USE OF ELECTRODERMAL WRISTBANDS TO MEASURE STUDENTS' COGNITIVE ENGAGEMENT IN THE CLASSROOM

Patrick Terriault [1], Anastassis Kozanitis [2], and Patrice Farand [3]
[1] École de technologie supérieure, [2] Université du Québec à Montréal, [3] Polytechnique Montréal
Corresponding author: Patrick.terriault@etsmtl.ca

Abstract *– A pilot project was conducted to study the feasibility of using electrodermal activity sensors embedded in a watch-like device to measure skin conductivity in real time. In the field of education, it may be interesting to use this technology to assess the students' cognitive engagement in the classroom. A few volunteer students as well as the professor were wearing an Empatica E4 wristband during some class periods where different activities were organized such as lectures, workshops and exams. Monitoring several individuals simultaneously makes possible to compare the collected data among students and between the students and the professor. Also, since the activities were weekly repeated, it was possible to assess to which extent the observed patterns were similar from one group to the other. In brief, the collected data is very difficult to interpret, since some external factors seem to have a significant effect on the measurements. Indeed, discrepancies are observed in the data curves representing the students' electrodermal activity. Also, the data generated by the professor is quite different from one group to the other, even if he repeated the exact same activities at two different times of the week. It is suggested to improve the understanding of all the phenomena that could affect the electrodermal activity measurements before trying to draw conclusions related to the students' cognitive engagement in the classroom.*

Keywords : Cognitive engagement, Electrodermal activity, Empatica E4 wristband, Engineering education.

1. INTRODUCTION

When the level of engagement of an individual in a physical activity has to be measured, physiological data such as the number of heartbeats per minute can be used. Unfortunately, the situation is not as simple when it comes to measure the level of cognitive engagement of an individual during a learning activity. Azevedo [1] as well as Dirican and Göktürk [4] identify several psychophysiological variables that can be used to measure cognitive engagement such as pupil diameter, blood pressure, cardiovascular measures and electrodermal activity.

Electrodermal activity (EDA) occurs when a person feels emotions, makes a physical effort or is engaged in a cognitive load. In these situations, the brain sends signals to the skin to initiate the sweating process, which modifies the electrical properties of the skin, even if the amount of sweat is barely noticeable [3]. It is an unconscious phenomenon over which the person has no control. Therefore, there is a real interest in using EDA in the field of education to measure the level of cognitive engagement in different situations: types of pedagogical strategies proposed to students (lectures, project-based approach, flip classrooms, etc.), types of activities (experimental manipulations, computer simulations, team problem solving, etc.), time of the day when students are involved in learning activities, etc. Knowing the level of cognitive engagement of students in different circumstances could allow the optimization of engineering education programs.

Laboratory equipment can be used to measure the EDA of a person, but the recent development of portable devices makes it possible to conduct experiments while the subject goes about his or her daily activities [6]. One of these devices is the E4 bracelet developed by the Italian firm Empatica in which the EDA sensor is embedded in a watch-like instrument worn on the wrist.

The use of such portable devices has been documented in the literature and some published results are of interest. For example, in the work of Poh et al. [11], a student's EDA was measured for one week without interruption. The results showed that the signal was sometimes higher while sleeping than during a lecture in class. Also, Potter et al. [12] have used the Empatica E4 bracelet in an academic context. These authors mentioned that it would be desirable to instrument several students simultaneously in order to make a judgment on the repeatability and generalization of the measurements. Unfortunately, there is no consensus on the potential of the electrodermal activity measurement to estimate the level of cognitive engagement of students during a learning activity. Indeed, Larmuseau et al. [7] as well as McNeal et al. [8] indicate that their results are correlated with students' self-reported mental effort during an activity, while others such as Menghini et al. [9] identify important limitations to use EDA measured at the wrist to estimate the level of cognitive engagement of students.

A research project was launched to contribute to the scientific literature concerning the repeatability of data collected from several students who follow the same learning activities at different times of the week. Prior to a large-scale study, a pilot project was set up to familiarize the research team with Empatica E4 sensors used in an educational research context. The results of this pilot project will be used to set the methodological aspects of the large-scale project in order to maximize the fidelity and reliability of the collected data. In this publication, the context in which the data was collected and the methodology are first described. The data itself is then presented. Finally, before the conclusion, a discussion related to the EDA measurements is carried out.

2. CONTEXT

The pilot project took place during the Fall 2019 semester, which includes 13 weeks of activities, excluding the final exam week. The professor taught the MEC423 course to two different groups of approximately 40 students each. This is a mandatory course in the mechanical engineering program at École de technologie supérieure and is usually taken at the end of the second year of a four-year curriculum. In this course, students learn the theoretical foundations of the finite element method, program the solution of simple problems in MATLAB, and use the commercial software ANSYS to simulate the mechanical and thermal behavior of real components. Each week, students are required to attend a 3-hour class period with the professor and another 3-hour practical work period with a teaching assistant. The class periods for Groups 1 and 2 took place respectively on Wednesday starting at 8:30 am and Tuesday starting at 1:30 pm. Students of these two groups were very homogeneous: 76 of the 79 students (96%) were male and 65 of them (82%) were between 22 and 26 years of age.

3. METHODOLOGY

The research project was presented to the students and participants were verbally invited to volunteer to wear a bracelet for five 3-hour class periods, specifically during weeks 4 to 8 of the semester inclusively. Only six Empatica E4 bracelets were bought with the available budget. Considering that the professor was going to wear a bracelet, five students from each group could participate in the study. Following the invitation, three students from group 1 (students #1, #2 and #3) and four students from group 2 (students #4, #5, #6 and #7) were recruited. Since each E4 bracelet is identified by a unique serial number, it was possible to ensure that the participants and the professor were systematically using the same bracelet. The participants were all men. The project obtained approval from the university Research Ethics Board (reference number H20190605) and all seven participants signed a consent form. The students agreed to wear the bracelet so that the level of electrodermal activity could be measured in real time during the class periods.

During the five weeks of data collection, four of these (Weeks 4, 5, 6 and 8) were typical class periods divided into two parts: first a lecture where the professor introduced the new material, and then a workshop where students were asked to gather a team and solve exercises while the professor moves from one team to the other to answer questions and to reexplain a misunderstood concept. For Week 7, a 2-hour exam was scheduled. By doing so, students' and professor's EDA can be observed and compared during three different learning activities: lecture, workshop and exam.

Participants were expected to arrive between 5 and 10 minutes before the start of the course. The professor himself proceeded to the installation of the E4 device on the participants' dominant arm. The bracelet is equipped with a series of notches and special care was taken to ensure that the installation was uniform for all participants. If it was possible to see the flashing green lights on the back of the bracelet by a simple rotation of the arm, then the bracelet was tightened one notch. Afterwards, the students had no further action to take during the entire course. In fact, the data was autonomously collected during the entire class period at a rate of 4 Hz.

At the end of a class period, participants gave back their E4 bracelet. Once returned to his office, the professor i) transferred the data from each bracelet to the Empatica website via a USB connection to his computer, ii) cleaned the bracelets with alcohol, and iii) recharged the internal battery for the next class period. Finally, the data was exported in a CSV format for processing with MATLAB.

4. RESULTS

The main results are presented in Fig. 1 to 9, all of which are grouped together in appendix A at the end of the paper. Figures 1, 3, and 5 correspond to the data for Group 1 collected during Weeks 4, 6, and 7, respectively, while Fig. 2, 4, and 6 correspond to the data for Group 2 for the same weeks. Note that the curves for weeks 5 and 8 are very similar to those for weeks 4 and 6 and are not provided in the appendix. Figures 5 and 6 present the data collected during the exam of Week 7. In these figures, the abscissa is the elapsed time, expressed in hours, since the beginning of the class period, while the ordinate is the skin conductivity expressed in μS. The symbols superimposed on the curves indicate the type of activity at the time the measurement is taken: the solid squares (■) represent lectures, the empty circles (○) correspond to the workshops and the solid triangles (▲) refer to the exam. For example, the Week 4 lecture lasted approximately 90 minutes for Group 1 (see Fig. 1) compared to approximately 75 minutes for Group 2 (see Fig. 2).

To evaluate to which extent the data is repeatable from one week to the next, the curves generated by an individual

are superimposed on the same graph. Thus, Figures 7, 8 and 9 show respectively all the curves obtained from the professor, student #3 of group 1 and student #7 of group 2 respectively. On these three figures, the red curve corresponds to the exam (Week 7), while the blue curves refer to the regular class periods comprising a lecture and a workshop (Weeks 4, 5, 6 and 8).

5. DISCUSSION

5.1. Professor's EDA measurements

Even though the professor gave the exact same learning activities to Groups 1 and 2 in Week 4, his EDA curves show a different pattern. In fact, for Group 1 (red curve in Fig. 1), the EDA increased from the beginning of the class and stabilized in the 30 minutes that followed. For Group 2 (red curve in Fig. 2), a period of approximately 45 minutes elapsed during the lecture before an EDA increase could be observed. After the lecture, a decrease in the EDA is observed in both groups during the workshop. This behavior was systematically observed during the following weeks, as shown by the red curves in Fig. 3 and 4. It should be noted that in Fig. 4, the teacher left the classroom to pick up a document from his office. The professor had to run through the university corridors and this physical activity, represented by the black portion of curve beginning at approximately 1.5 hours, resulted in an EDA increase. As expected, this shows that physical activity results in an EDA increase that is not related to cognitive engagement.

During the exam (Week 7), the professor only watches the students to avoid plagiarism. The teacher was therefore cognitively unengaged and this is shown by the red curve in Fig. 6 for Group 2. In Fig. 5 for Group 1, the pattern of the red curve is quite different, but this can be easily explained. Indeed, the professor had to find an unused classroom only a few minutes before the exam due to an error in the reservation system. This clearly shows that a stressful event also affects the EDA for a relatively long period of time (more than 30 minutes in this example).

When superimposing all the curves generated by the professor, Figure 7 shows consistent results. Indeed, the EDA is high during the lectures, moderate during the workshops and low during the exam. At first glance, this result is not surprising, since the mental load required to give a lecture is higher than for the other two types of activities. Moreover, the trends in the results is reproducible over the weeks, although the amplitude of the signal may vary from week to week. These variations in amplitude and the delay before the EDA increases during the lectures in Group 2 remain unexplained. As pointed out by Boucsein [2] and Doberenz et al. [5], medications including caffeine affect EDA measurements. The professor does not consume caffeine, but takes medication every morning. This raises the question of whether the effect of medications could cause the observed differences.

5.2. Students' EDA measurements

In general, when compared to the professor's EDA, students' EDA is rather low and stable during the typical class periods (Fig. 1, 2, 3 and 4), regardless of the type of learning activity (lecture or workshop). Indeed, the data collected showed no significant difference between learning activities for most students (e.g., Fig. 8), whereas we observed a stronger signal during exams for only a few students (e.g., Fig. 9). Based on our intuition, the level of EDA was expected to be relatively low in the lectures, higher in workshops, and even higher in the exam, which is obviously not the case. Could the room temperature be insufficient to observe significant variations in students' EDA when different learning activities occur? After checking with the technical service of the university, the temperature of the two classrooms was always controlled at about 21°C. With this regards, Pijeira-Díaz et al. [10] indicate that the room temperature should be constant at about 23°C for adequate EDA measurements. It would therefore be important to conduct some tests to assess whether the EDA of a student can be adequately measured at 21°C.

Some students show a high EDA level at the beginning of the lectures and a significant amount of time (approx. one hour) was required to bring back the signal to the baseline. This is the case for Students #1 and #3 at Week 4 (see Fig. 1). One must conclude that events prior to the beginning of the course influence the measurements. The 5-10 minutes period at the beginning of the class periods to stabilize the signal seems insufficient for some students, posing a logistical challenge considering that students generally have other obligations prior to a class. In this context, it would be ideal to have an accurate knowledge of students' lifestyle habits before class, particularly with regard to physical activity, caffeine and medication consumption. This information could be obtained during individual interviews scheduled immediately after a class period to try to identify the reasons for the significant variations in EDA signal.

The EDA of some students shows instantaneous peaks followed by gradual decreases until the baseline is reached back. This is particularly the case for Student #6 at Week 6 (see Fig. 4). To understand what causes these peaks, participants could be asked to note all the particular events that occur during the class period and interviewed just after the class period. Also, it could be advisable to record the class periods so that it would be possible to go back in time and accurately identify what was happening in the classroom at specific times.

It seems obvious that sensitivity is different from one student to another. Indeed, Student #2 in Group 1 (blue curves in Fig. 1, 3 and 5) generates systematically a weaker signal than the other students in his group, whereas the opposite situation is observed with Student #5 in Group 2 (blue curves in Fig. 2, 4 and 6). Also, by analyzing the

amplitude of the EDA signal, significant disparities are observed between students. Indeed, the skin conductivity of Student #3 in Group 1 (Fig. 8) varies between 0 and 12 µS, while that of Student #7 in Group 2 (Fig. 9) is almost always less than 1 µS. It is therefore important not to use data from a single student to draw conclusions, nor using signal amplitude as an indicator of cognitive engagement. It seems more appropriate to try to interpret trends (increases, decreases, peaks, etc.) because, as mentioned by Boucsein et al. [2], a variety of factors affect EDA measurements such as weight, height, exercise, hydration, diet, etc. Finally, the ethical certification issued for this pilot project prevented us from linking any data collected to a specific student. As a result, we were not allowed to contact Student #5 to try to understand why its signal was steadily increasing during the exam (Fig. 6).

6. CONCLUSION

This paper presents the results of a pilot project aimed at analyzing the potential of wrist-worn EDA sensors used to estimate the level of cognitive engagement of students in the classroom. The objective of the pilot project was to collect preliminary data in order to develop a robust protocol to maximize the validity of data subsequently collected in a large-scale study.

The data collected during the pilot project show that it is imperative to know the participants' lifestyle habits (physical activity, caffeine and medication consumption, stress levels, etc.) in order to properly interpret the data. In addition, means should be deployed to be able to identify the events that generated variations in the EDA signal, for example by recording the class periods or organizing individual interviews with participants. It would also be desirable to develop a way to compare EDA signals based on metrics (area under the curve, mean values, etc.), not on global, qualitative comparisons of curves versus time.

The large-scale research project was supposed to be carried out during the Summer and Fall 2020 semesters, but unfortunately, restrictions due to the Covid-19 pandemic forced all the courses to be given online and remotely. Since this publication is based exclusively on preliminary data, it is impossible to conclude, on one hand, that the Empatica E4 wristbands is an ineffective technology in measuring the level of students' cognitive engagement. On the other hand, it is possible to conclude that this technology is far from being "plug and play". Indeed, important control and follow-up measures with participants must be included in the experimental protocol for a proper interpretation of the results.

7. REFERENCES

[1] Azevedo, R., "Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues," *Educational psychologist*, vol. 50, no. 1, pp. 84-94, 2015.

[2] Boucsein, W., Fowles, D.C., Grimnes, S., Ben-Shakhar, G., Roth, W.T., Dawson, M.E. and Filion, D.L., "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017-1034, 2012.

[3] Critchley, H.D., "Electrodermal responses: what happens in the brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132-142, 2002.

[4] Dirican, A.C. and Göktürk, M., "Psychophysiological measures of human cognitive states applied in human computer interaction," *Procedia Computer Science*, vol. 3, pp. 1361-1367, 2011.

[5] Doberenz, S., Roth, W.T., Wollburg, E., Maslowski, N.I. and Kim, S., "Methodological considerations in ambulatory skin conductance monitoring," *International Journal of Psychophysiology*, vol. 80, no. 2, pp. 87-95, 2011.

[6] Garbarino, M., Lai, M., Bender, D., Picard, R.W. and Tognetti, S., "Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition," *4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies*, 2014.

[7] Larmuseau, C., Vanneste, P., Cornelis, J., Desmet, P. and Depaepe, F., "Combining physiological data and subjective measurements to investigate cognitive load during complex learning," *Frontline Learning Research*, vol. 7, no. 2, pp. 57-74, 2019.

[8] McNeal, K.S., Zhong, M., Soltis, N.A., Doukopoulos, L., Johnson, E.T., Courtney, S., Alwan, A. and Porch, M., "Biosensors Show Promise as a Measure of Student Engagement in a Large Introductory Biology Course," *CBE—Life Sciences Education*, vol. 19, no. 4, 2020.

[9] Menghini, L., Gianfranchi, E., Cellini, N., Patron, E., Tagliabue, M. and Sarlo, M., "Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions," *Psychophysiology*, vol. 56, no. 11, 2019.

[10] Pijeira-Díaz, H.J., Drachsler, H., Kirschner, P.A. and Järvelä, S., "Profiling sympathetic arousal in a physics course: How active are students?" *Journal of Computer Assisted Learning*, vol. 34, no. 4, pp. 397-408, 2018.

[11] Poh, M.-Z., Swenson, N. C. and Picard, R.W., "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE transactions on Biomedical engineering*, vol. 57, no. 5, pp. 1243-1252, 2010.

[12] Potter, L., Scallon, J., Swegle, D., Gould, T. and Okudan Kremer, G.E., "Establishing a Link between Electrodermal Activity and Classroom Engagement," 2019. From https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1203&context=imse_conf

APPENDIX A: FIGURES

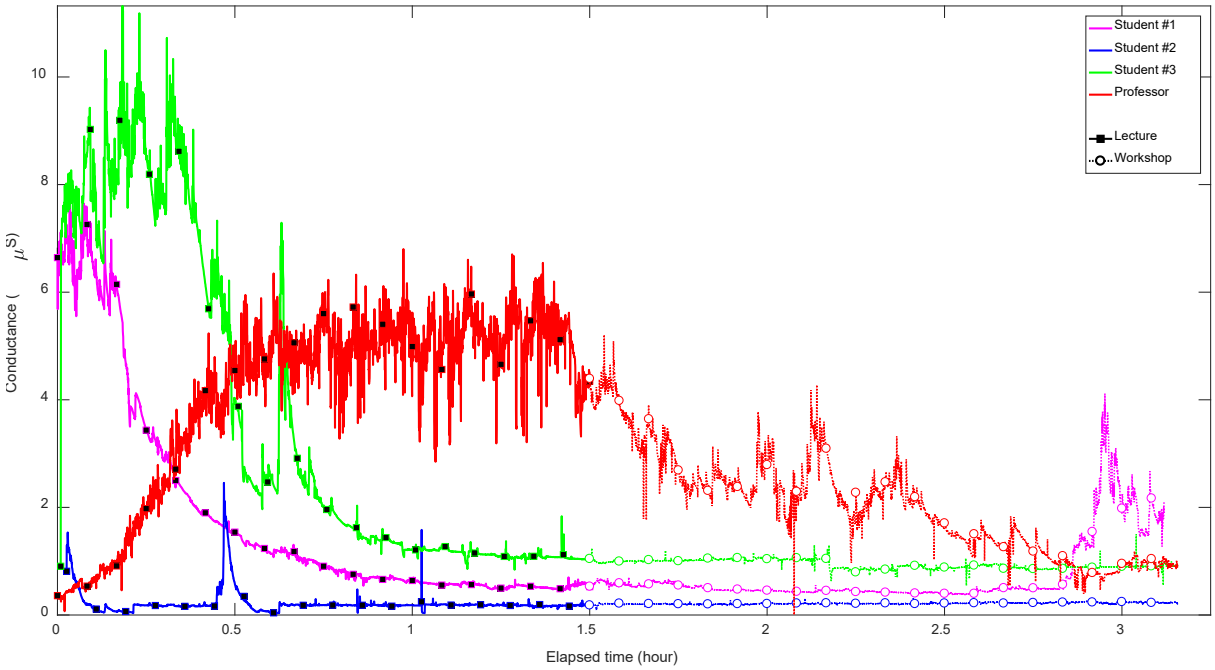

Fig. 1. EDA curves - Week 4 - Group 1.

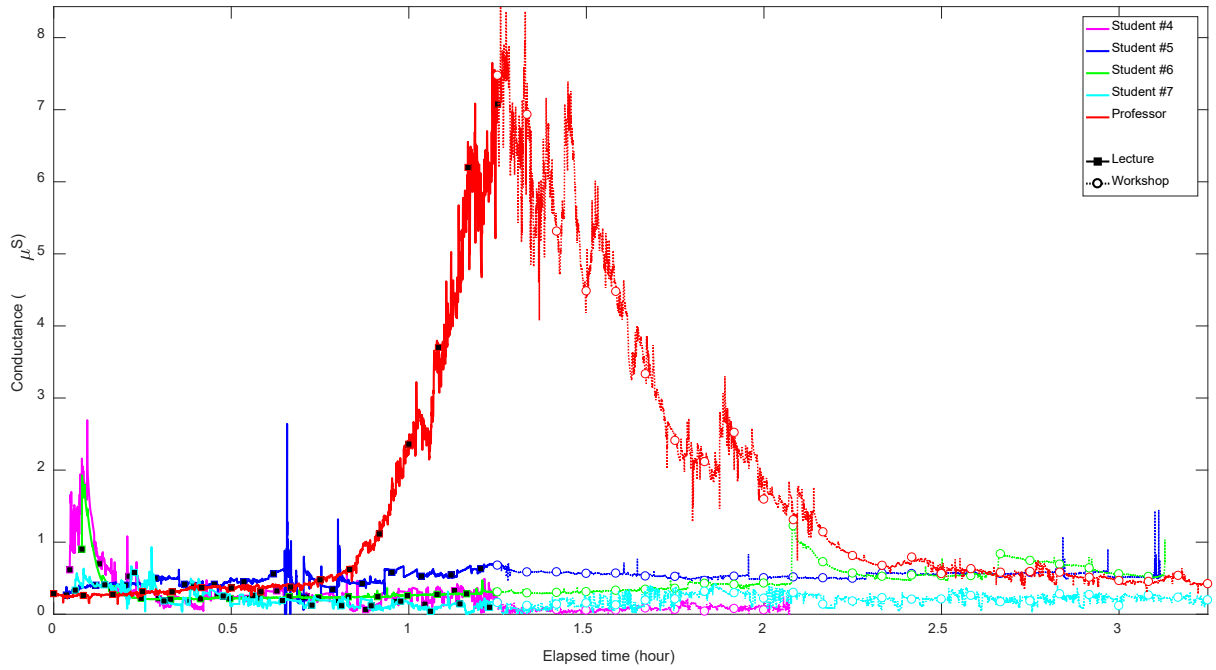

Fig. 2. EDA curves - Week 4 - Group 2.

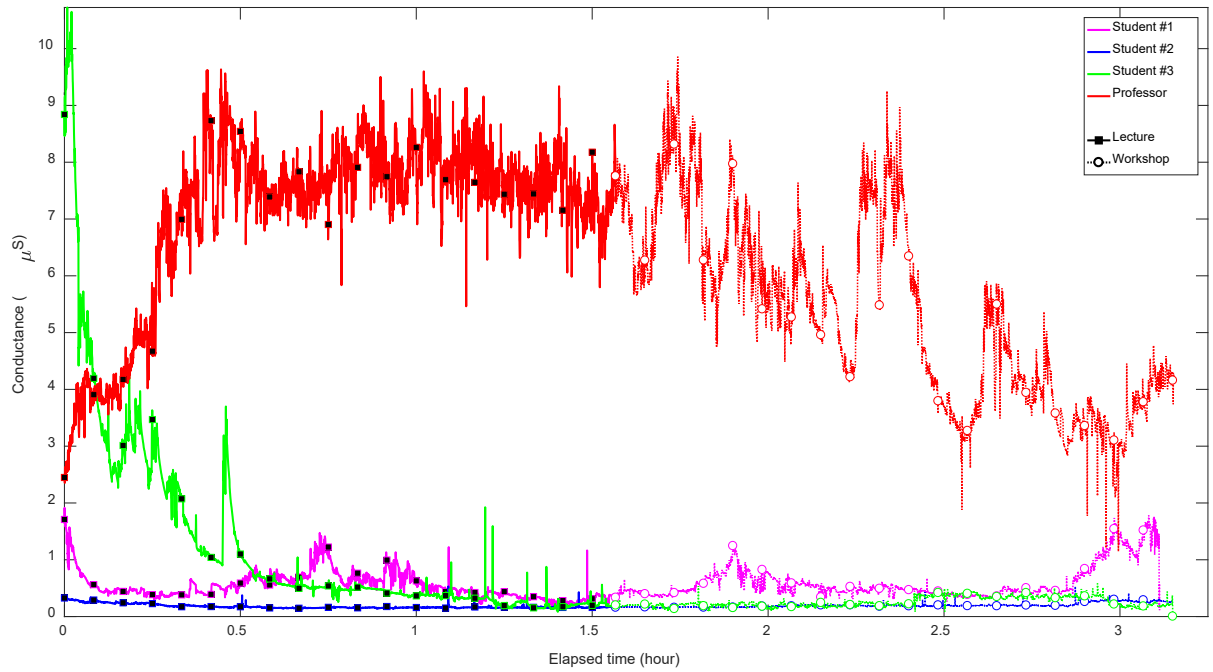

Fig. 3. EDA curves - Week 6 - Group 1.

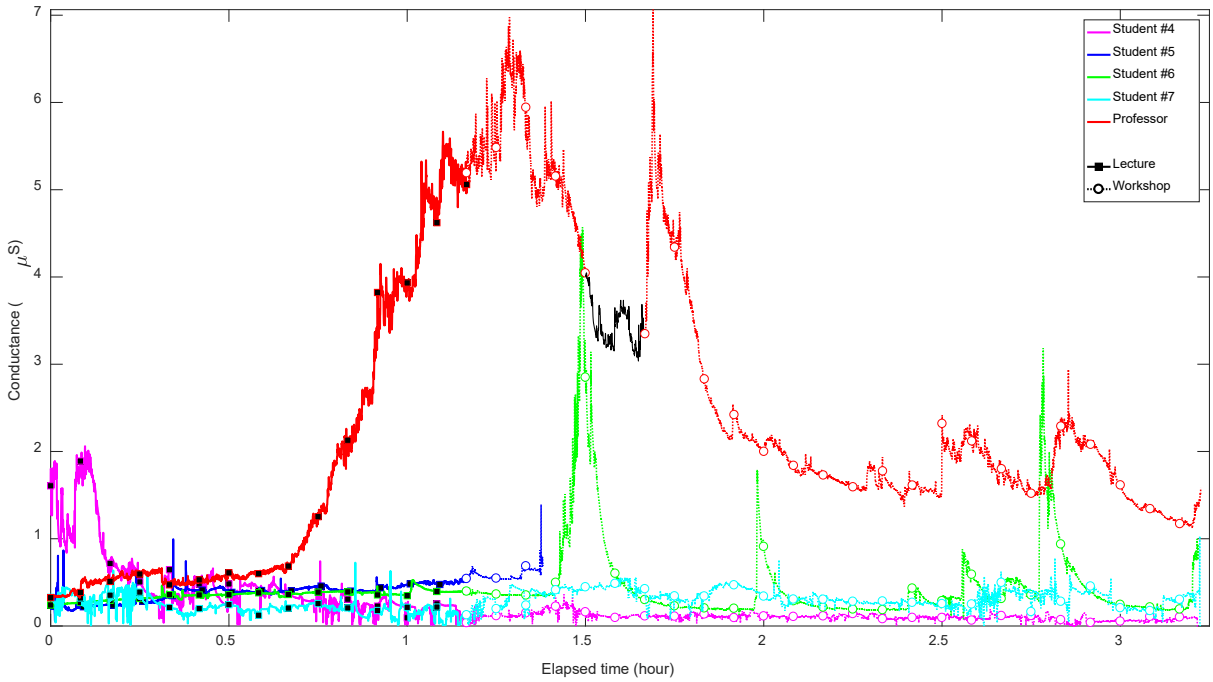

Fig. 4. EDA curves - Week 6 - Group 2.

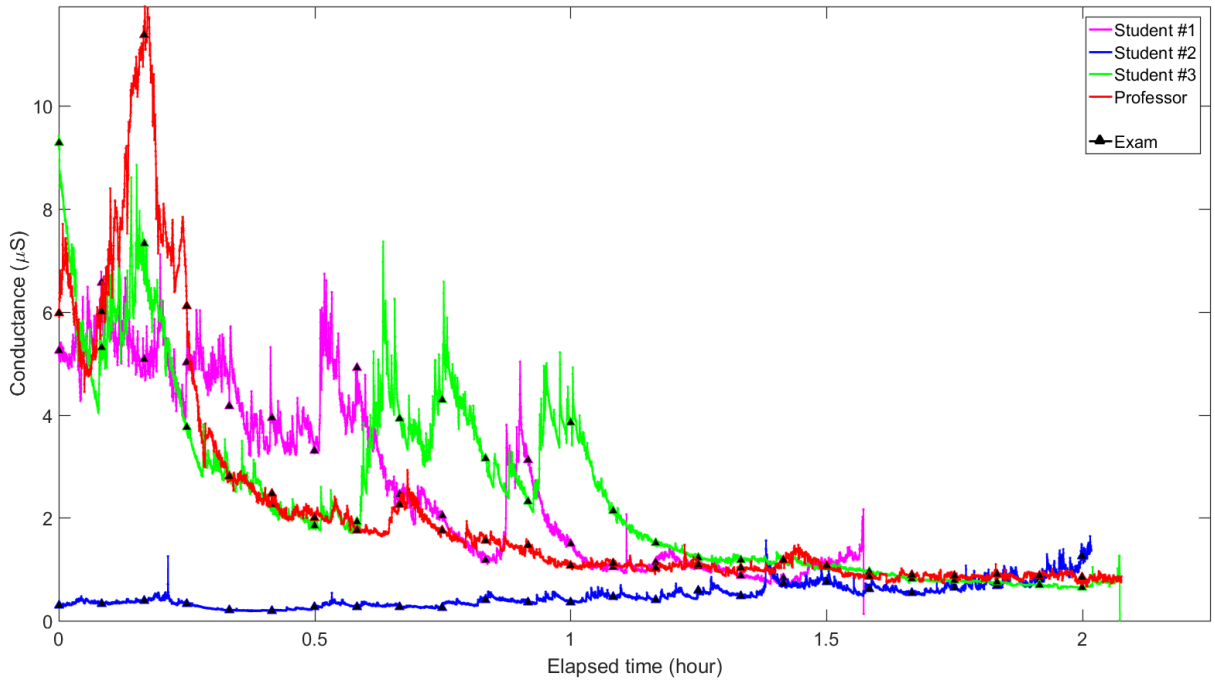

Fig. 5. EDA curves - Week 7 - Group 1.

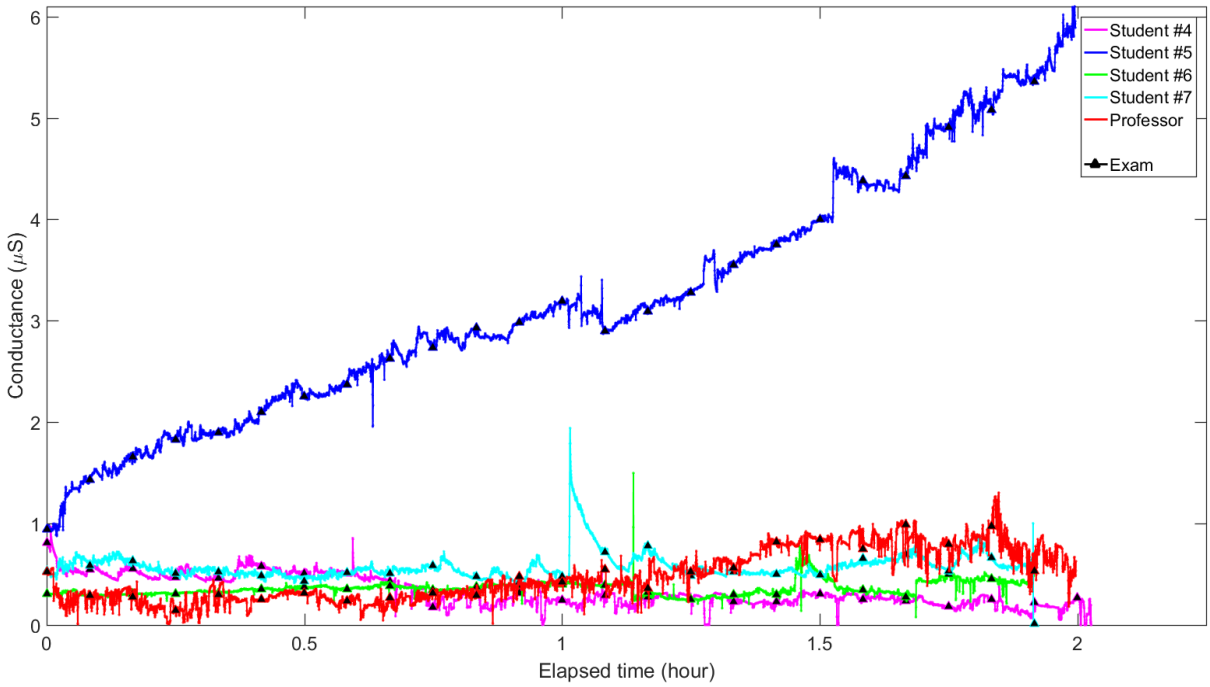

Fig. 6. EDA curves - Week 7 - Group 2.

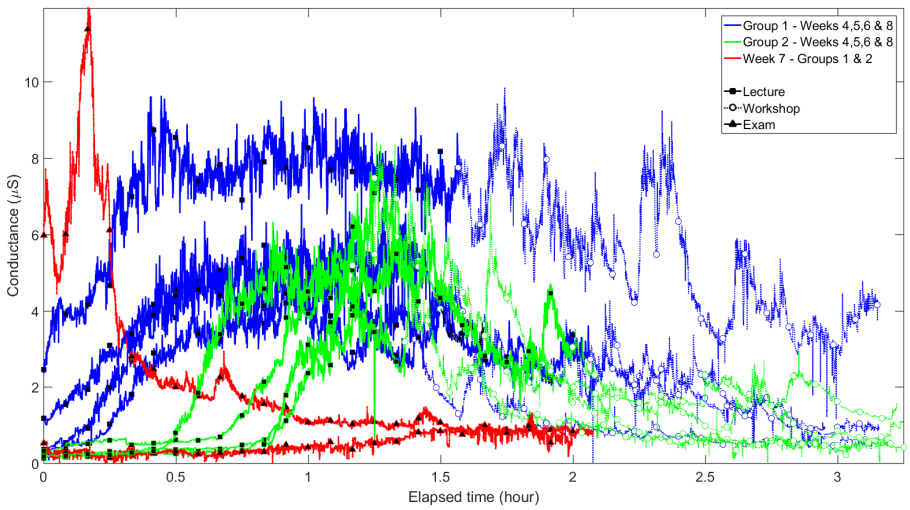

Fig. 7. Comparison of the EDA curves of the professor (Groups 1 and 2).

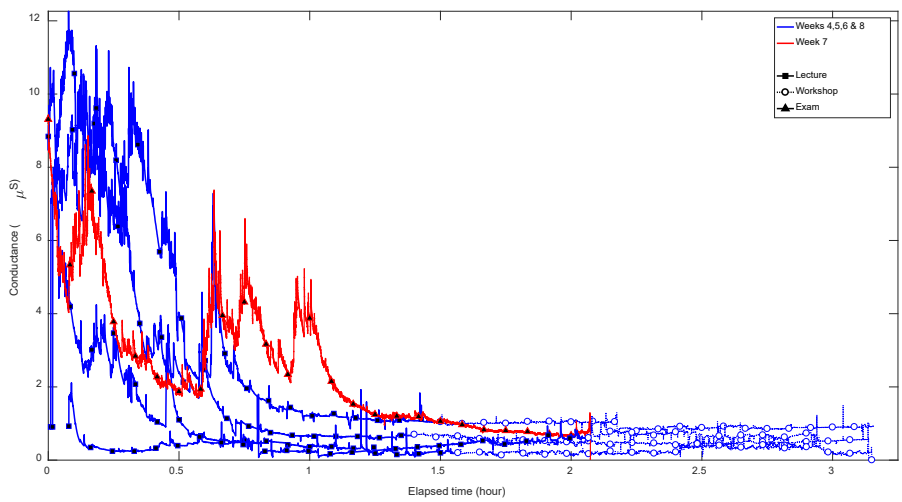

Fig. 8. Comparison of the EDA curves of student #3 (Group 1).

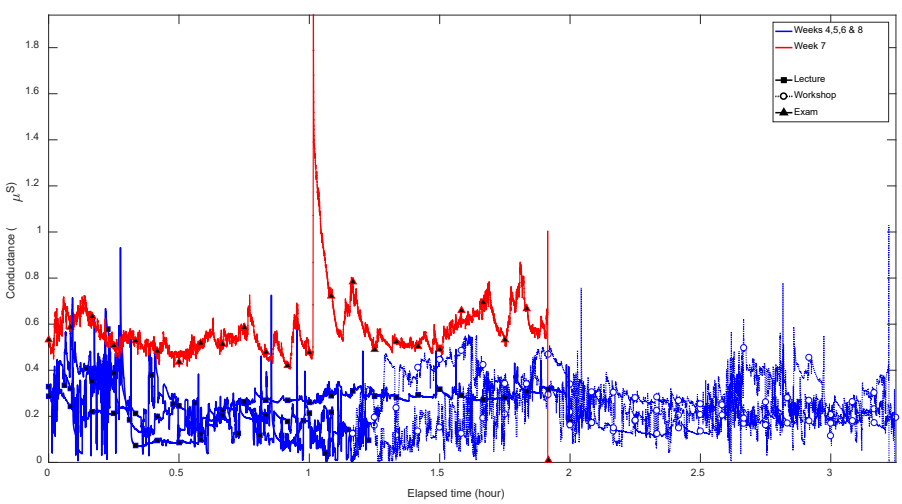

Fig. 9. Comparison of the EDA curves of student #7 (Group 2).