in a formative fashion (Ochoa & Wise, 2021). The research project from which we drew our two cases is devoted to producing a mobile dashboard that can be used to deliver data-driven, real-time formative feedback to facilitators and learners (Mallavarapu et al., 2019). The dashboard currently relies on data drawn from the simulation itself, but the social dynamics play a major role in the success or collapse of the simulated ecosystem. Formative feedback that presents only system data is thus impoverished because much of the learning is about how to coordinate actions. We cannot expect learners to think about and learn how to manage their social engagement unless we make social information a more active part of the feedback we supply, in alignment with calls for more “collaborative translucence” (Echeverria et al., 2019).

Anecdotally, some visitor groups have been observed to organically devise their own social strategies, from dictatorships to crisis response teams to representative democracies, and we strongly suspect that information on social structures will further raise their consciousness of coordination strategies. COT data could also be used in real time to highlight opportunities for interactions—oftentimes, passive visitors do not necessarily wish to be disengaged; they have just been shouldered aside (vom Lehn et al., 2001) or feel uncertain about how to make a contribution (Lyons, 2009) and just need an invitation. Facilitators made aware of these scenarios by COT could intervene to help such participants become more active.

COT has the potential to make large-group collaboration more translucent, allowing learners to reflect on the ways they are (or are not) collaborating to improve shared outcomes. Arguably, our society as a whole needs to make managing collaboration a more conscious endeavour, at ever-larger group sizes. This work is a step in that direction.

9. Conclusion

The concept of COT differs from the other ways social configurations have been studied in museums in that it is concerned with representing the overall composition of social subgroups that may emerge in response to the designed learning environment. We argue that these macro-level compositions of meso-level subgroup structures can shed new light on potentials for collaborative engagement. Our conceptualization of COT is as a social analogue to the quick and easy timing and tracking studies museums have used for a century to evaluate and study exhibit designs (Robinson, 1928). We showed COT to be useful for both characterizing and comparing the social subgroup structures that occur in different design variations of an immersive learning environment using paired cases. Further, we demonstrated the potential of COT as a bridge to triangulate across other forms

of data. Another advantage of this low-cost, low-touch method is that it inherently preserves privacy. We do not need to maintain any identifiable information pertaining to learners (e.g., images of their faces or bodies), allowing us to capture and automatically produce collaborative metrics without recording and maintaining video records, as has been shown possible for small-group classroom data collection (Chan et al., 2020). This allows us to obtain implied consent by posting a sign at the entrance indicating that research is in progress, which offers advantages over consent-requiring approaches, where a single non-consenting participant can leave “holes” in the data that violate the interpretation of collective phenomena.

We derived COT from CONs, which are theoretically grounded in Barab and Roth (2006)’s ecological perspective on learning environments. They conceive of learning environments as comprising “affordance networks,” or the “collection[s] of facts, concepts, tools, methods, practices, agendas, commitments, and even people” that support accomplishing an educational goal (Barab & Roth, 2006), but some of these resources are easier to design and supply than others—social resources, in particular, are subject to flux, emerging dynamically during enactment. CONs allow us to define moment-to-moment affordances for collaborative engagement in a learning environment. By creating a method that documents spatially located collaborative affordance networks, we have opened the door to analyses that join such networks with the spatially located physical affordance networks built into the learning environment (Lyons & Mallavarapu, 2021).

We posit that our approach to studying large-group collaboration will be particularly interesting to museum-based researchers, who have a long history of inferring meaning from visitor movements (Yalowitz & Bronnenkant, 2009) and who have already begun using tracking data to study visitor engagement with exhibitions (Shapiro et al., 2017) and small-group collaborative learning (Marshall et al., 2011; Kuflik & Dim, 2013; Dim & Kuflik, 2014). But we also claim that our work could benefit the designers of large-group mixed-reality computer-supported collaborative learning experiences (Smordal et al., 2012; Roberts et al., 2014; Gnoli et al., 2014; Enyedy et al., 2015; Kynigos et al., 2010), both to evaluate and study their designs and to support the formative orchestration of social aspects of the learning experience (Slotta et al., 2013; Matuk, 2016). This work adds to other analytic techniques designed to make the development of large-group computer-supported collaborative learning experiences more efficient (Lyons & Mallavarapu, 2021).

Acknowledgements

This work would not have been possible without the help of Rinat Levy-Cohen and Wren Thompson.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This material is based upon work supported by the National Science Foundation under Grant Nos. 1623094 and 1822864.

References

- Alldieck, T., Magnor, M., Xu, W., Theobalt, C., & Pons-Moll, G. (2018). Video based reconstruction of 3D people models. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 18–22 June 2018, Salt Lake City, Utah, USA (pp. 8387–8397). IEEE. <https://doi.org/10.1109/CVPR.2018.00875>
- Allen, S. (2003). Looking for learning in visitor talk: A methodological exploration. In G. Leinhardt, K. Crowley, & K. Knutson (Eds.), *Learning Conversations in Museums* (pp. 265–309). Routledge. <https://doi.org/10.4324/9781410606624>
- Allen, S., & Gutwill, J. P. (2004). Designing with multiple interactives: Five common pitfalls. *Curator*, 47(2), 199–212. <https://doi.org/10.1111/j.2151-6952.2004.tb00117.x>
- Aviv, R., Erlich, Z., Ravid, G., & Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal of Asynchronous Learning Networks*, 7(3). <https://doi.org/10.24059/olj.v7i3.1842>
- Bakharia, A., & Dawson, S. (2011). SNAPP: A bird’s-eye view of temporal participant interaction. *Proceedings of the First International Conference on Learning Analytics and Knowledge (LAK 2011)*, 27 February–1 March 2011, Banff, Alberta, Canada (pp. 168–173). ACM. <https://doi.org/10.1145/2090116.2090144>
- Barab, S. A., & Roth, W.-M. (2006). Curriculum-based ecosystems: Supporting knowing from an ecological perspective. *Educational Researcher*, 35(5), 3–13. <https://doi.org/10.3102/0013189X035005003>
- Barriault, C., & Pearson, D. (2010). Assessing exhibits for learning in science centers: A practical tool. *Visitor Studies*, 13(1), 90–106. <https://doi.org/10.1080/10645571003618824>
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

- Bitgood, S. (2006). An analysis of visitor circulation: Movement patterns and the general value principle. *Curator: The Museum Journal*, 49(4), 463–475. <https://doi.org/10.1111/j.2151-6952.2006.tb00237.x>
- Block, F., Hammerman, J., Horn, M., Spiegel, A., Christiansen, J., Phillips, B., Diamond, J., Evans, E. M., & Shen, C. (2015). Fluid grouping: Quantifying group engagement around interactive tabletop exhibits in the wild. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015)*, 18–23 April 2015, Seoul, Republic of Korea (pp. 867–876). ACM. <https://doi.org/10.1145/2702123.2702231>
- Cafaro, F., Panella, A., Lyons, L., Roberts, J., & Radinsky, J. (2013). I see you there! Developing identity-preserving embodied interaction for museum exhibits. *Proceedings of the 2013 ACM Annual Conference on Human Factors in Computing Systems (CHI 2013)*, 27 April–2 May 2013, Paris, France (pp. 1911–1920). ACM. <https://doi.org/10.1145/2470654.2466252>
- Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *arXiv:1611.08050*. <https://arxiv.org/abs/1611.08050>
- Carolan, B. V. (2013). *Social Network Analysis and Education: Theory, Methods & Applications*. Sage Publications. <https://doi.org/10.4135/9781452270104>
- Chan, M. C. E., Ochoa, X., & Clarke, D. (2020). Multimodal learning analytics in a laboratory classroom. In M. Virvou, E. Alepis, G. Tsihrintzis, & L. Jain (Eds.), *Machine Learning Paradigms. Intelligent Systems Reference Library* (pp. 131–156). Springer. https://doi.org/10.1007/978-3-030-13743-4_8
- Chua, Y. H. V., Dauwels, J., & Tan, S. C. (2019). Technologies for automated analysis of co-located, real-life, physical learning spaces: Where are we now? *Proceedings of the Ninth International Conference on Learning Analytics and Knowledge (LAK 2019)*, 4–8 March 2019, Tempe, Arizona, USA (pp. 11–20). ACM. <https://doi.org/10.1145/3303772.3303811>
- Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers and Education*, 116, 93–109. <https://doi.org/10.1016/j.compedu.2017.08.007>
- Dado, M., & Bodemer, D. (2017). A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educational Research Review*, 22, 159–180. <https://doi.org/10.1016/j.edurev.2017.08.005>
- D'Angelo, C. M., Smith, J., Alozie, N., Tsiartas, A., Richey, C., & Bratt, H. (2019). Mapping individual to group level collaboration indicators using speech data. In K. Lund, G. P. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *Proceedings of the Computer-Supported Collaborative Learning Conference (CSCL 2019)*, 17–21 June 2019, Lyon, France (pp. 628–631). ISLS. <https://repository.isls.org/handle/1/1637>
- Dawson, S., Bakharia, A., & Heathcote, E. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. *Proceedings of the Seventh International Conference on Networked Learning*, 3–4 May 2010, Aalborg, Denmark (pp. 125–133). <https://www.lancaster.ac.uk/fss/organisations/netlc/past/nlc2010/abstracts/Dawson.html>
- De Laat, M., Lally, V., Lipponen, L., & Simons, R. J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 87–103. <https://doi.org/10.1007/s11412-007-9006-4>
- Derry, S. J., DuRussel, L. A., & O'Donnell, A. M. (1998). Individual and distributed cognitions in interdisciplinary teamwork: A developing case study and emerging theory. *Educational Psychology Review*, 10(1), 25–56. <https://doi.org/10.1023/A:1022806130931>
- Diamond, J., Bond, A., Schenker, B., Meier, D., & Twersky, D. (1995). Collaborative multimedia. *Curator*, 38(3), 137–149. <https://doi.org/10.1111/j.2151-6952.1995.tb01050.x>
- Diamond, J., Luke, J. J., & Uttal, D. H. (2009). *Practical Evaluation Guide: Tools for Museums & Other Informal Educational Settings* (2nd ed.). AltaMira Press.
- Dim, E., & Kuflik, T. (2014). Automatic detection of social behavior of museum visitor pairs. *ACM Transactions on Interactive Intelligent Systems*, 4(4), 1–30. <https://doi.org/10.1145/2662869>
- D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, 52(2), 104–123. <https://doi.org/10.1080/00461520.2017.1281747>
- Dráždilová, P., Martinovič, J., Slaninová, K., & Snášel, V. (2008). Analysis of relations in eLearning. *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology—Workshops, WI-IAT Workshops 2008*, 9–12 December 2008, Sydney, Australia (pp. 373–376). IEEE. <https://doi.org/10.1109/WIIAT.2008.250>
- Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, 18, 7–22. <https://doi.org/10.1007/s10956-008-9119-1>

- Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), 15274–15278. <https://doi.org/10.1073/pnas.0900282106>
- Echeverria, V., Martinez-Maldonado, R., & Buckingham Shum, S. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019)*, 4–9 May 2019, Glasgow, UK (pp. 1–16). ACM. <https://doi.org/10.1145/3290605.3300269>
- Eghbal-Azar, K., & Widlok, T. (2013). Potentials and limitations of mobile eye tracking in visitor studies: Evidence from field research at two museum exhibitions in Germany. *Social Science Computer Review*, 31(1), 103–118. <https://doi.org/10.1177/0894439312453565>
- Engeström, J. (2005). Why some social network services work and others don't—Or: The case for object-centered sociality. <http://www.zengestrom.com/blog/2005/04/why-some-social-network-services-work-and-others-dont-or-the-case-for-object-centered-sociality.html>
- 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. <https://doi.org/10.1007/s11412-015-9207-1>
- Erlin, B., Yusof, N., & Rahman, A. A. (2009). Students' interactions in online asynchronous discussion forum: A social network analysis. *Proceedings of the 2009 International Conference on Education Technology and Computer (ICETC 2009)*, 17–20 April 2009, Singapore (pp. 25–29). IEEE. <https://doi.org/10.1109/ICETC.2009.48>
- Falk, J. H., & Dierking, L. D. (2018). *Learning from Museums: Visitor Experiences and the Making of Meaning* (2nd ed.). Rowman & Littlefield.
- Filippini Fantoni, S., Jaebker, K., Bauer, D., & Stofer, K. (2013). Capturing visitors' gazes: Three eye tracking studies in museums. In N. Proctor & R. Cherry (Eds.), *Proceedings of the 2013 Annual Conference of Museums and the Web (MW2013)*, 17–20 April 2013, Portland, Oregon, USA. Museums and the Web LLC. <https://mw2013.museumsandtheweb.com/paper/capturing-visitors-gazes-three-eye-tracking-studies-in-museums/>
- Fisher, D., & Dourish, P. (2004). Social and temporal structures in everyday collaboration. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2004)*, 24–29 April 2004, Vienna, Austria (pp. 551–558). ACM. <https://doi.org/10.1145/985692.985762>
- Gergle, D., Rosé, C. P., & Kraut, R. E. (2007). Modeling the impact of shared visual information on collaborative reference. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, 28 April–3 May 2007, San Jose, California, USA (pp. 1543–1552). ACM. <https://doi.org/10.1145/1240624.1240858>
- Gnoli, A., Perritano, A., Guerra, P., Lopez, B., Brown, J., & Moher, T. (2014). Back to the future: Embodied classroom simulations of animal foraging. *Proceedings of the Eighth International Conference on Tangible, Embedded and Embodied Interaction (TEI 2014)*, 16–19 February 2014, Munich, Germany (pp. 275–282). ACM. <https://doi.org/10.1145/2540930.2540972>
- Greenberg, S., Boring, S., Vermeulen, J., & Dostal, J. (2014). Dark patterns in proxemic interactions: A critical perspective. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS 2014)*, 21–25 June 2014, Vancouver, British Columbia, Canada (pp. 523–532). <https://doi.org/10.1145/2598510.2598541>
- Hagberg, A., Swart, P., & S Chult, D. (2008). *Exploring Network Structure, Dynamics, and Function Using NetworkX* (tech. rep.). Los Alamos National Laboratory (LANL), Los Alamos, New Mexico, USA. <https://www.osti.gov/biblio/960616-exploring-network-structure-dynamics-function-using-networkx>
- Hall, E. T. (1966). *The Hidden Dimension*. Anchor Books.
- Haythornthwaite, C. (1999). Collaborative work networks among distributed learners. *Proceedings of the Hawaii International Conference on System Sciences (HICSS 1999)*, 5–8 January 1999, Maui, Hawaii, USA. IEEE. <https://doi.org/10.1109/hicss.1999.772707>
- 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. <https://doi.org/10.1177/0963662505047343>
- Hecking, T., Ziebarth, S., & Ulrich Hoppe, H. (2014). Analysis of dynamic resource access patterns in a blended learning course. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge (LAK 2014)*, 24–28 March 2014, Indianapolis, Indiana, USA (pp. 173–182). ACM. <https://doi.org/10.1145/2567574.2567584>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv:1704.04861*. <https://arxiv.org/abs/1704.04861>

- Intille, S. S., & Bobick, A. F. (1995). Closed-world tracking. *Proceedings of the IEEE International Conference on Computer Vision (ICCV 1995)*, 20–23 June 1995, Cambridge, Massachusetts, USA (pp. 672–678). IEEE. <https://doi.org/10.1109/ICCV.1995.466874>
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49(4), 1037–1065. <https://doi.org/10.1016/j.compedu.2006.01.004>
- Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, 38(1), 59–88. <https://doi.org/10.1007/s11251-008-9078-1>
- Jorion, N., Roberts, J., Bowers, A., Tissenbaum, M., Lyons, L., Kumar, V., & Berland, M. (2020). Uncovering patterns in constructionist collaborative learning activities via cluster analysis of museum exhibit log files. *Frontline Learning Research*, 8(6), 77–87. <https://doi.org/10.14786/flr.v8i6.597>
- Kendon, A. (2010). Spacing and orientation in co-present interaction. In A. Esposito, N. Campbell, C. Vogel, A. Hussain, & A. Nijholt (Eds.), *Development of Multimodal Interfaces: Active Listening and Synchrony. Lecture Notes in Computer Science* (pp. 1–15). Springer-Verlag. https://doi.org/10.1007/978-3-642-12397-9_1
- Kim, T., McFee, E., Olguin, D. O., Waber, B., & Pentland, A. (2012). Sociometric badges: Using sensor technology to capture new forms of collaboration. *Journal of Organizational Behavior*, 33(3), 412–427. <https://doi.org/10.1002/job.1776>
- Kim, Y., Butail, S., Liu, L., Tscholl, M., Hwang, J., Cafaro, F., Trajkova, M., Kwon, K., Espino, D., Lee, S., Hamilton, E., D'Angelo, C., Ochoa, X., & Kline, A. (2020). Multimodal data analytics for assessing collaborative interactions. In M. Gresalfi & I. S. Horn (Eds.), *Proceedings of the 14th International Conference of the Learning Sciences (ICLS 2020)*, 19–23 June 2020, Nashville, Tennessee, USA (pp. 2547–2554). ISLS. <https://repository.isls.org/handle/1/6619>
- Korn, R. (1995). An analysis of differences between visitors at natural history museums and science centers. *Curator: The Museum Journal*, 38(3), 150–160. <https://doi.org/10.1111/j.2151-6952.1995.tb01051.x>
- Kufflik, T., & Dim, E. (2013). Early detection of pairs of visitors by using a museum triage. In N. Proctor & R. Cherry (Eds.), *Proceedings of the 2013 Annual Conference of Museums and the Web (MW2013)*, 17–20 April 2013, Portland, Oregon, USA. Museums and the Web LLC. <https://mw2013.museumsandtheweb.com/paper/early-detection-of-pairs-of-visitors-by-using-a-museum-triage/>
- Kynigos, C., Smyrniou, Z., & Roussou, M. (2010). Exploring rules and underlying concepts while engaged with collaborative full-body games. *Proceedings of the Ninth International Conference on Interaction Design and Children (IDC 2010)*, 9–12 June 2010, Barcelona, Spain (pp. 222–225). ACM. <https://doi.org/10.1145/1810543.1810576>
- Lanir, J., Kufflik, T., Sheidin, J., Yavin, N., Leiderman, K., & Segal, M. (2017). Visualizing museum visitors' behavior: Where do they go and what do they do there? *Personal and Ubiquitous Computing*, 21(2), 313–326. <https://doi.org/10.1007/s00779-016-0994-9>
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>
- Lepri, B., Staiano, J., Rigato, G., Kalimeri, K., Finnerty, A., Pianesi, F., Sebe, N., & Pentland, A. (2012). The sociometric badges corpus: A multilevel behavioral dataset for social behavior in complex organizations. *Proceedings of the 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust and 2012 ASE/IEEE International Conference on Social Computing (SocialCom/PASSAT 2012)*, 3–5 September 2012, Amsterdam, Netherlands (pp. 623–628). IEEE. <https://doi.org/10.1109/SocialCom-PASSAT.2012.71>
- Levy-Cohen, R., Mallavarapu, A., Lyons, L., & Uzzo, S. (2021). Studying shared regulation in immersive learning environments. In C. Hmelo-Silver, B. de Wever, & J. Oshima (Eds.), *Proceedings of the 14th International Conference on Computer-Supported Collaborative Learning (CSCL 2021)*, 8–11 June 2021, Bochum, Germany (pp. 100–115). ISLS. <https://repository.isls.org/handle/1/7349>
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Lyons, L. (2009). Designing opportunistic user interfaces to support a collaborative museum exhibit. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Proceedings of the Ninth International Conference on Computer Supported Collaborative Learning—Volume 1 (CSCL 2009)*, 8–13 June 2009, Rhodes, Greece (pp. 375–384). ISLS. <https://doi.org/10.3115/1600053.1600110>
- Lyons, L., & Mallavarapu, A. (2021). Collective usability: Using simulation tools to explore embodied design challenges in immersive, shared mixed-reality experiences. *Educational Technology & Society*, 24(2), 120–135. https://drive.google.com/file/d/1_ZxK1eCjOKqyjcd3_UyXlC6mH2Jg7s2/view
- Lyons, L., Tissenbaum, M., Berland, M., Eydt, R., Wielgus, L., & Mechtley, A. (2015). Designing visible engineering: Supporting tinkering performances in museums. *Proceedings of the 14th International Conference on Interaction*

- Design and Children* (IDC 2015), 21–24 June 2015, Boston, Massachusetts, USA (pp. 49–58). ACM. <https://doi.org/10.1145/2771839.2771845>
- Magnussen, R., Zachariassen, M., Kharlamov, N., Larsen, B., Kharlamov, N., & Larsen, B. (2017). Mobile eye-tracking methods in studies of audience learning in health promotion exhibitions. *Electronic Journal of E-Learning*, 15(1), 46–58. <https://academic-publishing.org/index.php/ejel/article/view/1818>
- Malcolm, P., Moher, T., Bhatt, D., Uphoff, B., & López-Silva, B. (2008). Embodying scientific concepts in the physical space of the classroom. *Proceedings of the Seventh International Conference on Interaction Design and Children* (IDC 2008), 11–13 June 2008, Chicago, Illinois, USA (pp. 234–241). ACM. <https://doi.org/10.1145/1463689.1463761>
- Mallavarapu, A., Lyons, L., Uzzo, S., Thompson, W., Levy-Cohen, R., & Slattery, B. (2019). Connect-to-connected worlds: Piloting a mobile, data-driven reflection tool for an open-ended simulation at a museum. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI 2019), 4–9 May 2019, Glasgow, UK (pp. 1–14). ACM. <https://doi.org/10.1145/3290605.3300237>
- Mansur, A. B. F., Yusof, N., & Othman, Mohd, S. (2012). Analysis of social learning network for wiki in moodle E-Learning. In H. Choi, S. Hong, & F. Ko (Eds.), *Proceedings of the Fourth International Conference on Information Sciences and Interaction Sciences* (ICIS 2011), 16–18 August 2011, Busan, Korea (pp. 201–209). IEEE. <https://ieeexplore.ieee.org/document/6014520>
- Marshall, P., Rogers, Y., & Pantidi, N. (2011). Using F-formations to analyse spatial patterns of interaction in physical environments. *Proceedings of the 2011 ACM Conference on Computer Supported Cooperative Work* (CSCW 2011), 19–23 March 2011, Hangzhou, China (pp. 445–454). ACM. <https://doi.org/10.1145/1958824.1958893>
- Martínez, R., Collins, A., Kay, J., & Yacef, K. (2011). Who did what? Who said that? Collaid: An environment for capturing traces of collaborative learning at the tabletop. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (ITS 2011), 13–16 November 2011, Kobe, Japan (pp. 172–181). ACM. <https://doi.org/10.1145/2076354.2076387>
- Martinez-Maldonado, R., Echeverria, V., Santos, O. C., Dos Santos, A. D. P., & Yacef, K. (2018). Physical learning analytics: A multimodal perspective. *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 375–379). ACM. <https://doi.org/10.1145/3170358.3170379>
- Martinez-Maldonado, R., Echeverria, V., Yacef, K., Dos Santos, A. D. P., & Pechenizkiy, M. (2017). How to capitalise on mobility, proximity and motion analytics to support formal and informal education? *Joint Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop and the Second Cross-LAK Workshop co-located with Seventh International Learning Analytics and Knowledge Conference* (LAK 2017), 14 March 2017, Vancouver, British Columbia, Canada (pp. 39–46). ACM. <http://ceur-ws.org/Vol-1828/paper-06.pdf>
- Matuk, C. (2016). The learning affordances of augmented reality for museum exhibits on human health. *Museums & Social Issues: A Journal of Reflective Discourse*, 11(1), 73–87. <https://doi.org/10.1080/15596893.2016.1142815>
- Meisner, R., vom Lehn, D., Heath, C., Burch, A., Gammon, B., & Reisman, M. (2007). Exhibiting performance: Co-participation in science centres and museums. *International Journal of Science Education*, 29(12), 1531–1555. <https://doi.org/10.1080/09500690701494050>
- Moher, T. (2008). Learning and participation in a persistent whole-classroom seismology simulation. *Proceedings of the 2008 Computer-Supported Collaborative Learning Conference, Part 2* (CSCL 2008), 24–28 June 2008, Utrecht, Netherlands (pp. 82–90). ACM. <https://dl.acm.org/doi/10.5555/1599871.1599882>
- Moussouri, T., & Roussos, G. (2013). Examining the effect of visitor motivation on observed visit strategies using mobile computer technologies. *Visitor Studies*, 16(1), 21–38. <https://doi.org/10.1080/10645578.2013.767732>
- Moussouri, T., & Roussos, G. (2015). Conducting visitor studies using smartphone-based location sensing. *Journal on Computing and Cultural Heritage*, 8(3), 1–16. <https://doi.org/10.1145/2677083>
- Munaro, M., Horn, A., Illum, R., Burke, J., & Rusu, R. B. (2014). OpenPTrack: People tracking for heterogeneous networks of color-depth cameras. *IAS-13 Workshop Proceedings: First International Workshop on 3D Robot Perception with Point Cloud Library*, 15–19 July 2014, Padova, Italy (pp. 1–13). http://www.dei.unipd.it/~munaro/media/Munaro_3DRP2014.pdf
- National Research Council. (2010). *Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary*. The National Academies Press. <https://doi.org/10.17226/12771>
- National Research Council. (2011). *Assessing 21st Century Skills: Summary of a Workshop*. The National Academies Press. <https://doi.org/10.17226/13215>
- Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), 8577–8582. <https://doi.org/10.1073/pnas.0601602103>

- Nicks, J. (2002). Curatorship in the exhibition planning process. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 345–347, 353–372). AltaMira Press.
- Norman, D. (1988). *The Design of Everyday Things*. Currency-Doubleday.
- Ochoa, X., & Wise, A. F. (2021). Supporting the shift to digital with student-centered learning analytics. *Educational Technology Research and Development*, 69(1), 357–361. <https://doi.org/10.1007/s11423-020-09882-2>
- Paullada, A., Raji, I. D., Bender, E. M., Denton, E., & Hanna, A. (2021). Data and its (dis)contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11), 100336. <https://doi.org/10.1016/j.patter.2021.100336>
- Peppler, K., Danish, J., Zaitlen, B., Glosson, D., Jacobs, A., & Phelps, D. (2010). BeeSim: Leveraging wearable computers in participatory simulations with young children. *Proceedings of the Ninth International Conference on Interaction Design and Children (IDC 2010)*, 9–12 June 2010, Barcelona, Spain (pp. 246–249). ACM. <https://doi.org/10.1145/1810543.1810582>
- Planey, J., & Lindgren, R. (2018). Embodying climate change: Incorporating full body tracking in the design of an interactive rates of change greenhouse gas simulation. In D. Beck, C. Allison, L. Morgado, J. Pirker, A. Peña-Rios, T. Ogle, J. Richter, & C. Gütl (Eds.), *Proceedings of the Fourth International Conference on Immersive Learning (iLRN 2018)*, 24–29 June 2018, Missoula, Montana, USA (pp. 23–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-93596-6_2
- Poquet, O., & Jovanovic, J. (2020). Intergroup and interpersonal forum positioning in shared-thread and post-reply networks. *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 187–196). ACM. <https://doi.org/10.1145/3375462.3375533>
- Povis, K. T., & Crowley, K. J. (2015). Family learning in object-based museums: The role of joint attention. *Visitor Studies*, 18(2), 168–182. <https://doi.org/10.1080/10645578.2015.1079095>
- Puntambekar, S., Erkens, G., & Hmelo-Silver, C. (2011). *Analyzing Interactions in CSCL: Methods, Approaches and Issues* (Vol. 12). Springer Science & Business Media. <https://doi.org/10.1007/978-1-4419-7710-6>
- Rabbany, R., Elatia, S., Takaffoli, M., & Zaïane, O. R. (2014). Collaborative learning of students in online discussion forums: A social network analysis perspective. In A. Peña-Ayala (Ed.), *Studies in Computational Intelligence* (pp. 441–466). Springer International Publishing. https://doi.org/10.1007/978-3-319-02738-8_16
- Reda, K., Tantipathananandh, C., Johnson, A., Leigh, J., & Berger-Wolf, T. (2011). Visualizing the evolution of community structures in dynamic social networks. *Computer Graphics Forum*, 30(3), 1061–1070. <https://doi.org/10.1111/j.1467-8659.2011.01955.x>
- Reilly, J. M., Ravenell, M., & Schneider, B. (2018). Exploring collaboration using motion sensors and multi-modal learning analytics. *Proceedings of the 11th International Conference on Educational Data Mining (EDM 2018)*, 15–18 July 2018, Buffalo, New York, USA. https://educationaldatamining.org/files/conferences/EDM2018/papers/EDM2018_paper.113.pdf
- Roberts, J., & Lyons, L. (2017). The value of learning talk: Applying a novel dialogue scoring method to inform interaction design in an open-ended, embodied museum exhibit. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 343–376. <https://doi.org/10.1007/s11412-017-9262-x>
- Roberts, J., Lyons, L., Cafaro, F., & Eydt, R. (2014). Interpreting data from within: Supporting human-data interaction in museum exhibits through perspective taking. *Proceedings of the 2014 Conference on Interaction Design and Children (IDC 2014)*, 17–20 June 2014, Aarhus, Denmark (pp. 7–16). ACM. <https://doi.org/10.1145/2593968.2593974>
- Robinson, E. (1928). *The Behavior of the Museum Visitor*. American Association of Museums.
- Rousseeuw, P. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(1), 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Schautz, A. M., van Dijk, E. M., & Meisert, A. (2016). The use of audio guides to collect individualized timing and tracking data in a science center exhibition. *Visitor Studies*, 19(1), 96–116. <https://doi.org/10.1080/10645578.2016.1144032>
- Schneider, B., & Blikstein, P. (2014). Unraveling students' interaction around a tangible interface using gesture recognition. In J. C. Stamper, Z. A. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the Seventh International Conference on Educational Data Mining (EDM 2014)*, 4–7 July 2014, London, UK (pp. 320–323). https://educationaldatamining.org/EDM2014/uploads/procs2014/short%20papers/320_EDM-2014-Short.pdf
- Schneider, B., & Blikstein, P. (2015). Unraveling students' interaction around a tangible interface using multimodal learning analytics. *Journal of Educational Data Mining*, 7(3), 89–116. <https://doi.org/10.5281/zenodo.3554729>
- Schneider, B., Dowell, N., & Thompson, K. (2021). Collaboration analytics—Current state and potential futures. *Journal of Learning Analytics*, 8(1), 1–12. <https://doi.org/10.18608/jla.2021.7447>

- Serrell, B. (1997). Paying attention: The duration and allocation of visitors' time in museum exhibitions. *Curator*, 40(2), 108–124. <https://doi.org/10.1111/j.2151-6952.1997.tb01292.x>
- Shapiro, B. R., Hall, R. P., & Owens, D. A. (2017). Developing & using interaction geography in a museum. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 377–399. <https://doi.org/10.1007/s11412-017-9264-8>
- Slotta, J. D., Tissenbaum, M., & Lui, M. (2013). Orchestrating of complex inquiry: Three roles for learning analytics in a smart classroom infrastructure. *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK 2013)*, 8–13 April 2013, Leuven, Belgium (pp. 270–274). ACM. <https://doi.org/10.1145/2460296.2460352>
- Smordal, O., Slotta, J., Moher, T., Lui, M., & Jornet, A. (2012). Hybrid spaces for science learning: New demands and opportunities for research. *Symposium for the 2012 International Conference of the Learning Sciences (ICLS 2012)*, 2–6 July 2012, Sydney, Australia (pp. 9–16). ISLS. https://www.researchgate.net/publication/233416697_Hybrid_spaces_for_science_learning_New_demands_and_opportunities_for_research
- Stake, R. E. (2013). *Multiple Case Study Analysis*. Guilford Press.
- Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., & Crofoot, M. C. (2015). Shared decision-making drives collective movement in wild baboons. *Science*, 348(6241), 1358–1361. <https://doi.org/10.1126/science.aaa5099>
- Suthers, D. D., Vatrappu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Computers and Education*, 50(4), 1103–1127. <https://doi.org/10.1016/j.compedu.2006.10.007>
- Thirde, D., Xu, M., & Orwell, J. (2006). Tracking football players with multiple cameras. In S. A. Velastin & P. Remagnino (Eds.), *Intelligent Distributed Video Surveillance Systems* (pp. 225–252). The Institution of Engineering and Technology. https://doi.org/10.1049/PBPC005E_ch8
- Tissenbaum, M., Kumar, V., & Berland, M. (2016). Modeling visitor behavior in a game-based engineering museum exhibit with hidden Markov models. *Proceedings of the Ninth International Conference on Educational Data Mining (EDM 2016)*, 29 June–2 July 2016, Raleigh, North Carolina, USA (pp. 517–522). https://educationaldatamining.org/EDM2016/proceedings/paper_104.pdf
- Tissenbaum, M., Berland, M., & Lyons, L. (2017). DCLM framework: Understanding collaboration in open-ended tabletop learning environments. *International Journal of Computer-Supported Collaborative Learning*, 12(1), 35–64. <https://doi.org/10.1007/s11412-017-9249-7>
- Tscholl, M., Lindgren, R., & Johnson, E. (2013). Enacting orbits: Refining the design of a full-body learning simulation. *Proceedings of the 12th International Conference on Interaction Design and Children (IDC 2013)*, 24–27 June 2013, New York (pp. 451–454). ACM. <https://doi.org/10.1145/2485760.2485807>
- vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic Interaction*, 24(2), 189–216. <https://doi.org/10.1525/si.2001.24.2.189>
- von Zadow, U., & Dachsel, R. (2017). GIANt: Visualizing group interaction at large wall displays. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI 2017)*, 6–11 May 2017, Denver, Colorado, USA (pp. 2639–2647). ACM. <https://doi.org/10.1145/3025453.3026006>
- Wang, M., Cheng, B., Chen, J., Mercer, N., & Kirschner, P. A. (2017). The use of web-based collaborative concept mapping to support group learning and interaction in an online environment. *The Internet and Higher Education*, 34, 28–40. <https://doi.org/10.1016/j.iheduc.2017.04.003>
- Wasserman, S., & Faust, K. (2012). *Social Network Analysis: Methods and Applications* (Vol. 8). Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
- Willging, P. A. (2005). Using social network analysis techniques to examine online interactions. *US-China Education Review*, 2(9), 1548–6613. <https://doi.org/10.13140/2.1.2643.4242>
- Wu, L., Waber, B., Aral, S., Brynjolfsson, E., & Pentland, A. S. (2008). Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an IT configuration task. *Proceedings of the 29th International Conference on Information Systems (ICIS 2008)*, 14–17 December 2008, Paris, France. SSRN. <https://doi.org/10.2139/ssrn.1130251>
- Xu, M., Orwell, J., & Jones, G. (2004). Tracking football players with multiple cameras. *Proceedings of the 2004 International Conference on Image Processing (ICIP 2004)*, 24–27 October 2004, Singapore (pp. 2909–2912). IEEE. <https://doi.org/10.1109/ICIP.2004.1421721>
- Yalowitz, S. S., & Bronnenkant, K. (2009). Timing and tracking: Unlocking visitor behavior. *Visitor Studies*, 12(1), 47–64. <https://doi.org/10.1080/10645570902769134>
- Yan, L., Martinez-Maldonado, R., Cordoba, B. G., Deppeler, J., Corrigan, D., Nieto, G. F., & Gasevic, D. (2021). Footprints at school: Modelling in-class social dynamics from students' physical positioning traces. *Proceedings of the 11th*

International Conference on Learning Analytics and Knowledge (LAK 2021), 12–16 April 2021, Irvine, California, USA (pp. 43–54). ACM. <https://doi.org/10.1145/3448139.3448144>

Zhang, J., Zhu, Z., Zou, W., Li, P., Li, Y., Su, H., & Huang, G. (2019). FastPose: Towards real-time pose estimation and tracking via scale-normalized multi-task networks. *arXiv:1908.05593*. <https://arxiv.org/abs/1908.05593>