"With-me-ness": A gaze-measure for students' attention in MOOCs

Kshitij Sharma[1], Patrick Jermann[2], Pierre Dillenbourg[1]
1. Computer Human Interaction in Learning and Instruction (CHILI)
2. Center for Digital Education (CEDE)
École Polytechnique Fédérale de Lausanne, Lausanne
Email: {kshitij.sharma, patrick.jermann, pierre.dillenbourg}@epfl.ch

Abstract: We propose a gaze-based indicator of students' attention in a MOOC video lecture. We report the results from an eye-tracking study during a MOOC lecture. We define the gaze-based indicator of students' attention as "with-me-ness". This answers a question from teachers' perspective *"how much are the students with me?"* With-me-ness is defined at two levels: perceptual, following teacher's deictic acts- and conceptual – following teacher discourse. We conducted an experiment with 40 participants and observed a significant and positive correlation between the two levels of with-me-ness and the posttest scores.

Keywords: Eye-tracking, MOOCs, Learning analytics.

Introduction

We use eye-tracking methods to measure the students' attention during MOOC lectures. In the background, the general question we address is "how can we help students to watch videos more efficiently?" In this study, we relate attention measures to how much students learn from a simple video.

There are many factors affecting classroom performance of the students: previous grades (Astin, 1971), students' efforts and motivation (Grabe and Latta, 1981), socioeconomic differences (Kaplan, 1982), quality of schooling (Wiley, 1976), attention (Good and Beckerman, 1978;) and participation (Finn, 1989). We do not claim we address all of them. We focus on attention as one dimension of students' performance that may be especially relevant when watching video lectures. We tackle this question from a teacher's perspective: "*How much the student is with me?*" Accordingly, we call this gaze-based measure "With-me-ness": is the student really "following" the lecturer, i.e. paying attention to the elements of the display that correspond to the instant behavior of the teacher? We selected two aspects of teacher's behavior that will influence the student's attention, what the teacher says and the teacher's deictic references: in several MOOCs, learners may see the hand or the pen of the teacher in overlay to the slides.

Related Work

Attention as Factor affecting Students' Performance

Attention had been proved to be associated with performance in many studies related to visual tasks (Yantis and Jonides, 1984; Prinzmetal et. al., 1986; Juola et. al., 1991). In a visual comparison of two line segments Prinzmetal et. al. (1986) and Juola et. al., (1991) showed that the more attentive participants were more correct in selecting the longer line segment. Yantis and Jonides (1984) found similar results in visual perception tasks.

In a classroom attention is *"listening, sitting and working on assigned tasks" -Homes et. al. (2006)*. In the context of academic performance previous research has shown strong association between students' attention and academic performance (Finn, 1989).

We do not claim that visual attention is a deep indicator of learning activity: looking at a piece of information is often a condition to interact with it (read it, select it, move it) but it does not indicate how deeply the learner processes this information (does he understand or give meaning to it). However, in the context of video watching, it's one of the rare behavioral information that can be collected. Moreover, as we will explain, "withmeness" is not simply measuring attention but co-attention, i.e. whether the learner is paying attention to the elements that the teacher is referring to, verbally or through deictic.

Eye-tracking and Performance/Expertise

Previous research provides insights about the relationship between the gaze patterns and the behavioral and performance indicators in diverse scenarios. Existing results show a clear relation between gaze patterns and expertise. In a collaborative Tetris game, Jermann et al. (2010) showed that experts pay more attention on the stack of Tetronimoes where pieces land while novices allocate more attention to new pieces falling from the top.

Existing results also show a clear relation between gaze patterns and task based performance. In a pair-programming task, Jermann et al. (2012) showed that the pairs with high quality collaboration have synchronized more their gaze on different parts of a program than the pairs that do not collaborate well. In a similar task, Sharma et al. (2011) showed that the good performing pairs pay more attention to the data-flow of the program than the poor performing pairs. Moreover, Sharma et al. (2013) showed that while describing the functionality of a

program the good performing teams had more gaze on the expressions in the program while poor performing teams have equal distribution of gaze on different parts of the program during similar phase of the task.

The Present Study and Research Question

We present the results of an eye-tracking study contextualized within a MOOC class. We choose MOOC videos as a stimulus for the eye tracking because the effectiveness of video as a medium for delivery of educational content is a controversial issue in the current rise of online education. This issue has already been studied and established in literature (Paivio (1991), Mayer (2003), Schwartz (2007)). Through this contribution we propose a gaze-measure to capture students' attention (introduced in the section "With-me-ness") in the context of MOOC lectures. The present study addresses following methodological question:

1. How can we define attention through a gaze-measure? At what levels can we define the attention or from a teacher's perspective the measure of "With-me-ness"?

We describe with-me-ness at two different levels: perceptual and conceptual. Apart from the methodological question, through this contribution we address following research question:

2. How is perceptual and conceptual levels of with-me-ness is related to performance?

With-me-ness

With-me-ness is defined at two levels: perceptual and conceptual. There are two pays a teacher may refer to an object: with deictic gestures, sometimes accompanied by words ("here", "this variable") or only by verbal references ("the counter", "the sum") Deictic references are implemented by using two cameras during MOOC recording, one that captures the teacher's face and one, above the writing surface, that captures the hand movements. In some MOOCs, the hand is not visible but teacher uses a digital pen whose traces on the display (underlining a word, circling an object, adding an arrow) act as a deictic gestures. Perceptual with-me-ness measures if the student looks at the items referred to by the teacher through deictic acts. Conceptual with-me-ness is defined by the discourse of the teacher: do students look at the object that the teacher is verbally referring to, i.e., .e. that the teacher is referring to a set of objects that are logically or semantically related to the idea he is referring to. Figure 1 shows the relative temporal granularities of the two levels of with-me-ness and different levels of perceptual with-me-ness.

The notion of with-me-ness is also comparable with measures of gaze coupling that were developed in studies involving dual eye-tracking. Cross-recurrence (Richardson et. al., 2007) reflects how much the gazes of two people follow each other during interaction. Recurrence is highest during references and recurrence level is related to the quality of interaction (Jermann & Nüssli, 2012).

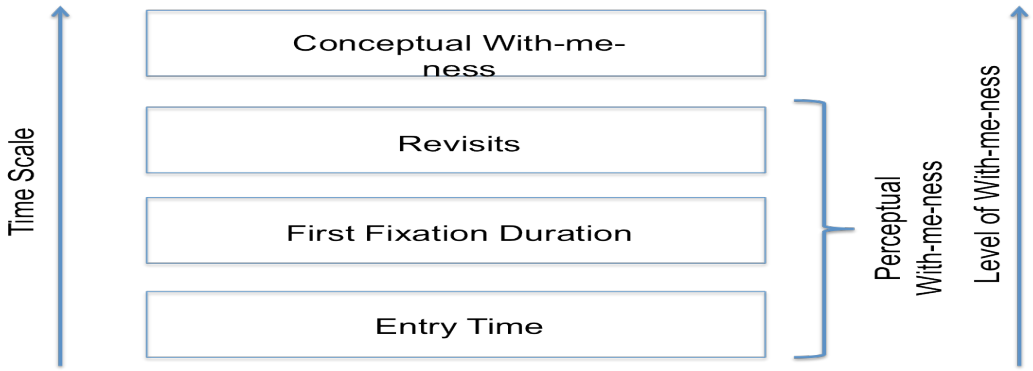

Figure 1: Temporal description of the two levels of with-me-ness and the sub-levels of perceptual with-me-ness.

Perceptual With-me-ness: The perceptual "With-me-ness" has 3 main components: entry time, first fixation duration and the number of revisits. Entry time is the temporal lag between the times a referring pointer appears on the screen and stops at the referred site (x,y) and the time student first time the student gaze stops at (x,y). First fixation duration is how long the student gaze stops at the referred site for the first time. Revisits are the number of times the student gaze comes back to the referred site.

Conceptual With-me-ness: The teacher may also refer. We measure how often a student looks at the objects verbally referred to by the teacher during the whole course of time (the complete video). In order to have a consistent measure of conceptual ``With-me-ness'' we normalize the time a student looks at the overlapping content by slide duration.

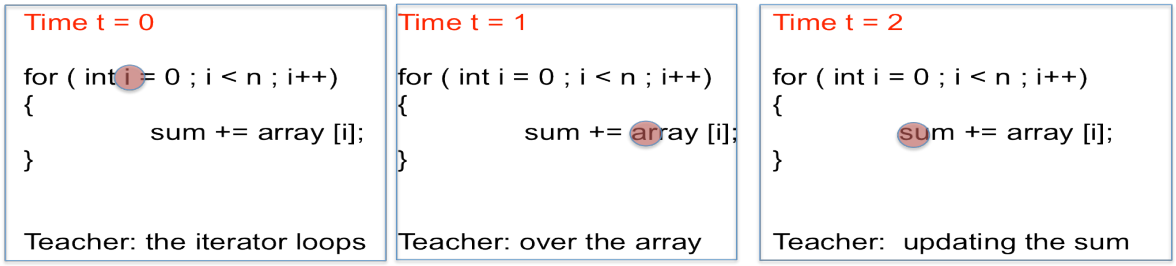

Figure 2: A typical example of conceptual with-me-ness over three consecutive time intervals. The semi-transparent red dot shows the gaze of a student while teacher is explaining the functionality of a program. We can see that the student is following the teacher's dialogues with his gaze.

Experiment

In the experiment, the participants watched two MOOC videos from the course "ANONYMUS" and answer programming questions after each video (particular to videos). The video was not made for this experiment but select from an existing popular MOOC. Participants' gaze was recorded, using SMI RED 250 eye-trackers, while they were watching the videos. Participants were not given controls for the playback for two reasons. First, the eye-tracking stimulus for every participant was the same, which facilitates the comparison participants. Second, the "time on task" remains the same for each participant.

40 university students from École Polytechnique Fédérale de Lausanne, Switzerland participated in the experiment. The only criterion of selecting the participant was the fact that each participant took the Java course in the previous semester, since this was a pre-requisite to understand the selected video. Upon their arrival in the experiment site the participants signed a consent form, then they answered three self-report questionnaires for a 20-point study processes (Biggs et. al., 2001), 10-point openness scale and 10-point conscientiousness scale. Then they took a programming pretest in Java. Then, they watched two videos from the MOOC course and after each video they answered programming questions based on what they were taught in the videos. In the following subsections, we describe different variables related to the present analysis.

The videos were in English: most participants were not English native speakers but are used to be taught in English. The first video included 13 slides and was 11 minutes 52 seconds long. The second video included 8 slides and was 10 minutes 7 seconds long. The content of slides was usually displayed in 3 to 4 steps. This increase conceptual with me ness since, at step 3 for instance student could look at objects that appeared on the slides in steps 1 and 2, but not yet at the object that will appear in step 4 and 5. The content of the slides was mostly text: sentences, lines of code and mathematical formulas. The content was mostly in black and white. Globally, slides were very light, with an average of only 55 words per slide.

Expertise and performance levels are given according the participants' scores from the pretest and the posttest respectively.

Results

1. Controlled variables: We observe no significant relation between the three variables. There is no significant relation between pretest score and posttest score. There is no significant relation between pretest score and learning strategy: the student with good pre-requisites, are not necessarily deep learners. More surprisingly, there is no significant relation between learning strategy and posttest score: deep learners do not learn more. This may certainly feed the debate on deep versus surface learning. However, our goal was not to enter into this debate but simply to control that there is not another variable, pre-requisite level or learning strategy, which would interfere massively with our key variables, perceptual and conceptual with-me-ness.

2. Pretest score and with-me-ness: We did not observe any significant relation between pretest score and the two levels of with-me-ness.

3. Learning strategy and with-me-ness: We also did not observe any significant relation between learning strategy and the two levels of with-me-ness.

4. Posttest score and with-me-ness: We observed significant correlations for the two different levels of with-me-ness and the posttest score.

a. **Entry Time:** We observe no correlation between entry time and the posttest score (Spearman's correlation = 0.1, $p>0.5$, Figure 2(b)). This can be explained using the saliency of the teacher's pointer. When a moving object appears on the screen, it constitutes a salient visual feature to which gaze is always attracted. This attraction does not reflect a deeper cognitive process and this is probably why it is not predictive to learning.

b. **First Fixation Duration:** We observe a significant correlation between the posttest score and the time spent for the first time the student look at the referred site (Spearman's correlation = 0.35, $p<0.5$, Figure 3(a)). The students who scores high in the posttest were paying more attention to the teacher's pointers. This behavior is indicative of more attention during the moments of deictic references.

c. **Number of revisits:** We observe a significant correlation between the posttest score and the number of times the student look at the referred site (Spearman's correlation = 0.31, $p<0.5$, Figure 3(b)). The students who scores high in the posttest came back to the referred sites more often than the students who scored less in the posttest. Having more revisits also resulted in having more fixations and thus more aggregated fixation duration as well. The revisiting behavior is indicative of rereading. Moreover, having more overall fixation duration on the referred sites is indicative of more reading time.

d. **Conceptual with-me-ness:** We observe a significant correlation between the posttest score and the time spent by the student following teachers' dialogues on the content of the slide (Spearman's correlation = 0.36, $p<0.5$, Figure 2(a)). The students who scores high in the posttest were paying more attention to the teacher's dialogue. This behavior is indicative of more attention during the whole video lecture.

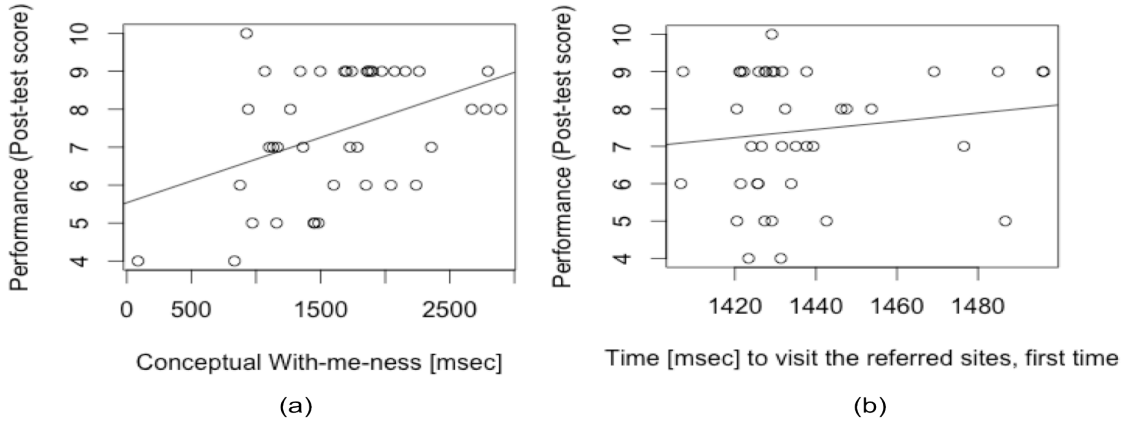

Figure 3: Posttest score and the With-me-ness. (a) Conceptual with-me-ness (x-axis) and posttest score (y-axis). (b) Entry time for perceptual with-me-ness (x-axis) and posttest score (y-axis).

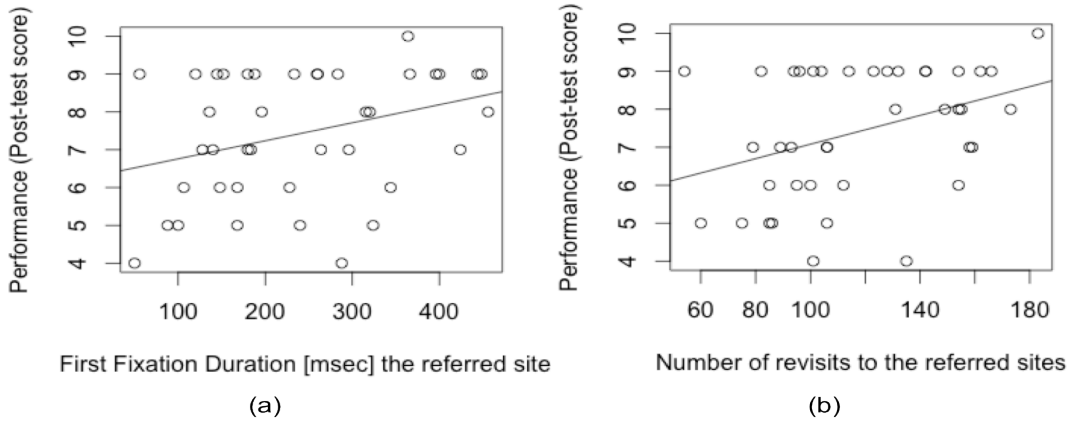

Figure 4: Posttest score and the With-me-ness. (a) First fixation duration for perceptual with-me-ness (x-axis) and posttest score (y-axis). (b) Revisits for perceptual with-me-ness (x-axis) and posttest score (y-axis).

Discussion

The entry-time component of the perceptual with-me-ness can be seen as the gaze behavior when there is a salient element present on the visual stimulus (Parkhurst et. al., 2002). The pointer of the teacher appears only a few times on the screen during the video lecture. We observe no correlation between the entry-time and the posttest scores. This can be explained by the fact that the pointer of teacher introduces a salient feature on the stimulus to which gaze is attracted. It does not reflect cognitive processing.

However, once the pointer is on the screen, the first fixation duration on the referred site is correlated with the posttest scores. The good-performers (those who scored high in the posttest) have more first fixation duration on the referred sites than the poor-performers. This is a typical situation during the moments of deictic references. Jermann and Nüssli (2012), in a pair-programming task, showed that better performing pairs have more concurrent gaze patterns during the moments of deictic references. Dale et. al. (2011), in listening comprehension task, showed that the pairs having more concurrent gaze during the period of references performed better than the other pairs.

The revisit component of the perceptual with-me-ness can be seen as rereading behavior. We observe a positive and significant correlation between the number of revisits to the referred sites and the posttest scores. The

participants scoring high in the posttest have higher number of revisits to the referred sites. Mills and King (2001) showed in their studies that rereading improves the comprehension. In the present study, the scenario is somewhat different than Mills and King (2001). In the present study, the students did not reread the study material. Instead, the students referred back to the previously seen content again in the duration the slide was visible to them. Thus the relation between rereading of the same content and the performance should be taken cautiously, clearly further experimentation is needed to reach a causal conclusion.

The conceptual with-me-ness corresponds to a deeper form of attention, in terms of both the temporal scale and the cognitive effort "to be with the teacher". We observe a positive and significant correlation between the conceptual with-me-ness and the posttest scores. The conceptual with-me-ness can be explained as a gaze-measure for the efforts of the student to sustain common ground within the teacher-student dyad. Dillenbourg and Traum (2006) and Richardson et. al. (2007) emphasized upon the importance of grounding gestures to sustain mutual understanding in collaborative problem solving scenarios. A video is not a dialogue; the learner has to build common grounds, asymmetrically, with the teacher. The correlation we observed between conceptual with-me-ness and the posttest score seems to support this hypothesis.

Conclusions

We found interesting relationships between gaze patterns and indicators for performance i.e., posttest scores. Those who achieved high scores in the posttest had more with-me-ness. In other words, from a teacher's perspective, students who scored high in the posttest were "with the teacher" for longer period of time than those who scored low in the posttest. They have more perceptual as well as conceptual with-me-ness than the poor-performers. However, the results reported are only correlations between variables; hence there is no causality claimed in the present contribution. Our conclusions is that have identified indicators that, on the one hand, can be captured by technology, and, on the other hand, related to learning performance.

The results also contribute towards our long-term goal of defining the student profiles based on their performance using the gaze data. The gaze-measure of students' attention can serve the purpose of a delayed feedback to the students based on their attention span. Perceptual with-me-ness can be used to give feedback to students about where they start lagging behind in the lecture. Moreover, revisits can be used to give feedback to students about their rereading behavior. The conceptual with-me-ness can be used to give feedback to the students about what they missed. Although, the results reported here are to taken cautiously and certainly more experimentation are needed to find any causality.

References

Astin, A. W. (1971). Predicting academic performance in college: Selectivity data for 2300 American colleges.

Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. British Journal of Educational Psychology, 71(1).

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in psychology*, *2*.

Dillenbourg, Pierre, and David Traum. "Sharing solutions: Persistence and grounding in multimodal collaborative problem solving." The Journal of the Learning Sciences 15.1 (2006.

Finn, J. D. (1989). Withdrawing from school. Review of educational research,59(2).

Grabe, M., & Latta, R. M. (1981). Cumulative achievement in a mastery instructional system: The impact of differences in resultant achievement motivation and persistence. *American Educational Research Journal*, *18*(1).

Holmes, R. M., Pellegrini, A. D., & Schmidt, S. L. (2006). The effects of different recess timing regimens on preschoolers' classroom attention. *Early Child Development and Care*, *176*(7).

Jermann, P., Nüssli, M. A., & Li, W. (2010). Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. In Proceedings of the 24th BCS Interaction Specialist Group Conference. British Computer Society.

Jermann, P., & Nüssli, M. A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM.

Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. Journal of Experimental Psychology: Human Perception and Performance, 17(1).

Kaplan, R. M. (1982). Nader's raid on the testing industry: Is it in the best interest of the consumer?. American Psychologist, 37(1).

Mayer, R.E., and Moreno, R., (2003) "Nine ways to reduce cognitive load in multimedia learning." Educational psychologist 38.1.

Millis, K. K., & King, A. (2001). Rereading strategically: The influences of comprehension ability and a prior reading on the memory for expository text. Reading Psychology, 22(1).

Paivio, A. (1991). "Dual coding theory: Retrospect and current status." Canadian Journal of Psychology/Revue canadienne de psychologie.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision research*, *42*(1).

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 361.

Schwartz, D.L., and Hartman, K. (2007). "It is not television anymore: Designing digital video for learning and assessment." Video research in the learning sciences.

Sharma, K., Jermann, P., Nüssli, M.-A., & Dillenbourg, P., (2011). Gaze Evidence of Different Activities of Program Understanding. In Procs. of the 24th Psychology of Programming Interest Group Conference.

Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2013). Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues. In Procs. of 10th International conference on Computer Supported Collaborative Learning.

Wiley, D. E. (1976). Another hour, another day: Quantity of schooling, a potent path for policy. *Schooling and achievement in American society*.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, *16*(1).