Multimodal AI Agent to Support Students' Motion-Based Educational Game Play

Serena Lee-Cultura[1[0000−0001−6277−1677]], Kshitij Sharma[1[0000−0003−3364−637X]], and Michail Giannakos[1[0000−0002−8016−6208]]

Norwegian University of Science and Technology, Trondheim, Norway
https://lci.idi.ntnu.no/ {serena.leecultura,kshitij.sharma,michailg}@ntnu.no

Abstract. Increased accessibility of lightweight sensors (e.g., eye trackers, physiological wristbands, and motion sensors), enable the extraction of student's cognitive, physiological, skeletal, and affective data, as they engage with Motion-Based Educational Games (MBEG). Real-time analysis of this Multi-Modal Data (MMD) leads to a deep understanding of student's learning experiences and affords new opportunities for timely, contextual, personalised feedback delivery to support the student. In this work-in-progress, we present the MMD-AI Agent for Learning; a MMD-driven Artificially Intelligent (AI) agent based eco-system, composed of 3 separate software components, which work together to facilitate student's learning during their interactions with MBEG. The Crunch Wizard, receives MMD from eye-trackers, physiological wristbands, web camera, and motion sensors worn by a student during game play, and derives relevant cognitive, physiological and affective measurements. The AI agent identifies and delivers appropriate feedback mechanisms to support a student's MBEG play learning experience. The Dashboard visualises the measurements to keep teachers informed of a student's progress. We discuss the foundational work that motivated the ecosystem's design, inform on our design and development accomplished thus far, and outline future directions.

Keywords: multimodal · ed-tech · artificial intelligence · motion-based games.

1 Introduction and Motivation

It is important to obtain a holistic understanding of student's educational experiences in order to provide them with appropriate, constructive, timely and personalised support during their interactions with learning activities. Current practices typically rely on researcher's observation, surveys, interviews and questionnaires. Though these methods have undoubtedly advanced our understanding of student's educational experiences, their affordances are limited to the student's memories, and perceptions of what researchers can "externally" observe. Additionally, data collected in this manner can only be analysed after the learning experience has completed; thereby preventing opportunities to scaffold student's learning in real-time through provision of feedback mechanisms.

In this regard, recent advancements in sensing technologies and their respective Multi-Modal Data (MMD), demonstrate the capacity to transform how we understand and conduct student's educational experiences. Specifically, Motion-Based Educational Games (MBEG) utilise sensing devices (e.g., Microsoft Kinect) to capture, map and interpret student's full-bodied movements as game input [1]. These games have gained traction across a multitude of educational domains in student's learning (e.g.,[10, 8]). Previous works suggests that compared to traditional classroom instruction (i.e., pen and paper), the "touchless" nature of MBEG offers student a more engaging and natural learning experience [9].

In addition to motion sensing devices, wearables like eye-tracking glasses and wristbands, allow for the unobtrusive, continuous, and automatic collection of student's MMD generated during their learning experiences. Leveraging the affordances of MMD (e.g., temporality and direct access to indicators of student's cognitive and affective processes [3]), provides researchers with a new vantage from which to observe and monitor student's "internal" states (e.g., cognitive, physiological) in real time [19]. Thus, investigating the inherent richness of student's natural interactions with MBEG by using MMD might enhance our understanding of student's learning experiences [13, 14], and contribute to the development of supportive measures (i.e., feedback mechanisms) which can be integrated to student's educational technologies/experiences. However, despite the potential of interlacing MBEG and MMD from sensing technologies, limited research has explored the confluence of these ideas/domains to better support/scaffold student's learning experiences through automatic, personalised, feedback generation/delivery.

To bridge this gap, we are creating a MMD-driven Artificially Intelligent (AI) agent and visualisation dashboard ecosystem that: 1) collects and assesses student's MMD during their interactions with MBEG, 2) identifies, prioritises and delivers appropriate feedback mechanisms to scaffold student's educational game-play experiences in real-time, and 3) visualises the MMD for teachers to keep informed on individual student's progress so that teachers can contextualise this data with their own personal accounts of each student. In this paper, we present our product vision, work completed thus far, and discuss directions of future work.

2 The MMD AI Agent Educational Ecosystem

The purpose of this work is to create an MMD-driven AI educational scaffolding ecosystem, that automatically and continually, provides personalised feedback to support student's learning during their interactions with MBEG. The MMD-AI Agent for Learning consists of three components, the Crunch Wizard, a Dashboard, and an AI agent (each described in section 2.2). Research, development and validation of the MMD-AI Agent for Learning occurs in three consecutive phase, where each phase builds on the previous results. In phase 1, we augmented student's MBEG play with MMD collection devices to explore how the

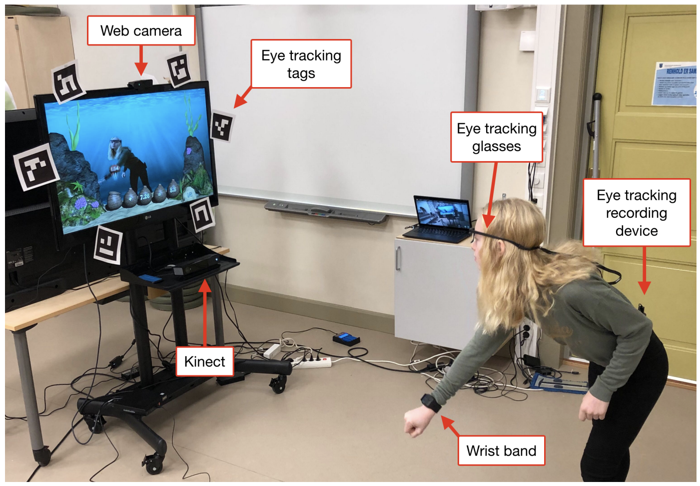


Fig. 1: Experimental setup of a student playing a MBEG. Labels indicate the MMD capture devices.

collective data (i.e., gaze from eye-tracking glasses, physiological from wristband, and skeletal from Kinect) might be used to understand and support student's learning. In phase 2, we commenced the design and development of the MMD-AI Agent for Learning, based on design implications that resulted from phase 1. We are currently partway through phase 2; initial versions of the Crunch Wizard and the Dashboard have been implemented, however, development of the AI agent has not started. Lastly, in phase 3, we will conduct an empirical study to assess the efficacy of the MMD-AI Agent for Learning, from the perspective of both students and teachers. Each phase is detailed below.

2.1 Phase 1: Discovering the possibilities of MMD and MBEG

The purpose of this stage was to explore how MMD contributes to understanding students' interactions with MBEG, and to identify ways to support students' game play experience through real-time, personalised, MMD-driven feedback mechanisms. To address this, we conducted a study in the winter of 2019, in which students wore Tobii eye-tracking glasses and Empatica E4 wristbands, as they played three commercial single-player MBEG from the Kinems Platform[1] focused development of maths and language skills. The MBEG tracked student's movement using Kinect.

Context & Participants: the study took place in a local science museum and public elementary school in a European city. The designated study space was arranged to accommodate two experimental set-ups running in parallel. Our sample included 46 students (28F, 18M) with an average age of 10.3 years (SD = 1.32, min= 8, max= 12 years). Sixteen students participated at the science

[1] https://www.kinems.com/

Table 1: Devices and the associated MMD measurements from the initial study.

Eye-Tracker	Cognitive load [4], Focus [15], Perceived difficulty [17, 18], Anticipation [4], On-task ratio
Wristband	Arousal [6], Stress [7], Hand movement, Engagement [6], Emotional regulation [2], Entertainment [20]
Kinect	Fatigue
Webcam	Emotion [5]

centre, and 30 at the elementary school. No students had prior exposure to MBEG. Students received a gift card in exchange for their participation.

Procedure & Data Collection: Each student was given a pair of Tobii eye-tracking glasses , and an Empatica E4 wristband to wear. They played 3 different MBEG, 3 times each (9 sessions total). Student participated in a practice session where experimenters assisted them in understanding the game's rules and objective. The cumulative play time of all 9 sessions ranged between 25-30 minutes. We collected sensor data from: Tobii eye-tracking glasses , Empatica E4 wristband (with sensors for HRV, blood-pressure, temperature and EDA levels), Microsoft Kinect (which tracked skeletal movement of 25 different joints), and Logitech web camera. We also screen recorded student's game sessions, examined their system logs (e.g., for specific event occurrences) and monitored in-game performance metrics (i.e., reaction times, game score). Figure 1 shows a student in action, wearing the data collection devices, together with experimental set up. Table 1 presents the measurements that were derived from the raw MMD.

Results: We analysed the data according to several research questions, but limit our presentation of results to only those which motivated, and are included in the design of, the AI agent discussed in Phase 2. Specifically, our results suggested the following possibilities (references concealed to preserve anonymity).

Problem Solving and Play: we examined the possibility of using MMD measurements derived from student's interactions with MBEG, to determine when students are 1) playing rather than problem solving, and 2) guessing rather than taking an informed decision [11, 12].

Prediction: the combination of gaze and physiological measures (from eye-trackers and wristband, respectively), can be used to predict student's in-game performance metrics (where performance is measured by correctness of answers). Prediction occurs with extreme accuracy (i.e., 94%) in a short amount of time (i.e., 47s) [13].

Prioritisation: relative to the student's correctness of answer (history of right or wrong) during MBEG play, we developed a novel MMD measurement classification scheme and corresponding algorithm which may be used together to prioritise the delivery of different (and possibly conflicting) types of feedback when multiple avenues of support are applicable for addressing the needs of different MMD measurements. We have broadly identified different types of feedback (e.g., scaling of content difficulty, offering hint,

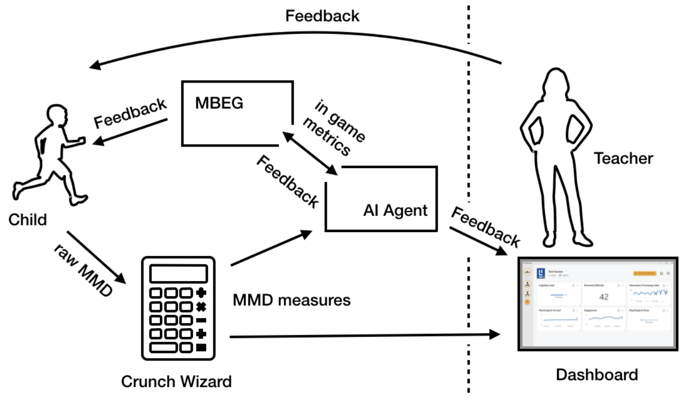


Fig. 2: MMD-AI Agent for Learning overview. The dotted line signifies that the student may participate remotely from the teacher.

informing the teacher to interact with the student in a specific manner, removal or inclusion of game distractors) corresponding to the different MMD measures. In practice, the exact realisation of each feedback type must be contextualised and thus, will largely depend on the focus of the MBEG [16].

2.2 Phase 2: Realising the MMD-driven AI ecosystem

Building on the above results, we designed and have started developing the MMD-AI Agent for Learning, purposed to provide students with personal, contextualised, real-time support delivered either in-game or from their teacher, as they interact with the MBEG. MMD-AI Agent for Learning consists of MMD devices (i.e., Tobii eye-tracker, Empatica E4 wristband, Logitech webcam, and Microsoft Kinect) and three software components; namely, the Crunch Wizard, Dashboard, and AI agent.

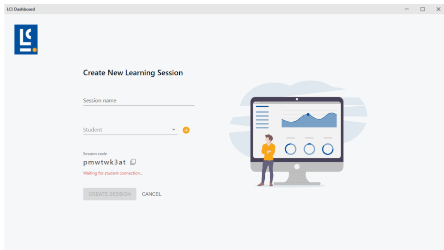
How it Works: the MMD-AI Agent for Learning works as follows. A student wears Tobii eye-tracking glasses and an Empatica E4 wristband as they interact with a MBEG. Throughout their game play, the student's raw MMD (gaze from eye-tracker and physiological from wristband) is continuously collected by the devices and sent to the Crunch Wizard program in real-time. Skeletal and video data are sent by the Kinect and Logitech webcam, respectively. The Crunch Wizard calculates measurements from the raw MMD, writes the measurements to a database, and sends them to the AI agent and the Dashboard. The AI agent assesses the student's measurements, in conjunction with in-game performance metrics sent from the MBEG, and determines whether the student requires support. If so, the AI agent determines the appropriate feedback mechanisms and mode of delivery. Feedback in the form of suggestions for the teacher, may be sent to Dashboard. Feedback may also be realised as in-game student support. The Dashboard visualises the incoming measurements according to the teacher's preferences (e.g., specifically selected students and measurements), as well as a notification centre where the teacher may receive feedback alerts indicating appropriate ways to scaffold the students experience.

The system *can* be used by a single student, or multiple students concurrently. In the case of multiple student's, each must participate from separate learning environments, such that they have access to their own eye-tracker, wristband, web camera, and MBEG. Each separate setup includes its own Crunch Wizard and AI agent program instances. However, all instances of Crunch Wizard and AI agent, pass their data to a single shared Dashboard. A student's Crunch Wizard and AI agent must be running on the same machine, but due to the decoupled nature of the software components, the Dashboard is able (but not required) to reside on a different machine, at a separate physical location. Thus, the MMD-AI Agent for Learning enables remote educational instruction, where each student interacting with a MBEG, and a teacher monitoring their learning experience (via MMD measurement viewed on the Dashboard), can participate from their separate locations. Figure 2 shows an overview of the MMD-AI Agent for Learning, together with the MBEG, and directional flow of data (raw MMD, derived measurements, in-game metrics) and feedback mechanisms. Following, we provide a detailed account of each software component: Crunch Wizard, Dashboard, and AI agent.

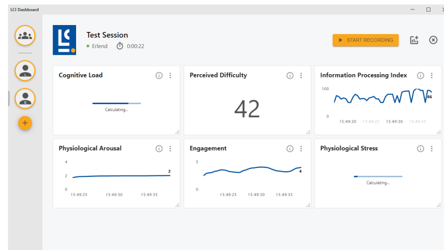
Crunch Wizard receives an influx of raw MMD (i.e., gaze, physiological, and skeletal) sent from the devices (Tobii eye tracker, Empatica E4 wristband, Logitech web camera, and Microsoft Kinect) worn by a student during game play throughout their learning experience, and calculates various corresponding MMD measurements (see 1). Since the devices monitor a student's state (e.g., HR, eye and joint movement), new data is continuously sent to Crunch Wizard, and corresponding new measurement updates are constantly calculated and propagated to the dashboard and AI agent. The current iteration of Crunch Wizard calculates the following measurements from incoming raw MMD: physiological stress, physiological arousal, engagement, emotional regulation, perceived difficulty, anticipation, cognitive load, information processing index, emotions, fatigue, motion stability.

Dashboard visualises students' MMD measurements in real-time, to help teachers know what necessary action they should take to improve the learning process of individual students. A future iteration will display recommendations to guide the teacher in providing appropriate feedback to specific students to support their learning session in real-time. The feedback recommendations come from the AI agent and are discussed in the following text.

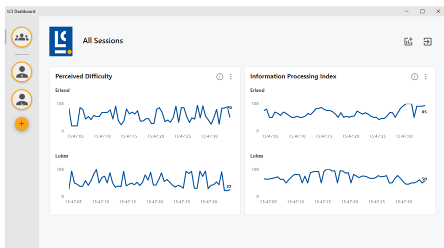
When the teacher enters the dashboard, they are prompted to fill out a form to create new *learning session* (see Figure 3a). The teacher enters a session name, and selects a student to participate. To add a student, they must have a learner profile. If no profile exists, the teacher must create one (by clicking the yellow + button). Each learning session has a unique *session code*, that is used to connect the student's software components (Crunch Wizard, and AI agent) to the teachers dashboard. After a connected has been established, the teacher is able to select between an "all sessions view" and individual session



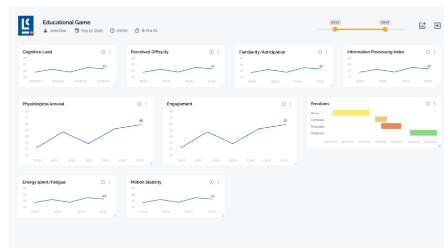
(a) View of new learning session and student profile creation.



(b) View of a single student's learning session.



(c) View of multiple student's learning sessions for measurement comparison.



(d) View of a student's recorded learning session.

Fig. 3: Different views offered by the dashboard.

views for each student. The all session view shows data for every participating student, whereas the individual student views allow the selection of any number of participating students. Figure 3b shows a live learning session view for a single student, Erland.

The teacher can customise which measurements (from those listed in the Crunch Wizard section) they wish to view during the learning session. This allows the teacher to assess the student's learning experience based on specifically selected metrics. If viewing multiple students, the teacher can identify, and direct their attention, the students that require additional support with the knowledge of where they are struggling. For example, a teacher may wish to examine the perceived difficulty and information processing index of two participating students, Erland and Lukas (see Figure 3c). Perceived difficulty is the level of difficulty the student attributes to the question. Thus, in seeing that Erland is experiencing the problem as harder than Lukas, the teacher can pay special attention to Erland's progress and support him if he appears to be struggling.

The Crunch Wizard is responsible for sending student's MMD measurements to the dashboard, and recording this the data in a database. The dashboard supports viewing learning sessions after a learning session has finished. The teacher can set the time interval they wish to view, which controls how much of the learning sessions data is shown on each measurement's selected graph. Figure 3d shows 34s of data for a single student's learning session (as set by the yellow slider on the top right). Viewing and comparison of multiple learning sessions belonging to the same student, or multiple students, is also possible.

AI agent orchestrates feedback delivery to scaffold the student's learning experience. It receives a continuous flow of real-time data from both the MBEG (e.g., the student's in-task performance metrics) and Crunch Wizard (e.g., MMD measures), collectively assesses the student's measurements, determines the appropriate feedback mechanisms required to facilitate the learning experience (if needed), and delivers the feedback accordingly. The type of feedback to deliver is determined by a combination of a student's performance prediction results, the MMD measurements classification and prioritisation algorithm, and play and problem solving design implications, derived from our previous publication, as mentioned above. The feedback is realised in numerous ways (i.e., in-game hints, scaling of content difficulty, informing the teacher that help is needed), according to the student's individual MMD measurements and in-game performance metrics. Moreover, feedback is provided either directly to the student through modification/supplementation of game content (e.g., scaling down content difficulty in response to elevated cognitive load or perceived difficulty), or to the teacher via the visualisation Dashboard (e.g., informing the teacher to encourage the student) so that they can use their personal knowledge of the student to determine a contextualised appropriate course of action. Thus, the AI agent must maintain a bi-directional stream of communication with both the MBEG and the Dashboard, but only receives data from the Crunch Wizard (see Figure 2). In order to enable the provision of feedback directly to the child through modification on the game content, the AI agent must communicate with the MBEG engine. In our first iteration of the AI agent, we will adapt, and connect directly to, the game engine for the Kinems Platform utilised in our study presented in Phase 1. Future versions will focus on developing the AI agent into a software SDK for easy integration into MBEG during their development phase. Development of the AI agent has not yet commenced.

2.3 Phase 3: User Testing

Thus far, the current version of the MMD-AI Agent for Learning (i.e., Crunch Wizard and Dashboard, without AI agent), has undergone a single user test with an adult acting as a teacher (though not a teacher). The motivation behind the usability testing was to identify issues with the Dashboard layout and functionality. Therefore, there was no need to include students playing a MBEG. Rather, the test was consulted remotely, with a researcher wearing the devices while web camera was capturing their movement. The test subject navigated to various views within the Dashboard during this experience, and was asked to identify and interpret different information on-screen. Test data was collected via video/audio recording and System Usability Scale (SUS) survey. Results indicated that there there *was* confusion related to recording, and viewing of recorded, data (as discussed above in the Dashboard section and shown in Figure 3d). Upon completion of development and integration of all three software applications (Crunch Wizard, Dashboard, and AI agent), the MMD-AI Agent for Learning will be extensively tested by teachers to assess its usability, and likelihood of class-room adoption. We require relevant test tasks (i.e., the system will be tested using the MBEG from the Kinems Platform), subjects (i.e., both teachers and students), and environment (i.e., in class-room).

3 Future Work and Conclusion

In future, we aim to extend the MMD-AI Agent for Learning to accept data from a stationary eye-tracker. Upon creation of a learning session, either the teacher would input or the system would detect, which eye-tracker a given student will use during their learning session. This will allows a student to sit directly in front of a computer screen which supports traditional computer-based learning experiences (e.g., non-motion based). A driving factor for augmenting the eye-tracking capabilities of the system, is to promote inclusion of students with limited mobility. Additionally, this would also enable future studies which compare traditional computer-based learning games with their motion-based counter parts. The Kinems Platform is a favourable candidate for this endeavour, as it includes learning games that are offered in both motion and non-motion -based form. Our work contributes to the understanding of multi-sensor devices “beyond mouse and keyboard” in learning contexts with the purpose of automatic feedback generation, adaptation, and personalisation in student’s learning, and good graphical visualisation of quantitative information, to make the learning processes more measurable and interpretable.

References

1. Bartoli, L., Corradi, C., Garzotto, F., Valoriani, M.: Exploring motion-based touchless games for autistic children’s learning. In: Proceedings of the 12th international conference on interaction design and children. pp. 102–111. ACM (2013)
2. Berntson, G.G., Cacioppo, J.T.: Heart rate variability: Stress and psychiatric conditions. Dynamic electrocardiography **41**(2), 57–64 (2004)
3. Cukurova, M., Giannakos, M., Martinez-Maldonado, R.: The promise and challenges of multimodal learning analytics (2020)
4. Giannakos, M.N., Papavlasopoulou, S., Sharma, K.: Monitoring children’s learning through wearable eye-tracking: The case of a making-based coding activity. IEEE Pervasive Computing pp. 1–12 (2020). https://doi.org/10.1109/MPRV.2019.2941929
5. HAGER, P.E.W.F.J.: Facial action coding system. the manual on cd rom (2002)
6. Hasson, U., Furman, O., Clark, D., Dudai, Y., Davachi, L.: Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. Neuron **57**(3), 452–462 (2008)
7. Herborn, K.A., Graves, J.L., Jerem, P., Evans, N.P., Nager, R., McCafferty, D.J., McKeegan, D.E.: Skin temperature reveals the intensity of acute stress. Physiology & behavior **152**, 225–230 (2015)
8. Homer, B.D., Kinzer, C.K., Plass, J.L., Letourneau, S.M., Hoffman, D., Bromley, M., Hayward, E.O., Turkay, S., Kornak, Y.: Moved to learn: The effects of interactivity in a kinect-based literacy game for beginning readers. Computers & Education **74**, 37–49 (2014)

9. Hsu, H.m.J.: The potential of kinect in education. International Journal of Information and Education Technology **1**(5), 365 (2011)
10. Lee, E., Liu, X., Zhang, X.: Xdigit: An arithmetic kinect game to enhance math learning experiences. Retrieved February **14**, 2013 (2012)
11. Lee-Cultura, S., Sharma, K., Cosentino, G., Papavlasopoulou, S., Giannakos, M.: Children's play and problem solving in motion-based educational games: Synergies between human annotations and multi-modal data. In: Proceedings of the Interaction Design and Children Conference. ACM (2021 (in press))
12. Lee-Cultura, S., Sharma, K., Giannakos, M.: Exploring children's play and problem solving in motion-based learning environments. In: International Journal of Child Computer Interaction (2021 (forthcoming))
13. Lee-Cultura, S., Sharma, K., Papavlasopoulou, S., Giannakos, M.: Motion-based educational games: Using multi-modal data to predict player's performance. In: 2020 IEEE Conference on Games (CoG). pp. 17–24. IEEE (2020)
14. Lee-Cultura, S., Sharma, K., Papavlasopoulou, S., Retalis, S., Giannakos, M.: Using sensing technologies to explain children's self-representation in motion-based educational games. In: Proceedings of the Interaction Design and Children Conference. pp. 541–555 (2020)
15. Pappas, I., Sharma, K., Mikalef, P., Giannakos, M.: Visual aesthetics of e-commerce websites: An eye-tracking approach (2018)
16. Sharma, K., Lee-Cultura, S., Giannakos, M.: Keep calm and don't carry-forward: Towards sensor-data driven ai agent to enhance human learning (2021 (under review))
17. Sharma, K., Papamitsiou, Z., Giannakos, M.: Building pipelines for educational data using ai and multimodal analytics: A "grey-box" approach. British Journal of Educational Technology (2019)
18. Sharma, K., Papamitsiou, Z., Giannakos, M.N.: Modelling learners' behaviour: A novel approach using garch with multimodal data. In: European Conference on Technology Enhanced Learning. pp. 450–465. Springer (2019)
19. Thorson, K.R., West, T.V., Mendes, W.B.: Measuring physiological influence in dyads: A guide to designing, implementing, and analyzing dyadic physiological studies. Psychological methods **23**(4), 595 (2018)
20. Yannakakis, G.N., Hallam, J., Lund, H.H.: Entertainment capture through heart rate activity in physical interactive playgrounds. User Modeling and User-Adapted Interaction **18**(1), 207–243 (2008)