Looking AT versus Looking THROUGH: A Dual Eye-tracking Study in MOOC Context

Kshitij Sharma, École Polytechnique Fédérale de Lausanne, Switzerland, kshitij.sharma@epfl.ch
Daniela Caballero, Pontificia Universidad Católica de Chile, Santiago, dccaball@uc.cl
Himanshu Verma, École Polytechnique Fédérale de Lausanne, Switzerland, himanshu.verma@epfl.ch
Patrick Jermann, École Polytechnique Fédérale de Lausanne, Switzerland, patrick.jermann@epfl.ch
Pierre Dillenbourg, École Polytechnique Fédérale de Lausanne, Switzerland, pierre.dillenbourg@epfl.ch

Abstract: We present an eye-tracking study to show the different gaze patterns across the MOOC learners. We hypothesize that the pretests can be used to shape the attention of the students. Moreover we also proposed a collaborative add on activity to the learning material for the students to help them reflect on the material they learnt in the MOOC video. What comes out of the present study is two different interaction styles that differentiate the good students from the poor students. The good students engage with the teacher/collaborating partner through the interface/display. While the poor students engage with the material only. We name these two interaction styles as “looking through” and “looking at” respectively.

Keywords: Eye-tracking, massive open online courses, student engagement

Introduction

In the last two years millions of students worldwide have signed up for Massive Open Online Courses (MOOCs). The major issues for the MOOC researchers are: “how to develop efficient measures to capture the attention and engagement of the students?” and “how to make the learning process more efficient?” We address these two questions using a dual eye- tracking study based on a MOOC lecture and other add-on activities. Before the students attend to the MOOC lecture we use a pretest to prime them about the course content and after they have watched the video we ask them to collaboratively create a concept map based on what they learnt in the MOOC lecture.

We capture the attention and engagement of students during the video lecture and during the collaborative activity using eye tracking. In the present decade, off the shelf eye-trackers have readily become available. Soon the eye tracking will no more only be a sophisticated research tool.

In this article we present an empirical study that sheds some light on the gaze features of the MOOC learners and the effect of priming the students in two different ways. As we will see, the priming method impacts the learning gain of the students; and the gaze features we propose are efficient enough to highlight the differences between the good MOOC learners from the poor ones.

The two main contributions of the paper are as follows. First, we distinguish between the categories of learners based on how efficiently the students use the display to connect to the teacher or their peers. Second, we validate our gaze measures for quantifying how much the student follows the teacher and how together the collaborating pair works.

The rest of this paper is organized as follows. The second section presents the related work from collaborative eye-tracking research and from eye-tracking research in online learning. The third section presents the salient features and research questions from the present study. The fourth section explains the experiment and it’s variables. The fifth section presents the results. The sixth section discusses the implications of the results. Finally, the seventh section concludes the paper.

Related work

Eye tracking for online collaboration

In previous studies Jermann, et. al., (2010), Nüssli, et. al., (2009), have shown that the gaze is predictive of the expertise and/or the task performance. In a collaborative Tetris task, Jermann et al., (2010) showed that the experts focus more on the stack than the novices. In a collaborative Raven and Bongard puzzle solving task, Nüssli et al., (2009) showed that the good performers switch more often between the problem figures and the solution figures than the bad performers. Molinari, et. al., (2008) showed that the gaze is predictive of the mutual modeling of knowledge in a pair. The participants used a knowledge awareness tool, to assess their partners' knowledge, to manipulate their own actions on the concept map.

In a collaborative concept map task, Liu et al., (2009) used the gaze data to predict the expertise level of the pair. In a reference disambiguation task, Kraljic & Brennan, (2005) showed that the good performers spent less time on the ambiguous objects then the bad performers. In another collaborative problem solving task, Schneider (2013) showed that the gaze features are predictive of the collaboration quality (a rating scheme proposed by (Meier, Spada, & Rummel, 2007).

In a collaborative task the moments of joint attention are the most important. The moments of joint attention provide the basis of creating a shared understanding of the problem at hand. Making references is a key process to initiate a moment of joint attention. Jermann & Nüssli, (2012); Richardson & Dale, (2005) and Richardson, et. al., (2007) showed in the different studies how the moments of joint attention affect the gaze of the collaborating partners. The cross-recurrence (the probability of looking at the same thing at the same time) was observed to be higher during the referencing moments than rest of the interaction (Richardson & Dale, 2005; Richardson et al., 2007), Moreover, Jermann & Nüssli, (2012) showed that the pairs with high quality of interaction have higher cross-recurrence during the moments of joint attention.

Apart from moments of joint attention there are many other episodes of interaction during a collaborative problem-solving task. These episodes can be based on an underlying cognitive process (Aleven, et. al., 2012; Sharma, et. al., 2012) or dialogues (Gergle & Clark, 2011). In a pair program comprehension study, Sharma, et. al., (2012) showed that gaze patterns of the pair can differentiate between the episodes of linear reading and episodes of understanding the data flow of the program. In a collaborative learning task, Aleven et al. (2012) showed that the gaze patterns are indicative of the individual and collaborative learning strategies. In a pair programming task Sharma, et. al., (2012 and 2013); Jermann & Nüssli (2012) showed that certain dialogue episodes correspond to the higher gaze proportions at certain area on the screen. In a collaborative elicitation task, Gergle & Clark, (2011) showed that the movement of mobile partners can help them as a coordination mechanism.

Eye-tracking for online education

Use of eye-tracking in online education has provided the researchers with insights about the students' learning processes and outcomes. Scheiter, et. al., (2010) emphasizes on the usefulness of the eye tracking methods as analytical tools in online education and collaborative problem solving. Sharma, et. al. (2014 a, 2014 b) proposed gaze measures to predict the learning outcome in MOOCs. Sharma et. al. (2014a) uses the low level gaze features (derived from the stimulus) to predict the learning outcome; while Sharma et. al. (2014b) used the fact that how closely the students follow the teachers' deictic and verbal references to predict the learning outcomes. van Gog & Scheiter (2010) used eye-tracking to analyze multimedia learning process and instruction design. Scheiter, et. al. (2010) used eye-tracking data to differentiate between conceptual strategies in relation with different expertise levels in multimedia learning. Van Gog, et. al. (2005a) used eye-tracking data to differentiate expertise levels in different phases of an electrical circuit troubleshooting problem and concludes that experts focus more on the problematic area than the novices.

Mayer, (2010) summarized the major eye-tracking results on online learning with graphics and concluded that there was a strong relation between fixation durations and learning outcomes and visual signal guided students' visual attention. In a study to compare the affect of color coded learning material Ozcelik, et. al., (2009) found that the learning gain and the average fixation duration were higher for the students who received the color coded material than those who received the non color coded material.

Present study

We present a dual eye-tracking study where the participants attended a MOOC lecture individually and then a pair of participants collaborated to create the concept map about the learning material. We use the pretest to shape the understanding of the participants in a specific way (paying more attention to textual or schema elements in the video). This is called priming effect. One of the major hypothesis is that the there are two factors shaping the learning gain of the students: 1) how closely the students follow the teacher? 2) how well they collaborate in the concept map task? The first factor is important because the more a student follows the teacher, the more he could learn. The second factor is important because the better a student collaborates with the partner, the more the pair could discuss the learning material and have a better understanding. The present contribution explores the following research questions:

Question 1: How does the priming affect learning? We want to see if there is a priming effect on the learning gain of the participants .

Question 2: How are the individual gaze patterns during the video related to the collaborative gaze patterns during the collaborative concept map phase?

Question 3: How are the individual and collaborative gaze patterns related to learning gain?

Experiment and methods

Independent variables and conditions

We used a pretest as a contextual priming method. We designed two versions of the pretest. The first version had usual textual questions. The second version had exactly the same questions as in the first version but they were depicted as a schema (Figure 1 (a)).

Priming condition

Based on the two priming types we had two priming conditions for the individual video lecture task: 1) textual priming and 2) schema priming.

Pair composition

Based on the two priming types we had three pair compositions for the collaborative concept map task: 1) Both the participants received the textual pretest (TT); 2) Both the participants received the schema pretest (SS); 3) Both the participants received different pretests (ST).

Participants and procedure

98 students from École Polytechnique Fédérale de Lausanne, Switzerland participated in the present study. The participants were paid an equivalent of CHF 30 for their participation in the study. There were 49 participants in each of the priming condition (textual and schema). There were 16 pairs in each of TT and SS pair configurations while there were 17 pairs in ST pair configuration.

Upon their arrival in the laboratory, the participants signed a consent form. Then the participants took an individual pretest about the video content. Then the participants individually watched two videos about "resting membrane potential". Then they created a collaborative concept map using IHMC CMap tools[1]. Finally, they took an individual posttest. The videos[2,3] were taken from "Khan Academy". The total length of the videos was 17 minutes and 5 seconds. The participants came to the laboratory in pairs. While watching the videos, the participants had full control over the video player. The participants had no time constraint during the video phase. The collaborative concept map phase was 10 minutes long. During the collaborative concept map phase the participants could talk to each other while their screens were synchronized, i.e., the participants in a pair were able to see what their partners' action. Both the pretest and the posttest were multiple-choice questions where the participants had to indicate whether a given statement was either true or false. The gaze was recorded using the SMI RED 250 eye-trackers.

Dependent variable: Learning gain

The learning gain was calculated simply as the difference between the individual pretest and posttest scores. The miming and maximum for each test were 0 and 10, respectively.

Process variables

With-me-ness during the video lecture

With-me-ness (Sharma et. al., 2014b) is a gaze measure for quantifying students' attention during the video lectures. With-me-ness has two components: 1) perceptual with-me-ness and 2) conceptual with-me-ness. The perceptual with-me-ness captures the students' attention especially during the moments when the teacher makes explicit deictic gestures. Whereas, the conceptual with-me-ness captures whether and how much the gaze of the student is following the teacher's dialogues. To compute conceptual with-me-ness, two authors mapped the teachers' dialogues to the different objects on the screen. We name these objects as objects of interest (figure 1 (b)). Once we have the objects of interest on the screen, we computed what proportion of gaze time to the dialogue length (+2 seconds) in time is spent by the participants on the objects of interest. This proportion is the measure of the conceptual with-me-ness.

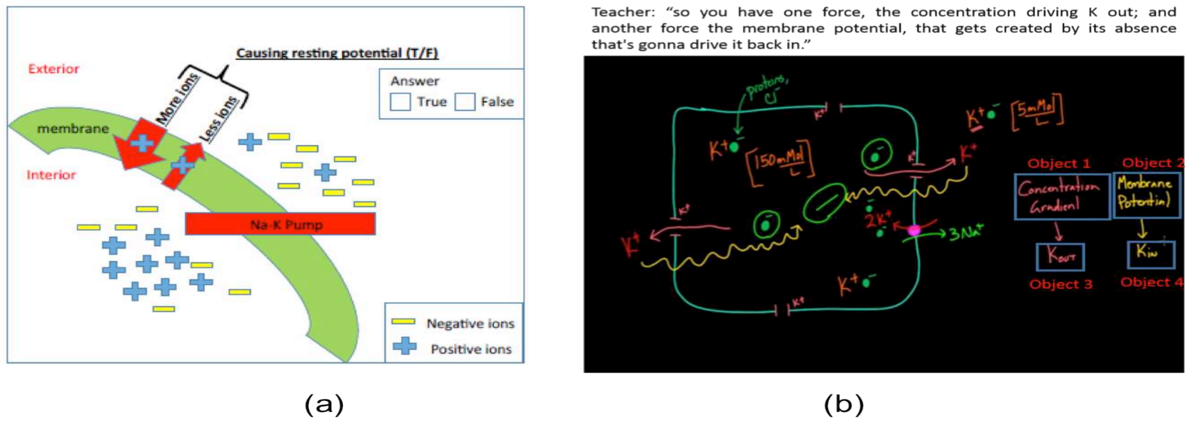

Figure 1. (a) Example question from the schema version of the pretest. The corresponding question in the textual version of pretest was "The original cause of the resting potential is the fact that the amount of the positive ions which diffuse to the interior is slightly more than the amount of the positive ions which diffuse to the exterior. (TRUE/FALSE)". (b) Example of objects of interest from the video lecture used in the experimental task. The example shows the four objects, related to the teacher's dialogue, on the screen.

Gaze similarity during collaborative concept map

The gaze similarity is the measure of how much the two participants in a pair were looking at the same thing at the same time (figure 2) or how similar their gaze patterns were during a short period of time. To compute the gaze similarity the whole interaction (during the collaborative concept map task) is divided into equal duration time windows. For each time window we compute a proportion vector, for each participant, containing the proportion of the window duration spent on each object of interest on the screen. Finally, the gaze similarity is computed as the scalar product of the proportion vector for the two participants in a pair. The gaze similarity is a similar measure as the cross-recurrence proposed by Richardson & Dale, (2005) but it is easier and faster to compute.

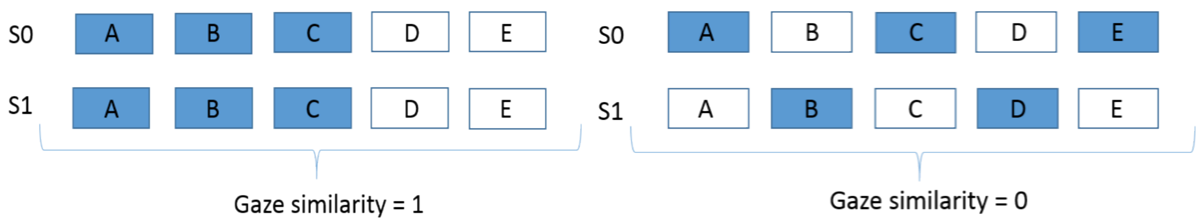

Figure 2. Typical cases in calculation of gaze similarity the filled rectangle denotes that the there was some time spent on the labeled object and the empty rectangle shows that there was no gaze on that particular object. If there are 5 objects (A, B, C, D, E) on the screen and the two participants (S0 and S1) are looking at the objects. The case with gaze similarity 1 shows that in a given time window both the participants spent equal amount of time on the respective objects. The case with gaze similarity 0 shows that in a given time window both the participants were looking at completely different sets of objects.

Results

The results show the relation between the priming, the two levels of with-me-ness during video phase, the gaze similarity and the learning gain. A fact that is worth mentioning at this point is that the time on task during the video lecture phase was varying across the participants. However, we do not observe a significant relation between the time spent on the video and the learning gain. For this contribution we are only focusing on the relation between the gaze patterns of the students during the video lecture and the collaborative concept map.

Priming effect on the learning gain

We observe a significant difference in the learning gain between the two priming conditions (figure 3a). The learning gain for the participants in the textual priming condition is significantly higher than the learning gain for the participants in the schema priming condition ($F [1, 96] = 16.77$, $p < .01$).

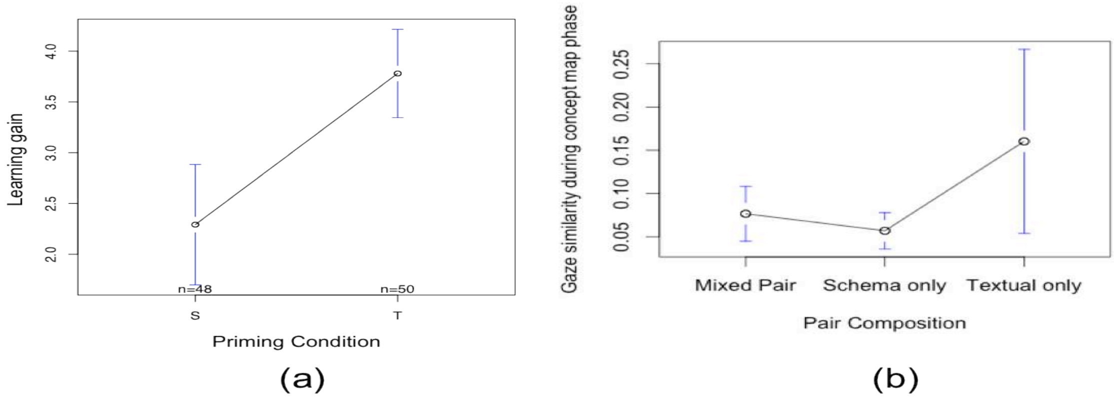

Figure 3. (a) Learning gains for the two priming conditions. (b) Gaze similarity for the three pair compositions

Pair composition and gaze similarity

We observe a significant difference in the gaze similarity be- tween the three pair configurations (figure 3b). The learning gain for the TT pairs is significantly higher than the gaze similarity for ST or SS pairs ($F [2, 46] = 3.44$,

With-me-ness and gaze similarity

Perceptual with-me-ness

We observe a significant positive correlation between the gaze similarity and the average perceptual with-me-ness of the pair ($r(47) = 0.48$, $p < .01$). The pairs having high gaze similarity have high average perceptual with-me-ness.

Conceptual with-me-ness versus gaze similarity

We observe a significant positive correlation between the gaze similarity and the average conceptual with-me-ness of the pair ($r(47) = 0.34$, $p < .05$). The pairs having high gaze similarity have high average conceptual with-me-ness.

With-me-ness and learning gain

Perceptual with-me-ness

We observe a significant positive correlation between the perceptual with-me-ness and the learning gain ($r(96) = 0.51$, $p < .01$). This difference is irrespective of the priming condition. The participants having high perceptual with-me- ness, irrespective of their priming type, have high learning gain.

Conceptual with-me-ness

We observe a significant positive correlation between the conceptual with-me-ness and the learning gain ($r(96) = 0.41$, $p < .01$). This difference is irrespective of the priming condition. The participants having high conceptual with-me- ness, irrespective of their priming type, have high learning gain.

Gaze similarity and learning gain

We observe a significant positive correlation between the gaze similarity between the pair and the learning gain ($r(96) = 0.39$, $p < .05$). The pairs having high gaze similarity have high learning gain.

Discussion

The research questions we addressed through the present study were about the relationships among the priming, the learning gain, with-me-ness during the video and the pair's gaze similarity during the collaborative concept map task. In this section we present the possible interpretation of the results presented in the previous section.

The first question concerns the effectiveness of priming on the learning gain of the participants. The learning gain of the participants in textual priming condition is significantly higher than that for the participants in the schema priming condition.

Moreover, during the collaborative concept map task, the pairs with both the participants from the textual priming condition have higher gaze similarity than the pairs in other two configurations (both from schema priming condition and both from different priming conditions). Once again, we can expect a better priming effect in textual priming condition than in the schema priming condition. The second question we address is about the relationship between the gaze patterns of the participants during the video watching phase and during the collaborative concept map phase. The pairs having high average with-me-ness also have a high gaze similarity. This is explained in terms of sharing a strong basis for shared understanding of the topic. Sharma et. al. (2014b) also found that with-me-ness is correlated with the learning outcomes. If both the participants followed the lecture in an efficient manner, i.e., with high with-me-ness, the pair has a strong base to build a shared understanding. Hence, the pair has more gaze similarity. This result is also consistent with the related research by (Richardson & Dale, 2005 and Richardson, et. al., 2007) where the gaze cross- recurrence is higher when the participants had a better level of mutual understanding.

The first part of the third question address the effect of the with-me-ness on the learning gain. Both the levels of with-me-ness are positively correlated with the learning gain, which is consistent with the results found by Sharma et. al. (2014b). The only difference is that in the present study we observe higher values for both the perceptual and conceptual with-me-ness than what Sharma et. al. (2014b) found in their experiments. The different levels of with-me-ness values are explained by the different types of the video lecture. The video used by Sharma et. al. (2014b) had only textual slides. The video in the present experiment had no slides; the teacher starts with a blank board and incrementally fills the board by writing the lecture material (schemas, tables, formulas). The low values of the correlation in the experiment can be explained by the nature of the videos. In the video from Sharma et. al. (2014b), the content is present on the screen from the beginning of slide resulting in the distraction as students might start reading from the slides and do not listen to the teacher. On the other hand, the video content in the video of the present experiment itself follows the flow of teacher's discourse hence resulting in higher values of with-me-ness for every student.

Finally, the second part of the third question inquires about the effect of the gaze similarity during the collaborative concept map task on the learning gain. The pairs with high gaze similarity also have high average learning gain. This can be explained using the fact that the high gaze similarity indicates a good shared understanding on the concerned topic. Hence, a similar pair (in terms of gaze) discusses the lecture points in a better manner than the pair with low gaze similarity. More specifically, the pair with high gaze similarity works on the same part of the concept map in a given time window, hence they develop a better mutual understanding about the concerned topic. Whereas, the pair with low gaze similarity work on less similar parts of the concept map and hence they fail to have a shared understanding.

Conclusions

We have studies the gaze properties of a learner during a MOOC lecture and during a collaborative concept map task. The good learner follows the teacher in both the perceptual and conceptual spaces of teacher-student interaction. More- over, a good learner is also well synchronized with his/her partner during a collaborative task. We also explored the effect of priming on the learning gain and the gaze patterns of the students. We observe that the textual priming has some advantages over the schema priming. A plausible explanation for textual priming emerging out as the better way of priming, as far as learning gain is concerned, is that the students are more used to the text than the schema tests in their regular studies. However, the effects of the textual priming on the collaborative gaze patterns might be attributed to the fact that textual priming has better effects than the schema priming.

From the present study, what emerged as a working hypothesis for the future research is a concept of "looking through" versus "looking at": some learners look "at" the display, as we look at a magazine, while other students seem to look "through" the display, that is, to look at the teacher or their partner in interaction as if they were actually present there. The latter seems to gain deeper engagement and hence a better learning outcome. The students who look "at" the display lag in following either the teacher or their partners. Whereas, the students who look "through" the display, use the display not only to follow the teacher or their partner but they use the display to create a shared understanding. Having a shared understanding in tern increases the learning gain for such students.

The concepts of "looking through" and "looking at" could be seen as new interaction style categories. "Looking at" the interface/ display indicates that the person is engaged with the material only, which is presented to him/her. "Looking through" the interface/ display indicates that the person is en- gaged with the peer. The peer in the video phase is the teacher and in the collaborative concept map is the collaborating partner. The "looking through" interaction resembles the social colocation of the interacting peers. As an analogy, to high- light the difference between the two interaction styles, we can compare the interaction with the teacher/collaborating partner to watching a movie. “Looking at” can be compared with liking the movie; whereas, “looking through” can be compared with appreciating the director.

Finally, we also validate the gaze measure for following the teacher during the MOOC lecture. We observe that the with- me-ness is a gaze measure that can be calculated for any kind of lecture independent of the lecture type. We also observe the consistency of the relation between the with-me-ness and the learning outcomes across the present study and our previous work (Sharma et. al., 2014b). Moreover, we also validate the collaborative gaze measure “gaze similarity” is a simpler and yet equally efficient measure to another mostly used collaborative gaze measure cross-recurrence (Richardson & Dale, 2005 and Richardson, et. al., 2007).

References

Aleven, V., Rau, M., & Rummel, N. (2012). Planned use of eye movement data to explore complementary strengths of individual and collaborative learning. *Proceeding of the DUET 2012*.

Gergle, D., & Clark, A. T. (2011). See what i’m saying?: using Dyadic Mobile Eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 435–444). ACM.

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* (pp. 1125–1134). ACM.

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* (pp. 36–44). British Computer Society.

Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, *50*(2), 194–231.

Liu, Y., Hsueh, P.-Y., Lai, J., Sangin, M., Nussli, M.-A., & Dillenbourg, P. (2009). Who is the expert? Analyzing gaze data to predict expertise level in collaborative applications. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on* (pp. 898–901). IEEE.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*(2), 167–171.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, *2*(1), 63–86.

Molinari, G., Sangin, M., Nüssli, M.-A., & Dillenbourg, P. (2008). Effects of knowledge interdependence with the partner on visual and action transactivity in collaborative concept mapping. In *Proceedings of the 8th international conference on International conference for the learning sciences-Volume 2* (pp. 91–98). International Society of the Learning Sciences.

Nüssli, M.-A., Jermann, P., Sangin, M., & Dillenbourg, P. (2009). Collaboration and abstract representations: towards predictive models based on raw speech and eye-tracking data. In *Proceedings of the 9th international conference on Computer supported collaborative learning-Volume 1* (pp. 78–82). International Society of the Learning Sciences.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education*, *53*(2), 445–453.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers’ and listeners’ eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407–413.

Scheiter, K., & Van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology*, *23*(9), 1209–1214.

Bertrand Schneider, R. P. (2013). Using Network Analysis Techniques on Social Eyetracking Data to Predict Levels of Collaboration. *Proceeding of the DUET*.

Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2012). Gaze Evidence for different activities in program understanding. In 24th Annual conference of Psychology of Programming Interest Group.

Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2013). Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues. In Computer Supported Collaborative Learning (CSCL 2013).

Sharma, K., Jermann, P., & Dillenbourg, P. (2014 a). “With-me-ness”: A gaze-measure for students’ attention in MOOCs. In International Conference of Learning Sciences (ICLS 2014).

Sharma, K., Jermann, P., & Dillenbourg, P. (2014 b). How Students Learn using MOOCs: An Eye-tracking Insight. In EMOOCs 2014, the Second MOOC European Stakeholders Summit.
Van Gog, T., Paas, F., & Van Merriënboer, J. J. (2005 a). Uncovering expertise-related differences in troubleshooting performance: combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*(2), 205–221.
Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*(2), 95–99.

Acknowledgements

This research is supported by the Swiss National Science Foundation grants CR12I1_132996 and 206021_144975. We would also like to thank all the participants who took part in the experiment.

