ABSTRACT

Human body-language is one of the richest and most obscure sources of information in inter-personal communication which we aim to re-introduce into the classroom’s ecosystem. In this paper we present our observations of student-to-student influence and measurements. We show parallels with previous theories and formulate a new concept for measuring the level of attention based on synchronization of student actions. We observed that the students with lower levels of attention are slower to react than focused students, a phenomenon we named “sleepers’ lag”.

Categories and Subject Descriptors
K.3.1 [Computer Uses in Education]: Computer-assisted instruction (CAI); K.3.m [Computers and Education]: Metrics—performance measures, Miscellaneous

General Terms
Measurement, Orchestration, Classroom attention

Keywords
Motion lag, audience synchronization, classroom attention, classroom orchestration

1. INTRODUCTION

What is a good indicator of comprehension? Not all students are willing to participate during the class period. At times the most passive of students can surprise you with their knowledge, and the most active students will underachieve. How to see this during the class when you have the best chance of correcting a misconception in the lecture’s material?

Our goal is to provide teachers with a new view of the classroom, one that lets them assess the impact of their lecture. This is important because the element that will change on every lecture is the audience mood - and the lecture needs permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

LAK ’14, March 24 - 28 2014, Indianapolis, IN, USA. Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-2664-3/14/03S15.00. http://dx.doi.org/10.1145/2567574.2567581

to change with it in order to be effective [4]. The adjustment of the teaching style can prove to be a challenging task for a novice teacher. While aiming to give each student the amount of attention needed to learn qualitatively, we would like to point the lecturer to the parts of the classroom which he or she is not engaging in the lecture. But with all the different styles of lecturing, how to measure performance without crippling the teachers with technological aids that distract more than help?

Our approach is based on the attempt to formalize pure observation. Without overloading the students with gadgets and formally structured procedures that dictate a certain format of the learning experience, we start with a system of cameras and measurements of human activity in it’s most basic form - movement.

In this paper we present the method for measuring movement in a class and the procedure used to relate the gathered information to the students’ subjective perception of attention. The main contribution is the concept of measuring the speed of student reactions in class to detect the students with lower attention. The concept is based on the idea that the students who are focused on the lecture would react at the moment the important information was presented, while the distracted students will be slower to note it. This is the concept that we call the “sleepers’ lag”. The higher the variability of reaction time to the stimuli (in our case - teacher’s presentation) - the lower the attention of the classroom. In order not to introduce some artificial measuring points during the lecture we propose to compare students against each other in a dyadic fashion.

Our other conclusions go further into exploration on how the geometry of the classroom and immediate surroundings affect the individual student. This sets up the ground for the “student-centred” observation of the classroom, as opposed to the dominant trend of exploration which considers the teacher as the only stimulus present in the classroom of approximately fifty persons (to which we refer to as “teacher-centric”).

1.1 Theoretical Background

Traditional classrooms (in both talk format and seating configuration) remain the dominant format of lecturing on all levels of formal education today. There have been many critiques on the format, saying that classroom’s geographical configuration makes it difficult to develop the teacher-student relationship and understanding beyond stereotypes [17]. And while some claim that the current organizational setup evolved into existence for practical reasons [20] we
can’t ignore difficulties that teachers are facing with keeping students attention over time [31][22] and on-task[27].

The set of theories we group under the name “teacher-centric” focus on teacher and the teachers impact on the classroom. As the primary orchestrator of the learning process [10][11], teachers take on the responsibility that starts from educational presentation, pedagogical guidance [7] to the level of students’ personal transformation [30]. Teacher’s role in the classroom has been characterized as emotional labour [17] and cognitively demanding [13]. In many occasions a good teacher is characterized with the ability to present the teaching material in a way that is acceptable to students, and this has been shown to be the major difference between a novice and a experienced teacher [4], confirming the need for the teacher to be a reflective practitioner [28].

Teachers still need to go beyond the instinctive reactions [19] in order to reach all the students. The geometry of the classroom, as mentioned before, was already presented as an emotional barrier for more natural interaction [17]. Students in the front rows are perceived as “more interested”[8]. Majority of communication is oriented to the T-shaped region with the highest concentration of interaction focused on the front-center of the classroom [1]. The effect does not only affect the teacher’s perception, but students also adjust to the geometry of the classroom, and the students who seek interaction will tend to sit in the high-interaction places [3]. It’s also shown that the seating arrangement will act as an amplification of students interactions - making high-verbalizers more active in the high-interaction zone, and making low-verbalizers even less active in the low-interaction zone (the edge regions of the classroom) [20]. The classroom environment greatly affects the perception of the teacher and students, but this doesn’t always go in favour of the learning process.

Being far away from the teacher goes beyond just teachers perception. On the “student-centric” side of research, Daum [9] found that distance from the teacher also has a real effect on the success of the students. Irrespective of the position or grades, students have difficulty maintaining the attention during the whole duration of the class [27]. Even if it’s not clearly quantifiable after how much time students lose attention, proposed measurements between 10 minutes [31] and 20 minutes [22] are far less then the average duration of class period. Finn et al. [14] found that smaller class sizes (less then 15 students) affect the quality of the lecture in two ways - the teachers took less time to manage the learning process, but more importantly the students’ interaction between themselves also changed for the better. As the students grow up in the school system, the relationship between the teacher and the students becomes less emotionally involved [17] and their participation in the classroom activities decreases [21]. This seems closely related to students becoming more accustomed to studying in a large group, where individual visibility is questioned [14] and situation is easy for diffusion of responsibility and social loafing [15]. It’s common for the students to have more practical goals than purely academic improvement [2].

Our research aims to scaffold teacher’s perception of the students and raise awareness about students’ reception of the lecture. Some of the current methods of dealing with the problem use web-oriented solutions [12], feedback devices in form of a clicker[6] or mobile phone application [26]. We look up to the research of unobtrusive measurements [29] in order not to disturb the classroom eco-system.

2. THEORETICAL ASSUMPTIONS

Our initial hypothesis for the experiment was that we can detect agreeing groups of students by common behavior patterns and that the people in the visible surroundings affect the person (student) by their non-verbal cues. We considered body language in it’s most basic form and compared the cooccurrences of motion between pairs of students. We also related our observations to students’ level of attention.

From the dual eye-tracking theory, we know that the quality of collaboration [24] and understanding [18] between two persons can assessed by analysing the consistency of their gaze patterns. We draw an analogy with these conclusions in

Figure 1: Motion intensity graphs. Horizontal axis represents the time and vertical axis 0-100% of relative motion of the person (explained in Section 3.1) a) Example of co-movement for two persons. Person 2 shifted hers seating position (blue line), 2 seconds later, neighbouring Person 1 (marked in green) also started re-adjusting herself. b) Motion of a single person (dark green trace) overlayed over the average motion of the whole classroom (gray trace). Horizontal red line marks the 30% threshold that we used for movement analysis. Color-coded labels on top indicate different events during the class, as described in Section 3.2. Annotations present here are: Blue rectangles - slide chage; Red periods - question answering periods or questionnaire filling periods; Green vertical lines - slide animations
the domain of motion in the classroom, with the hypothesis that students who listen to the teacher will be more likely to move in synchronized manner, while an absent-minded student will act on his/hers own internal rhythm. Synchronized motion is not limited to any specific action, but can be explained on the example of taking notes - attentive students would turn pages on the hand-outs and note important facts as they are presented in class. More than a reaction to the lectures audio/visual stimulus, motion can be seen as an agreement of the audience. If the students agree that an event outside the classroom (e.g. loud noise, truck) is more important than the lecture, they would still have a synchronized motion (everybody looking through the window) but caused by a different stimulus than the teacher.

Synchronization in the class was studied in a dyadic fashion, by comparing pairs of any two students. Depending on relative location between the two students we divided the dyads into three conditions based on their mutual visibility (as described in Section 3.3).

Given that learning is not a strictly formalized activity, reactions of students can vary or be completely blank. In dual eye-tracking, a delay of 2 seconds between the speaker’s and listener’s gaze during the moments of referencing has been identified [25], with the conclusion that the comprehension between participants is inversely proportional to the time-lag. Based on this, we define two movements of the students as co-movement if it happens within ±4 seconds from each other (depicted in Figure 1a). We differentiate between i) perfect synchronization (< 2 sec apart), ii) synchronization (2-4 sec apart), iii) weak synchronization (4-6 sec apart). Three periods are displayed on the Figure 2b as the vertical axis.

The additional, third, period was introduced to take into account indirect synchronization - when the person is not reacting to the teachers stimulus but is following the reaction of others, in which we add 2 seconds for the person to observe the reaction of others and then reproduce it. This is what we call the “sleepers lag” - the idea that the persons who are mimicking attention instead of actually paying attention will have a delay (a “lag”) in their actions.

3. METHOD

Our setup and method for gathering data is novel in the classroom environment. We will describe the main points of the technological part of the method, the data-gathering methodology and our current working sample.

3.1 Motion Analysis

Analysis of the motion is based on tracking feature points in the video [5]. Our setup consists of three cameras used for coverage of the students and one observing the teachers actions. Initial steps of analysis - synchronization of video streams from all sources and annotating visible regions in which students reside during the lecture are described in [23].

Our main challenges in the process of extracting a measurement of motion for further analysis were i) inter-personal occlusions, ii) perspective distortion, iii) normalization of the amount of movement recorded from a single person into a comparable measurement between several persons.

i) Inter-personal occlusions are handled by taking several pre-processing steps before assigning the motion to a person. The main idea is that by grouping the motion vectors into motion tracks (T) which consist of "cloud" of motion vectors over several frames. The criteria for grouping is based on prox-
Figure 3: Motion detection and grouping. a) Individual motion vectors shown as purple arrows b) Motion vectors grouped over time into motion tracks which can be assigned to an individual c) Marked student areas and centres of Gaussian probabilities which model the probability of motion belonging to each student.

Imity, direction similarity and intensity of the vectors. For visualisation purposes, a set of cloud centres from several frames are connected into a track which is shown in Figure 3b. Finally the entire track is assigned to the student of highest probability \( g_f \) defined by Formula 1. Each student \( g \) has a Gaussian distribution centred on the position of his head (depicted in Figure 3c). The entire track is assessed over every center (i.e. student) and motion is assigned to the student with the highest probability.

\[
g_f = \arg \max_g \sum_{v \in T} p(v|g) \tag{1}
\]

In cases where a student was occluded on more than 80% of tracked area the movements were undistinguishable from the person in front of him/her. Depending on the quality of the measurements for the person in front, either one or both students were removed from further processing if they were below a set threshold.

Taking into account that our primary interest was motion between students, it is important to notice that this method was designed so that

- a motion occurring between two students would not be assigned to both students
- large motions spanning several tracked areas would be assigned to a single person, and not to a group of people.

ii) Perspective distortion To compensate for the perspective effect, the number of tracking points remains constant over all annotated tracked areas. Second measure was to normalize the intensity of the motion vector by the diagonal of the student region. This ensures that the hand-motion of the student in the back row will be registered with the same intensity as the hand-motion of the student in the first row.

iii) Normalizing the amount of motion Normalizing the motion of a person has proven to be difficult. We based our normalization on two premises i) student is on average sitting still during the class ii) student has at least one full-body movement in the recorded footage (e.g. pose shift). To scale this to a range of 0-100% motion, we take the median value of movement intensity as the 5% motion (which corresponds to small motion/sitting still being registered as 5% motion), and we verify that given this basic motion intensity the student reaches 100% motion at least once during the class. Motion which registers above the threshold of 100% is clipped to the maximum value. The final motion intensity over time can be visualized as shown in Figure 1b.

3.2 Experimental procedure

We observed each lecture for the duration of 30 minutes. After a random interval (average duration 7 minutes) a tone signal was given which interrupted the lecture. At that time the students were asked to fill out a questionnaire sampling their activities and self-reported perception of the classroom.

In addition to student samples, we hand-annotated events during the class, which were a product of teacher’s action or a teacher-student interaction. Events were annotated into following categories: i) slide change ii) slide animation iii) Question begin/end period iv) Answer begin/end period v) Other events. Our questionnaire filling periods (which typically lasted around 1 minute) were designated as “Question answering” periods. Since they don’t represent a normal part of a lecture, student activity in those periods were not taken into consideration in further analysis of data. The events are shown as annotations in the top part of the timeline visualization in Figure 1b.

3.2.1 Questionnaires

By using a 10-point Likert scale, participants were registering their

- attention level,
- perception of the teacher (energetic/boring),
- perception of the classroom attention (high/low),
- importance of the material being presented (important/irrelevant).

In addition to this, the questionnaire enumerated activities that the students did during the previous time-period:

- listening,
- taking notes,
- repeating key ideas,
- thinking about other things,
- interacting with people around you (which was not scheduled by the classroom activity),
- using your laptop/phone.

Students could check more than one activity.
3.2.2 Student sample

We base our results on analysis of two classes, described in Table 1. Both student groups were in the bachelor program of EPFL, Switzerland. The teachers were two experienced lecturers teaching on subjects from social sciences in case of Class 1 and technical sciences in Class 2. The lectures were given in different times of the day - one being in the morning, the other in the late afternoon time, and in different rooms.

<table>
<thead>
<tr>
<th>Class</th>
<th>Size</th>
<th>Analysed</th>
<th>Female rate</th>
<th>Rows</th>
<th>Columns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>29</td>
<td>36.8%</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>14</td>
<td>22.2%</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1: Basic information about analysed classes

Even though we initially consider both classes comparable, small number of students in the second class rendered conclusions from that observation to be statistically invalid. We show the results found in the Class 2 sample here to demonstrate that there is a consistent trend in both cases.

3.3 Location and surroundings

One of our main considerations when thinking how the student perceives the lecture and came from proxemic zones [16]. Since the perception of the teacher changes significantly depending on how far they are from the front, we decided against normalizing the space in the way it was done for [1], which would allow us to make one big sample by making the two classes comparable.

Emulating the proxemics concept in the classroom environment we defined 3 zones depicted in Figure 4:

- immediate neighbour - which models “personal space”. Person to the immediate left or right of the student, with whom the student shares the desk-space and leg-space. This is partially dictated by the dimensions of the student desks which are made for two persons per desk;
- visible neighbourhood - represents the zone of two rows in front of the student ±2 persons wide. This represents the “social zone” of proxemics theory (which spans from 1.2m - 3m). The zone practically models the people who would be intentionally or unintentionally observed by the student who’s following the material on the slides or looking towards the teacher;
- non-visible students - students who are either too far to the side or behind the individual and can not be seen without intentional action.

4. OBSERVATIONS

4.1 Questionnaire data

The collected questionnaire data was used primarily as the basis for further analysis of the collected video material. Nevertheless we report the condensed findings to depict the general situation in the classrooms. A general note on the findings is that because of the small number of samples, we are reporting our findings with Kendall’s correlation.

Reported levels of attention in both cases were high, with the mean shifted to around 7. In case of Class 1 $\mu = 6.822, \sigma = 2.344$, and in case of Class 2 the normal distribution has parameters $\mu = 7.444, \sigma = 1.100$ (shown in Figure 5).

There is a significant correlation between the personal level of attention and the perceived level of attention of the entire class (Class 1 $\tau = 0.477 (p < 0.05)$; Class 2 $\tau = 0.413 (p < 0.05)$). We considered this to be an interesting way of expressing dissatisfaction with personal or class performance in way that the student would mark bigger difference between personal and classroom attention if there is a bigger dissatisfaction with the learning conditions.

Classes were perceived by participants as exhibiting both high teacher-energy and high levels of student attention.

<table>
<thead>
<tr>
<th>Class</th>
<th>Class attention mean (variance)</th>
<th>Teacher energy mean (variance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.776 (3.711)</td>
<td>7.783 (1.866)</td>
</tr>
<tr>
<td>2</td>
<td>7.125 (3.266)</td>
<td>8.347 (1.920)</td>
</tr>
</tbody>
</table>

Table 2: Parameters of perceived class quality

Activities students reported (shown in Figure 6) show an expected tendency to report material related activities (listening to lecture, taking notes and repeating ideas) in higher attention levels. Off-task activities (“thinking about other things”, “talking to others”) were reported on all levels up to the maximum level of attention. Note that the students in Class 2 were using tablets as part of their regular studies to view the class material, while that was not required for Class 1.

We also studied the variation of attention levels over time in hope of capturing the reported drop of concentration after 10 minutes [31], but found no clear trend.
Figure 5: Average attention of students in both classes was subjectively perceived as high. a) Class 1 ($\mu = 6.822, \sigma = 2.344$) and b) Class 2 ($\mu = 7.444, \sigma = 1.100$).

4.2 Motion data

We compared the number of synchronized movements during the class period. Synchronized movement is defined as body movement with more than 30% intensity from each of the two persons being compared (shown as the horizontal red line in Figure 1b). The 30% threshold was taken to separate minor body movements and motion that is likely to be noticed by others. We took into consideration the visibility of two persons, meaning that in order for the movement of Person 1 to be considered as stimulus, it has to be visible to Person 2. Visibility reasoning was done based on the sitting location of the two persons.

We compared the average number of synced movements between pairs sitting immediately next to each other (marked as the dark-blue region in Figure 4) and other pairs. We found that immediate neighbours had higher probability of synchronized moving than a non-neighbouring pair with a $t$-test ($p \leq 0.05$), shown in Table 3.

We analysed but found no significant difference in number of synchronized moments between the pair from a visible neighbourhood and the non-visible students.

To compare the motion metrics with the previous findings of Adams [1] on student activity, we also tested the influence of teachers proximity to the movement of the students. The further away students are from the front-center of the classroom (the point which is the closest to the teacher in both cases, represented as distance $d$ in Figure 4) the less active they are (Kendall correlation is $\tau = -0.284$ ($p = 0.03$) for Class 1; and $\tau = -0.172$ ($p = 0.45$) for Class 2). Analysing the samples we have seen the same trend in both cases, even

<table>
<thead>
<tr>
<th>Class</th>
<th>Neighbouring pairs mean (variance)</th>
<th>Other pairs mean (variance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>76.54 (32.47)</td>
<td>54.43 (15.64)</td>
</tr>
<tr>
<td>2</td>
<td>63.33 (24.33)</td>
<td>44.88 (18.42)</td>
</tr>
</tbody>
</table>

Table 3: Average number of synchronized moments between immediate neighbours and other pairs.
Figure 7: Correlation between distance from teacher and motion intensity in Class 1; Kendall correlation $\tau = -0.284 \ (p = 0.03)$

tough the correlation was insignificant for the second classroom. Figure 7 shows the correlation for Class 1.

Our third test was to find the correlation of the average reported level of attention to the reaction speed. The question was whether students with lower attention levels were more likely to lag behind a other students in their visible field. The correlation found had the expected trend in Kendall correlation $\tau = -0.259$ but was marginally insignificant ($p = 0.06$) on the sample size of 29 students of the Class 1. The result is shown in Figure 8. Class 2 correlation had a similar trend but was not statistically significant $\tau = -0.222 \ (p = 0.32)$. The data thus suggests that there is a phenomenon of “sleeper’s lag”, but the current sample is not conclusive.

5. CONCLUSIONS

In this paper we demonstrated our concept of measuring speed of reaction in the student population of classroom. We gave insight about the subjective perception of the classroom attention gathered with the questionnaire, and shown that the students will project their level of attention onto others in their reports. Our first conclusion about existing synchronization of motion between immediate neighbours shows that there might be an underlying reason for this, and that two persons can affect each other just by sitting together without actual direct interaction.

We found a similarity with previous studies on the affect of teacher proximity on students [1] and found that students who are further away not only participate less, but also move less.

Finally, we proposed a new way of evaluating the overall attention of the classroom by comparing pairs of students and analysing how synchronous they move. By comparing the motion results to the data gathered in the questionnaire we showed that there is a correlation between slower reaction time and lower levels of reported attention - the “sleeper’s lag”, but our data was not conclusive.

We haven’t touched on the subject of presenting the information to the teachers during the lecture, and we’re planning on starting a dialogue with the participating teachers to find the best representation for displaying the information during the lecture. Our next steps are to confirm the findings on a broader sample of students and continue to refine the technological methods. In addition to the “sleeper’s lag” we would also like to explore further the phenomenon which we call the “distraction ripples” - assuming the transitivity of the motion-syncing, we would like to capture the spreading of influence of one class-member to people around him/her. We are also interested to correlate how well do these “ripples” spread in high-attention and low-attention groups of students, in order to make a new metric of class attention.

In addition to motion, we aim to introduce additional cues into our reasoning about student attention and perception of the class, specifically - gaze direction. The end goal is to have a holistic image of the classroom life in order to find the most salient cues that can be unobtrusively collected. Our intention is that in the end, the entire system would act as a training experience for the novice teachers.

Stepping back from the trend of individual learning with massive on-line open courses (MOOCs), classrooms remain the dominant form of lecturing on all educational levels. Introducing technological solutions to the classroom has the potential of huge impact on the way the students learn. By not excluding the teachers but supplementing their observations with advanced measurements we hope to create a blend which would be superior to the current methods and beneficial for both students and teachers.

6. ACKNOWLEDGEMENTS

This work has been sponsored by the ProDoc SNF Grant, project PDFMP1 135108. We would also like to thank all of the experiment participants (students and teachers) for their help in gathering the data.
7. REFERENCES