Study limitations:

- ➤ We don't have a theoretical underpinning / data aligned with a research question
- ➤ We don't have a hypothesis / experimental groups aligned with data analytics questions
- ➤ We don't have high quality data
- ➤ We have very little time and resources
- ➤ We have little domain knowledge
- ➤ We don't have experience / significant technical knowledge in the specific technical area: person / body / movement recognition
- ➤ We don't have capacity for significant technical tests
- ➤ We don't have an easy ground truth (manual coding?)

3 types of papers for lit review:
- domain: children free play; children body & movement
- data analysis in domain
- openPose data analysis

- Body Movement and Inattention in Learning-Disable and Normal Children

Robert P. Rugel, Douglas Cheatam and Annette Mitchell

Fiske and Maddi (1961), Berlyne (1960), and Schultz (1965) all suggest that organisms experience a subjective state that can be described along a boredom-excitation continuum.

Raters unaware of group membership observed the subjects' movement at 1-minute intervals during the vigilance task. Six categories of movement were rated: movement of right and left arms (including hands), movement of fight and left legs (including feet), head movement, and trunk movement.
If the actual movement the subject went through was more than 25% of the potential range, then the movement was considered "large." If the actual movement was less than 25% the movement was considered moderate. "No movement" was also recorded. Two 1 or 0 points were assigned to "large," "moderate," or "no movement," respectively, Total body activity per observation was the sum of the points in the six categories. Mean body movement levels during each quarter of the vigilance task were obtained by averaging the observations made during that quarter.

The results suggested that body movement increased throughout the vigilance task, inereased rates of external stimulation result in decreased level of body movement, and learning-disabled children differed from controls in showing higher levels of body movement and poorer vigilance performance.

- Multimodal Learning Analytics research with young children: a systematic review

Lucrezia Crescenzi-Lanna

Kinetic analytics

Students' position and gesture recognition (e.g. hand and wrist movements) are usually studied using motion-based technologies and moving or infrared cameras worn by them. With few exceptions (Junokas, Lindgren, Kang & Morphew, 2018), in the analysed literature, motion and gesture movement are studied using Kinect, a range of devices produced by Microsoft that employ skeletal tracking and wearable "fiducial markers" to record gestures. As a result, according to Kosmas, Ioannou and Retalis (2017), the analytics associated with these devices have come to be called "kinetic analytics".

The Kinect system offers a very limited set of games (called "Kinems suite"), a selection of which were used in several MMLA studies (Spikol et al., 2018; Worsley, 2018; Kourakli, Altanis, Retalis, Boloudakis, Zbainos & Antonopoulou, 2017). In addition, Kourakli and colleagues (2017) conducted a study that included twenty Greek children between six and eleven years of age who had special educational needs, concluding that Kinect users have to be high-functioning in order to interact with the Kinems games.

These limits pose a restriction for applied research exploring "the potential of engaging the body in the learning process" (Kosmas, Ioannou & Retalis, 2017, p. 121). Overcoming this constraint should be a priority, since embodied learning is a central part of early childhood education.

Existing literature on embodied cognition (EC) and embodied learning shows promising effects of bodily engagement and movement on children's cognitive and academic outcomes. Embodied learning appears as a multimodal and playful process that requires the involvement of the human body in the cognitive process (Foglia & Wilson, 2013; Wilson, 2002).

In that way, as Atkinson (2010) states, "we experience, understand, and act on the world through our bodies."

According to Lindgren and Johnson-Glenberg (2013), the primary principles of the implementation of embodied learning are the following: the sensorimotor activity, the relevance of gestures to the theme that is to be reproduced, and the emotional involvement of participant in the whole process.

- Using Embodied Learning Technology to Advance Motor Performance of Children with Special Educational Needs and Motor Impairments

Panagiotis Kosmas, Andri Ioannou, and Symeon Retalis

Gains in (a) Psychomotor Abilities (Gp) - the ability to perform physical body motor movements with precision, coordination, or strength, and (b) Psychomotor Speed (Gps) - the speed and fluidity with which physical body movements can be made, based on the Cattell-Horn-Carroll Integrated Model classification of skills, which is widely accepted as the most comprehensive and empirically supported model of cognitive abilities [6].
Psychomotor Ability (Gp) is the ability to perform physical body motor movements with precision, coordination, or strength, operationalized in this study as the motor stability of the hand. Psychomotor Speed (Gps) is the speed and fluidity with which physical body movements can be made, operationalized in this work as the time for successful completion of the task.
(Visual-motor coordination; Hand stability; Speed improvement in task completion)

Embodied cognition has become a significant learning paradigm in contemporary theory of cognitive sciences. The fascinating insight of this theory is that behavior is not simply the output of someone's isolated brain [7]. Rather, embodied cognition holds that cognitive processes are deeply rooted in the body's interactions with the world [1]. Consequently, the body plays a central role in shaping the mind and therefore, learning scientists should consider ways of engaging the body in the learning activity [8], known as embodied learning. As Nguyen, and Larson [9] noted, in embodied learning environments where learners use their bodies "learners are simultaneously sensorimotor bodies, reflective minds, and social beings".

- Does body movement engage you more in digital game play? And Why?

Nadia Bianchi-Berthouze, Whan Woong Kim, Darshak Patel

Whilst this recent trend suggests that game designers expect these new consoles to result in more intuitive and natural games, engagement is still a novel area in game research, and the relationship between engagement and body movement has not been studied. The degree of involvement in technology is currently described using a variety of terms: immersion, engagement, presence or fun, to name just a few. The concept of presence or immersion, and their measurement, has mostly been studied in the context of virtual environments (e.g., [1-5]).

Finally, engagement was also described in terms of three categories: participation, narration and co-presence of others, thus stating the social aspect of engagement [12]. We would like to stress the fact that most theories of engagement have focused purely on its mental aspects. Tellingly, Koster [13] defined “fun” as the act of mastering the game mentally. One aim of this paper is to suggest that body movements should play an important role in engagement.

In recent years, however, this idea has been questioned by psychology studies showing body posture to be a very good indicator for certain categories of emotions, see [16-19]) for examples. And accordingly, recent studies (see [20-25] for some examples) have set to establish a framework for grounding the recognition of affective states into body postures. Our own studies, in particular, have proposed a general description of posture based on angles and distances between body joints and used it to create an affective posture recognition system that maps the set of postural descriptors into affective categories. In addition to classification rates that favourably compared with those obtained using facial expressions, we also showed how posture could provide for the discrimination of affective categories across cultures [26], thus showing posture as a very powerful communicative modality.

But interestingly, another line of work suggests another important role of body posture. And that is that changes in posture can induce changes in affective states or have a feedback role affecting motivation and emotion. A study by Riskind and Gotay [27], for example, revealed how “subjects who had been temporarily placed in a slumped, depressed physical posture later appeared to develop helplessness more readily, as assessed by their lack of persistence in a standard learned helplessness task, than did subjects who had been placed in an expansive, upright posture.”

These two facets of bodily activity in general (posture, movement) provide the theoretical justification for our hypothesis on the existence of a (possibly bilateral) relationship between engagement and body movement. The question we specifically address here is whether an increase in task-related body movement imposed, or allowed, by the game controller will result in an increase of the player’s engagement level.

The participants were fitted with a lightweight (6kgs) exoskeleton -- GIPSY by Animazoo (UK) -- on their upper body, arms and head, so as to provide angular measurements for each of the upper-body joints. In addition, a video camera was placed in front of the participant to record his/her body movements during play.
The amount of body movements in each condition was quantified by a measure (denoted Gypsy score thereafter) computed as the normalized sum of the total angular movement over the entire duration of the song. Concretely, a sum of angular differences between each consecutive frame was computed, summed up over all frames (60 frames per second), and normalized by the number of frames in a song to account for differences in song duration.

Table 1. Typical body movements observed in the clips corresponding to the clusters depicted in Figure 5 and the emotion labels used by the observers:

Cluster	Body gestures	Affective Labels
A	Raising arms up to mid air	Excited, joyful, happy
B	Shaking body in a rhythmic fashion (dancing)	Excited, content, aroused
C	Thumbs-up and arm bent	Happy, satisfied, joyful
D	Leaning back and shaking body	Amused, excited, happy, content, surprised, satisfied
E	Shaking head	Relaxed, content satisfied
F	Dropping arms	Disappointed, frustrated, calm
G	Shaking/shivering body while leaning back	Disappointed, frustrated
H	Very little movement	Bored, disappointed

Unstructured play is considered important for the social, physical and cognitive development of children. Traditional observational research examining play behaviour at playtime (recess) has been hampered by challenges in obtaining reliable data and in processing sufficient quantities of that data to permit credible inferences to be drawn. This successful case study demonstrates that sensor based movement data can be used to explore children's social behaviour during naturalistic play.

There is an emerging literature concerning the use of such technologies to study individual differences in childhood social behaviour. Veiga and colleagues have used RFID to establish that frequent proximity to other children during play is a good predictor of social competence [9].

The basic idea of the study was to collect location data while the children were engaged in free-play on the school playground. We introduced a basic ‘loose parts play’ intervention to provide information about how changes to the play environment influence children’s interactions. Loose parts play introduces to the play space loose materials that have an ambiguous affordance [3]; the aim is to foster creativity and collaboration. In the baseline condition (10 sessions) children were left to play without any toys or other materials. In the intervention condition (10 sessions) loose materials, balls, and access to 2 large fixed play structures were provided and children were facilitated by two adult play workers.

Sensor data were collected over 27 sessions of at least 45 minutes each; of these, 7 initial sessions were assigned as pilot or habituation, 10 sessions as ‘baseline’ and 10 sessions as ‘intervention’. Due to the constraints of the school day it was not possible to randomise the hour assigned for data collection, instead we collected data at the same time each day (12:00-13:00). The playground (Figure 1) is 24×35 metres and was divided into zones defining areas of interest (shown in green in Figure 1), inaccessible zones (shown in red and dark red) and boundary zones (shown in blue and orange). During data collection, recordings were made using two cameras that covered the area, for ground truth data collection and validation.
Therefore, two bespoke sensors/devices were designed and developed in-house for this experiment: (i) a wearable head-mounted sensor with IMU and GNSS for precise positioning; and (ii) a wearable shoe-mounted IMU sensor for activity monitoring, motion capture and GNSS augmentation (see Figures 2 and 3).

Based on the video recordings and field observations, we observed and noted three main social play types that our method would need to characterise: (i) Hangout Play, (ii) Pursuit Play and (iii) Focal Play.

In this section, we propose and formulise different spatio-temporal metrics to measure the associations between children engaged in each type of play:
Warping Turning Distance
Warping Euclidean Distance
Warping Fréchet Distance

Creating a Movement-based Social Affiliation Network:
Considering the WFD method as likely to be the most informative irrespective of play type, the second piece of analysis involved creating a social affiliation network

summarising spatio-temporal data from each child across the 10 baseline and 10 intervention episodes, and using similarity between pairs of children to add links between nodes as described in 3.3.1.

The Overall Affiliation Matrices (visualised as clustergrams and networks in section 4), show there is interesting social information contained in the trajectories of children in a playspace.

From this exploratory case study, it is clear that using curve similarity to define links in an affiliation network is a potentially useful tool for exploring social structures and dynamics as they occur during naturalistic play. This could help to test hypotheses in developmental social psychology that have previously been impractical to explore, especially those that apply to whole group behaviour. For example, the hypothesised influence of play type and activity levels on the emergence of sex segregated behaviour in pre-pubescent groups could be studied [14], or theories of distinct social-behavioural phenotypes associated with different neurodevelopmental conditions could be tested [15]. The method also has potential as an objective test of change in social behaviour following intervention.

- Unraveling Students' Interaction Around a Tangible Interface Using Multimodal Learning Analytics
Bertrand Schneider, Paulo Blikstein

Students' gestures have received a great deal of attention from learning scientists over the last decades. Body movements and postures can provide valuable clues about students' prior knowledge, misconceptions, and problem-solving strategies when learning new concepts.
Numerous studies have unraveled links between students' understanding of various topics (e.g., Church, 1999; Abrahamson, Trninic, Guti;rrez, Huth & Lee, 2011) and specific gestures (e.g., deictic, iconic, and metaphoric gestures; Roth, 2001). More generally, there has been a plethora of studies about people's intuitive representations of everyday situations and bodily language (e.g., embodied cognition; Anderson, 2003). This line of research has provided new ways to understand how students integrate new concepts into their everyday understanding of science phenomena using gestures and, more generally, body actions.

Our goal is to apply data mining techniques on those two datasets to: 1) find patterns

that characterize productive groups; 2) use various data mining algorithms and explore their potential to extract predictors for learning; and 3) investigate the social component of our dataset in terms of body synchronization and proxemics. Thus our approach is mostly data-driven, in the sense that we don't have exact predictions of the type of patterns we will find.
However, our analyses are grounded by several psychological and educational theories (such as proxemics: Hall, 1966; student status in small collaborative groups: Shaer, Strait, Valdes, Feng, Lintz & Wang, 2011; bimanual coordination: Worsley & Blikstein, 2013; and more generally embodied cognition: Anderson, 2003), which makes our work partially theory-driven.

Gestures: Grafsgaard, Wiggins, and Boyer (2014) found that facial expression and gestures (e.g., hand-to-face gestures) were predictive of engagement and frustration, while facial expressions and body posture (e.g., distance from the computer screen) were predictive of learning.
Interaction: There is a large body of work on behavioral convergence looking at **multimodal indicators of behavioral synchronization between group member**s (e.g., Pentland & Heibeck, 2008).

In general, we believe that merely replacing human coders with computational tools does not take full advantage of the potential offered by sensors and algorithms. Our approach is to consider computational techniques as an augmentation of traditional research methods: algorithms can replace some qualitative analyses, but most of them are just too complex to be replicated automatically. Thus an additional goal is to use computers to provide an alternative perspective on educational datasets, which can then be used as lenses for constructing new hypotheses and analyses that couldn't be generated with traditional methods.

In this section, we describe various measures that we extracted from the Kinect logs. More specifically, we looked at: 1) the amount of movements generated by the students and the Kinect's use as a proxy for engagement; 2) prototypical body postures and students' likelihood to transition between them; 3) student's bimanual coordination and its relationship to group dynamics; 4) dyads synchronization and proxemics in small group collaboration. We predict that several of those measures should be related to learning, at least through an indirect effect via students' engagement, quality of collaboration, and cognitive states.

Our approach was to take our entire dataset (1 million entries; i.e., one entry is a line recorded by the Kinect sensor) and transform it into (joint) angles instead of positions in a three-dimensional Cartesian coordinate system. We then fed this matrix into a simple and fast clustering algorithm (K-means) that provided us with prototypical body positions. As a first step, we generated 2 to 9 clusters and visually inspected the results; we decided to keep three clusters, because the postures found were all perceptibly different, relatively easy to interpret, and there was no overlap between them.

The way we interpret those three clusters seems to correlate with statistical measures: the "active" posture is positively associated with students' learning gains $r(34) = 0.329$, $p < 0.05$ while the "passive" one is negatively correlated with students learning gains $r(34) = -0.420$, $p< 0.05$. Additionally, we found that the number of times students transitioned from one posture to another was also significantly correlated with their learning gains: $r(34) = 0.335$, $p < 0.05$.

Our sixth hypothesis is that students' leadership can be detected from students' gestures and are associated with the dyad's learning gains. In a related study, Worsley & Blikstein (2013) have shown that bimanual coordination was predictive of participants' expertise in solving an engineering problem. Based on these results, we decided to compute a similar metric for our dataset. More specifically, the idea is to compute and compare the amount of movement generated by each hand. Figure 8 shows all the graphs generated by this approach: some students barely use their left arm while others use both arms during the entire activity. It is interesting to see the variety of the graphs produced; all the students have a very distinct signature in terms of their hand movements.

To make sense of this metric, we need to introduce additional results that we found in the initial study (Schneider & Blikstein, 2015). Previous research has shown that each student working in groups can often be categorized as either being the "driver" or the "passenger" of the interaction (Shaer, Strait, Valdes, Feng, Lintz & Wang, 2011). One coder used several indicators to categorize each dyad's members: 1) who started the discussion when the experimenter leaves, 2) who spoke most, 3) who managed turn-taking (e.g., by asking "what do you think?"; "how do you understand this part of the diagram?"), and 4) who decides the next focus of attention (e.g., "so to summarize, our answers are..."; "I think we need to spend more time on this...").

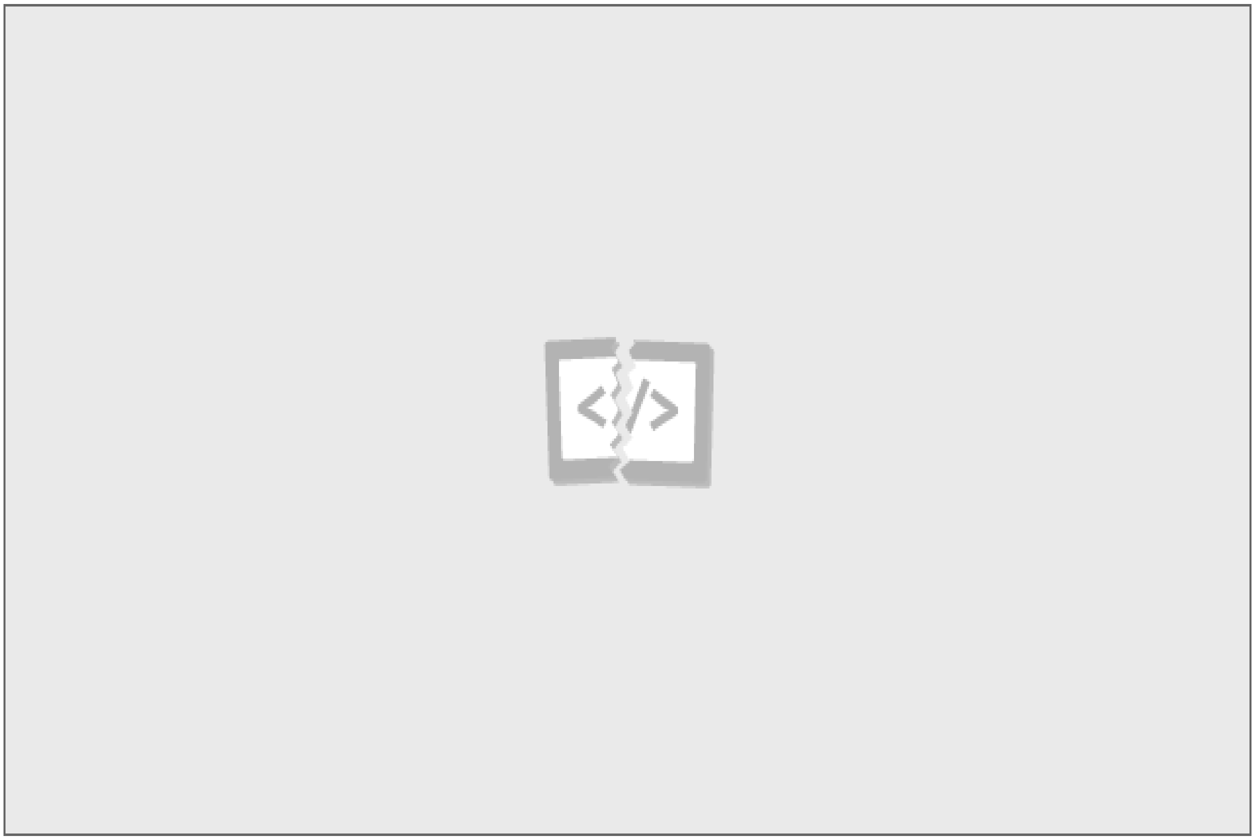
Second, we explored how Kinect data can inform the way we understand "in-situ" interactions around a tabletop: we found that the raw amount of movement was not a relevant predictor for our purposes; however, we found that bimanual coordination was predictive of students' leadership in a group. Even more interestingly, clustering body position with k-means provided us with interesting categories: we found that "active" positions were correlated with learning gains, "passive" positions were negatively correlated with learning gains, and that the number of transitions between those states was predictive of learning. Third, we explored students' body language on a social level: contrary to common social psychology theories, we found that body synchronization was not correlated with any of our measures. Fourth, we found that the distance between students' bodies during the activity was associated with their pre-existing knowledge on the topic taught: students with low scores on the pre-test tended to be further away from their partner compared to students who obtained a high score. We interpret this result as a sign of defensiveness regarding an unfamiliar and possibly difficult topic for them to learn about.

- Automated and unobtrusive measurement of physical activity in an interactive playground

Alejandro Morenoa, Ronald Poppe, Jenny L. Gibsond, Dirk Heylena

Promoting physical activity is one of the main goals of interactive playgrounds. To validate whether this goal is met, we need to measure the amount of physical player activity. Traditional methods of measuring activity, such as observations or annotations of game sessions, require time and personnel. Others, such as heart rate monitors and accelerometers, need to be worn by the player. In this paper, we investigate whether physical activity can be measured unobtrusively by tracking players using depth cameras and applying computer vision algorithms.

In general, these approaches are designed to put body movement as a core part of the gameplay in order to motivate players to exert themselves (e.g., Landry and Parés, 2014; Müller et al., 2012a). This does not necessarily mean that players engage in appropriate levels of exertion (Peng et al., 2013). Players could move very little, or players might move too much and burn out quickly. Knowing beforehand how to stimulate players appropriately is difficult, and is likely to differ between individuals. One promising alternative to control the level of exertion is to adapt the stimulation of the players in real-time, based on measurements of the players' levels of physical activity (Altimira et al., 2017). While these methods are suitable for the study of play in a laboratory setting, the requirement of fitting sensors and registering them to the game session hinders their use in everyday play. In the current paper, we present an approach that overcomes this limitation by measuring exertion in real-time and unobtrusively, using overhead cameras.



https://www.youtube.com/watch?v=7_mcWCB76Ps

- Studies referring to related work:

Veiga and colleagues have used RFID to establish that frequent proximity to other children during play is a good predictor of social competence [9].

Moreno and colleagues have demonstrated using simulation, how social roles during games can be inferred from movement data alone [7].

Argyle and Dean, who proposed the equilibrium theory [1]. This theory stated that people dynamically adapt their postures, gestures, gaze and proximity to others depending on how intimate their relationship is.

Hall proposed that we regulate our distance to others when interacting, and that these distances are dependent on our social intimacy with them [12]

Groh et al. developed a system that classified when people were engaging in social interactions using interpersonal distances and body orientation [11].

Bazzani et al. also used body orientation, represented by the subjective view frustum, along with spatial cues to infer when people were interacting [4].

Kong et al. proposed the use of semantic motion relationships between two interacting people to recognize human interactions [15].

Wang et al. presented an algorithm that exploited repetitions in social games to recognize stages during play and pairwise social interactions even in unstructured collections of videos [21].

Chang et al., whom proposed a probabilistic group level motion analysis for recognizing group behavior in unconstrained surveillance environments [5].

Tetteroo et al. designed an interactive playground that tracks children's movement and recognized some actions using cameras and accelerometers [20].

- "You're it!": Role Identification using Pairwise Interactions in Tag Games

Alejandro Moreno and Ronald Poppe
All these theories state that an individual's behavior is influenced by other people present in their surroundings, and that modeling these relationships helps to explain the exhibited behavior.
In surveillance settings, the use of proxemics to define groups is common, however our setting is that of games and playgrounds, where proxemic conventions do not hold and the identification of teams is more meaningful.

- Social interactions by visual focus of attention in a three-dimensional environment

L. Bazzani, M. Cristani, D. Tosato, M. Farenzena, G. Paggetti, G. Menegaz, V. Murino

This is especially regrettable when other domains, for example affective computing (AC) (Picard, 2000) or social signal processing (SSP) (Vinciarelli et al., 2009), pay significant attention to social, affective and emotional aspects of human behaviour.

In particular, SSP aims at developing theories and algorithms that codify how human beings behave while involved in social interactions, putting together perspectives from sociology, psychology and computer science (Pentland, 2007; Pantic et al., 2009; Vinciarelli et al., 2009). Here, the main tools for the analysis are the social signals (Vinciarelli et al., 2009), that is temporal co-occurrences of social cues (Ambady & Rosenthal, 1992) that can be basically defined as a set of temporally sequenced changes in neuromuscular, neurocognitive and neurophysiological activity. Social cues are organised into five categories that are heterogeneous, multimodal aspects of a social interplay (Vinciarelli et al., 2009): (1) physical appearance, (2) gesture and posture, (3) face and eyes behaviour, (4) vocal behaviour and (5) space and environment.

- Tetteroo, Daniel & Reidsma, Dennis & Dijk, Betsy & Nijholt, Anton. (2011). Design of an Interactive Playground Based on Traditional Children's Play. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering. 78. 10.1007/978-3-642-30214-5_15.

Almost without exception, they were prepared to give up computer gaming time in exchange for spending time with the interactive playground. Considering the motivation of this project, this promises a hopeful future for interactive playgrounds.