RITPU • IJTHE

Shaping learners' attention in Massive Open Online Courses

Kshitij **Sharma**
Computer Human Interaction in Learning and Instruction, École Polytechnique Fédérale de Lausanne, SWITZERLAND

Daniela **Caballero**
Pontificia Universidad Católica, CHILE

Himanshu **Verma**
Computer Human Interaction in Learning and Instruction, École Polytechnique Fédérale de Lausanne, SWITZERLAND

Patrick **Jermann**
Center for Digital Education, École Polytechnique Fédérale de Lausanne, SWITZERLAND

Pierre **Dillenbourg**
Computer Human Interaction in Learning and Instruction, École Polytechnique Fédérale de Lausanne, SWITZERLAND

Research paper with empirical data

Abstract

We present an eye-tracking study in the context of Massive Open Online Course (MOOC) videos. We propose to use a pre-test as a way of priming students about the video content before they watch the video. In this study, we used two versions of the same pre-test: textual and schema based. The results show that priming has an effect on the learning gain as well as the gaze patterns of students as they watch the videos; the type of priming has an effect on whether students spend more time looking at textual or schema video elements.

Keywords

Eye tracking, Massive Open Online Courses, Student engagement, Priming

Résumé

Nous présentons une étude oculométrique du visionnement de vidéo par des apprenants dans le contexte des cours massifs en ligne. Les questions du prétest sont proposées en deux versions, l'une textuelle et l'autre schématique. Nous faisons l'hypothèse que le format du prétest influence la perception de la vidéo par les étudiants. Les résultats montrent que l'amorçage a un effet sur le regard des apprenants de même que sur les gains d'apprentissage. Le type d'amorçage a un effet sur le type d'éléments (textuel ou schématique) que les apprenants regardent le plus.

Mots-clés

Oculométrie, Cours Massifs En Ligne, Amorçage

Introduction

In the last several years, millions of students worldwide have signed up for Massive Open Online Courses (MOOCs). The major issues for MOOC researchers are: how to efficiently engage and maintain students' attention, and how to optimise the learning process. 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 the MOOC lecture, we use a pre-test to prime them on the course content. After they have watched the video we ask them to collaboratively in pairs to create a concept-map based on what they learned in the MOOC lecture.

We use the Activating Student Knowledge (Tormey & LeDuc, 2014) method to pre-test. We compare two versions of priming (textual and schema). We capture the students' attention and engagement during the video lecture and during the collaborative activity by using eye-tracking. Off-the-shelf eye-trackers have become readily available. Soon, eye-tracking will no longer be seen exclusively as a sophisticated research tool.

In this article, we present an empirical study that sheds light on the gaze patterns of MOOC learners, as well as the effect of priming the students in two different ways. We show that the priming method affects the learning gain of the students as well as their gaze patterns throughout the video lecture. The gaze features we propose are capable of highlighting differences between high and low achievers. This distinction could be helpful in designing automated feedback systems for helping low achievers.

The rest of this paper is organised as follows. The second section reviews literature on collaborative eye-tracking, and eye-tracking in online learning. The third section presents the salient features and research questions of the study. The fourth section explains the experiment and its variables. The fifth section presents the results. The sixth section discusses the implications of the results. Finally, the seventh section concludes the paper.

Literature Review

In this section we review pertinent research. First, we present findings on eye-tracking and online collaboration. Second, we present findings on eye-tracking in online learning environments.

Eye-tracking for online collaboration

Previous studies (Jermann, Nüssli, & Li, 2010; Nüssli, Jermann, Sangin, & Dillenbourg, 2009) have shown that gaze is predictive of expertise and task performance, both for learning tasks and otherwise. In a collaborative Tetris task[1], Jermann et al. (2010) demonstrated that experts focus more on the stack than novices do. In a collaborative Raven and Bongard puzzle-solving task, Nüssli et al. (2009) demonstrated that good performers switch more often than bad performers between problem figures and solution figures.

In a collaborative task, the moments of joint attention are the most important. The moments of joint attention provide the basis for creating a shared understanding of the problem at hand. Making references is a key process in establishing a moment of joint attention. References are deictic gestures (pointing in collocated collaboration; and selecting a part of the screen space in computer mediated remote collaboration) usually followed by a verbal explanation to what has been referred. Jermann & Nüssli (2012), Richardson & Dale (2005) and Richardson, Dale, & Kirkham (2007) showed in 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 during the rest of the interaction (Richardson & Dale, 2005; Richardson et al., 2007), Moreover, Jermann & Nüssli (2012) showed that the pairs with high-quality interaction have higher cross-recurrence during the moments of joint attention.

Other than moments of joint attention, there are many other episodes of interaction during a collaborative problem-solving task. These episodes can be based on an under-

1 Tetris is a computer game where "Tetriminos" or geometric shaped game pieces, composed of 4 square blocks each, falling from the top of the game screen had to be arranged as a stack at the bottom of the screen in order to create a 10 block horizontal line without gaps (source: http://en.wikipedia.org/wiki/Tetris).

lying cognitive process (Aleven, Rau, & Rummel, 2012; Sharma, Jermann, Nüssli & Dillenbourg, 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 episodes of linear reading and episodes of understanding the program's data flow. In a collaborative learning task, Aleven et al. (2012) showed that gaze patterns are indicative of individual and collaborative learning strategies. In a pair-programming task Sharma et al. (2012) and Sharma, Jermann, Nüssli & Dillenbourg (2013); Jermann & Nüssli (2012) demonstrated that certain dialogue episodes correspond to higher gaze proportions at certain areas on the computer screen. In a collaborative elicitation task on a mobile screen, Gergle & Clark (2011) showed that the movement of mobile partners can act as a coordination mechanism for their explicit deictic gestures.

Eye-tracking for online education

Use of eye-tracking in online education has provided researchers with insights into students' learning processes and outcomes. Scheiter, Gerjets, & van Gog (2010) emphasize the usefulness of eye-tracking methods as analytical tools in online education and collaborative problem-solving. With a few exceptions, eye-tracking studies are normally done using a computer screen. Sharma, Jermann, & Dillenbourg (2014a, 2014b) proposed gaze measures to predict learning outcome in MOOCs. Sharma et al. (2014b) used low level gaze features derived from stimulus on a computer screen to predict learning outcomes; while Sharma et al. (2014a) monitored how closely students followed the teachers' deictic and verbal references to predict learning outcomes. Van Gog & Scheiter (2010) used eye-tracking to analyse multimedia learning process and instruction design. Scheiter et al. (2010) used eye-tracking data to differentiate between conceptual strategies in relation to different expertise levels in multimedia learning. Van Gog, Paas, & van Merriënboer (2005) used eye-tracking data to differentiate expertise levels in different phases of troubleshooting an electrical circuit and concluded that experts focus more on the problematic area than the novices do.

Van Gog, Paas, van Merriënboer, & Witte (2005) used eye-tracking data to provide feedback to students about their actions while troubleshooting an electrical circuit and found that feedback improved learning outcomes. Van Gog, Jarodzka, Scheiter, Gerjets, & Paas (2009) found that displaying an expert's gaze during problem-solving guided novices to invest more mental effort than when no gaze is displayed. Amadieu, van Gog, Paas, Tricot, & Mariné (2009) used eye-tracking data to determine the effect of expertise, in a collaborative concept-map task, on cognitive load. The authors concluded that the average fixation duration was lower for experts, indicating a greater cognitive load on experts than on novices. In an experiment in which the participants had to learn a game, Alkan & Cagiltay (2007) found that good learners focussed more on the contraption areas of a game while they thought about possible solutions. Slykhuis, Wiebe, & Annetta (2005) found that students spend more time on complementary pictures in a presentation than they do on a decorative picture.

Research Questions and Methodology

We present a dual eye-tracking study in which the participants attended a MOOC lecture individually and then worked in pairs to create concept-maps of the learning material. We use a pre-test to shape the understanding of the participants in a specific way with more emphasis on either textual or schema elements in the video. This is called the priming effect. One of the important hypotheses of the experiment was that priming affects the learning process of students, particularly their gaze patterns. A parallel hypothesis is that there are two factors shaping the learning gain of students: 1) how closely the students follow the teacher and 2) how well they collaborate in the concept-map task. The first factor is important because the better a student follows the teacher, the more (s)he can learn. The second factor is important because a pair that works well together will spend more time discussing the learning material, and so will develop a better understanding of it. For this paper, we concentrate exclusively on analysing the video-watching phase of the experiment. Through this study we explore the following research questions:

Question 1: How does priming affect the learning outcome of students?

Question 2: How does priming affect the gaze patterns of students while watching the video?

Experiment

In this section, we present the details of the experiment. First, we present the procedure and tools used for the experiment. Second, we present a detailed description of the independent, dependent and process variables used in the study.

Participants and procedure

Ninety-eight masters students from École Polytechnique Fédérale de Lausanne in Switzerland participated in the present study. The participants comprised 20 females and 78 males. The participants were compensated with an equivalent of CHF 30 for their participation in the study. There were 49 participants in each of the priming conditions (textual and schema). For the collaborative concept-map task, we had three pair configurations: one in which both participants had textual priming (TT), another in which both participants had schema priming (SS), and another in which each participant had different priming (ST). There were 16 pairs in each of the TT and SS pair configurations while there were 17 pairs in the ST pair configuration.

Upon their arrival at the laboratory, the participants signed a consent form. Then the participants each took a pre-test about the video content. Each participant then watched two videos about "resting membrane potential". The participants then worked in pairs to create a concept-map using IHMC CMap tools. Finally, each participant took an individual post-test. The videos were taken from "Khan Academy[2]". The total length of the videos was 17 minutes and 5 seconds. One important point worth mentioning here is that the teacher neither appeared in the video, nor was (s)he present during the screening

The participants came to the laboratory in randomly assigned pairs. While watching the videos, the participants had full control over the video player. The participants had no time constraint during the video-watching phase. The collaborative concept-map phase was 10-12 minutes long. During the collaborative concept-map phase, the participants in each pair could not see each other, but could talk with one another. Moreover, their screens were synchronised, allowing each participant to see their partner's action. Both the pre-test and the post-test were composed of questions which participants had to answer by indicating whether a given statement was either true or false.

Independent variables and conditions

As mentioned previously, we wanted to observe the difference in gaze patterns associated with different modes of priming. We used a pre-test as a contextual priming method. We designed two versions of the pre-test. The first version had standard textual questions. The second version had exactly the same questions as the first version but they were depicted as a schema. Figure 1 shows one question from the schema-based pre-test. The corresponding question in the textual pre-test was: "State whether the following statement is true or false: The main cause for the creation of resting membrane potential is more positive ions move inside the membrane than outside of the membrane."

2 The video was about resting membrane potential from Khan Academy.
Sources: https://www.youtube.com/watch?v=PtKAeihnbv0 https://www.youtube.com/watch?v=eROhIFBGKuU

Priming

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

Dependent variable: learning gain

The learning gain was calculated as the difference between the individual pre-test and post-test scores. The minimum and maximum scores for each test were 0 and 10, respectively.

Process variables

Time on video

We measured the total time spent on the video lecture by each participant. As the participants were allowed to interact with the video in any manner they wanted, the time spent on the video-watching phase is an important variable in comparing the two priming conditions.

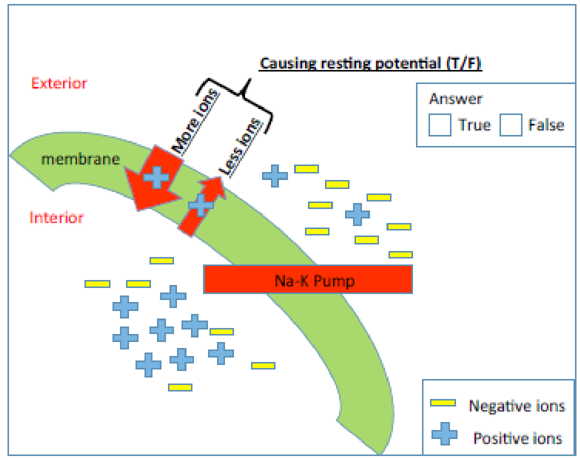

Figure 1: Example question from the schema version of the pre-test. The corresponding textual question was "State whether the following statement is true or false: The main cause for the creation of resting membrane potential is more positive ions move inside the membrane than outside of the membrane."

Teacher: "so you have one force, the concentration driving K out; and another force the membrane potential, that gets created by its absence that's gonna drive it back in."

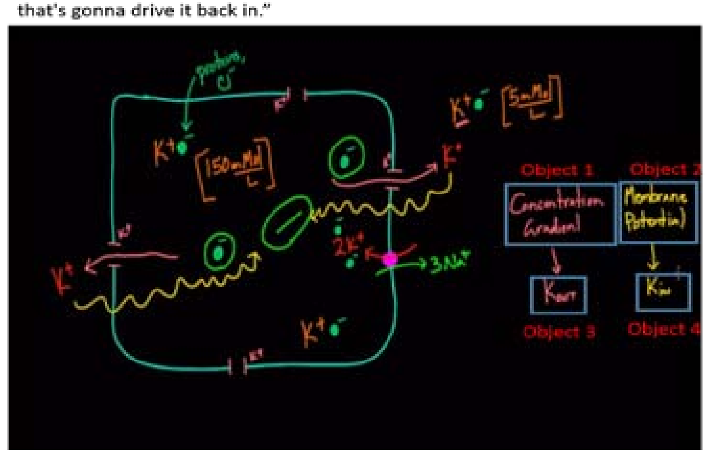

Figure 2: Example of areas of interest used in the experimental task. Objects 1 and 2 are textual elements, while object 3 and 4 are schema elements. The main schema in the middle of this snapshot was also divided into different schema elements like "ions", "membrane" and "channels".

Gaze on text

The video lecture had a mix of textual and schema elements. The teacher drew some figures and charts during the lecture and also made some tables and wrote some formulae. We categorised the tables, formulae and sentences written by the teacher as the textual elements of the video; the graphs, figures and charts were categorised as schema elements. For example, figure 2 is a snapshot of the video we used in the experiment. The objects on the screen were divided into schema or textual area of interest (AOI). We measured the time spent on the textual elements by each participant during the video lecture. This helped us to verify our hypothesis concerning the effect of priming on the gaze patterns of the participants.

Gaze compensation index

The proportion of time that the participants spent looking at the textual elements of the video did not correctly reflect the compensation in the gaze patterns, as the schema and textual elements do not appear in the same proportions on the screen throughout the video lecture. Initially, for a few minutes, the video contains only schema elements. Later the teacher keeps adding textual elements. This means that the proportions of schema and textual elements change over time. Hence, we need to take that change into account in order to calculate the real compensation effect. We propose a gaze compensation index calculated as follows:

$$Gaze\ Compensation\ Index = \sqrt{\sum \frac{\left(\frac{Gt}{Gs} - \frac{Pt}{Ps}\right)^2}{\frac{Pt}{Ps}}}$$

Where,

Gt := Gaze on textual elements in a given time window;

Gs := Gaze on schema elements in a given time window;

Pt := Percentage of screen covered with textual elements;

Ps := Percentage of screen covered with schema elements;

A gaze compensation index that is equal to zero reflects that a participant was spending the same proportion of time on textual and schema elements as were present on the computer screen. On the other hand, a higher gaze compensation index indicates a divergence between the proportion of time spent on the textual and schema elements compared to the proportions of screen space they cover.

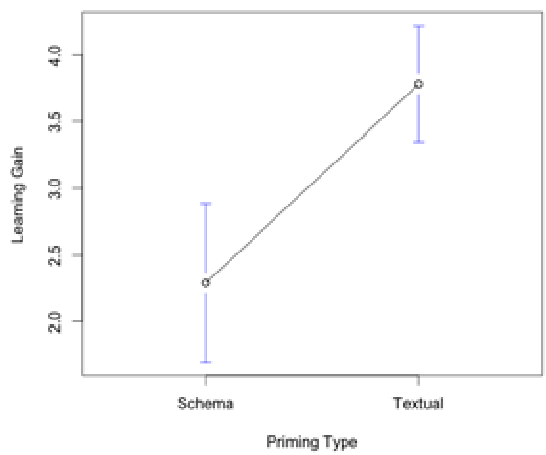

Figure 3: Learning gain for two priming conditions.

Results

In this section we present the relations between the priming conditions, the learning gain and the gaze variables.

Priming and learning gain

We observed a significant difference in learning gain between the two priming conditions (figure 3). 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 ($F [1, 96] = 16.77$, $p < .01$).

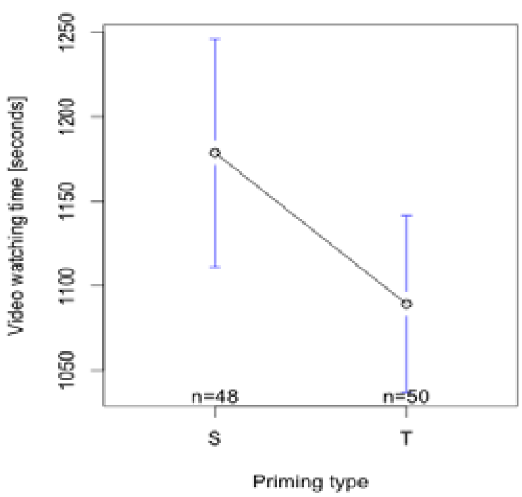

Figure 4: Video watching time (time on task) for the two priming conditions (S = schema priming, T = textual priming).

Priming and time on video

We observed a significant difference in the time spent on the video-watching phase between the two priming conditions (figure 4). The time spent on the video-watching phase for the participants in the textual priming condition is significantly lower than the learning gain for the participants in the schema priming condition ($F\ [1, 96] = 4.49$, $p < .05$).

Priming and time on text

We observed a significant difference in the time spent on textual elements in the video between the two priming conditions (figure 5). The time spent on the video-watching phase for the participants in the textual priming condition is significantly lower than the learning gain for the participants in the schema priming condition ($F\ [1, 96] = 4.91$, $p < .05$).

Priming and Gaze Compensation Index

We compared the gaze compensation index across the two priming conditions (figure 6). The participants in the textual priming condition had a higher compensation index than the participants in the schema priming condition ($F[1,96] = 56.198$, $p<.001$).

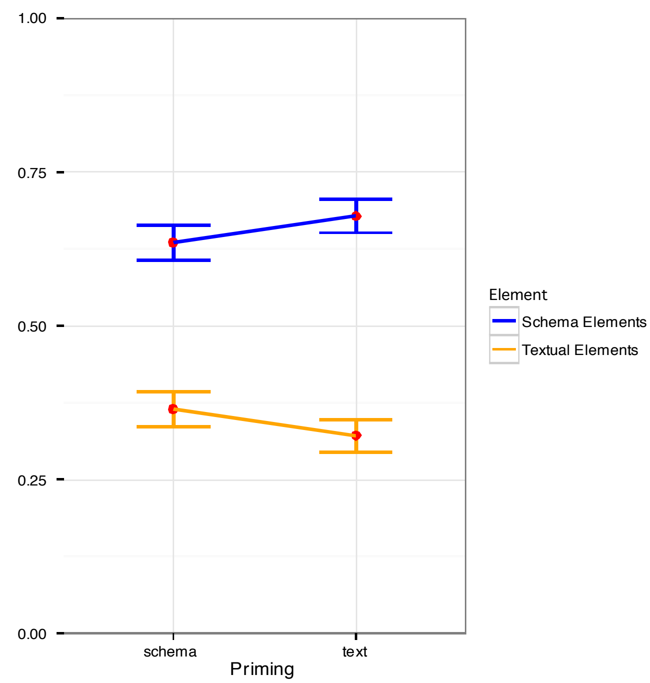

Figure 5: Gaze on different video elements for the two priming conditions.

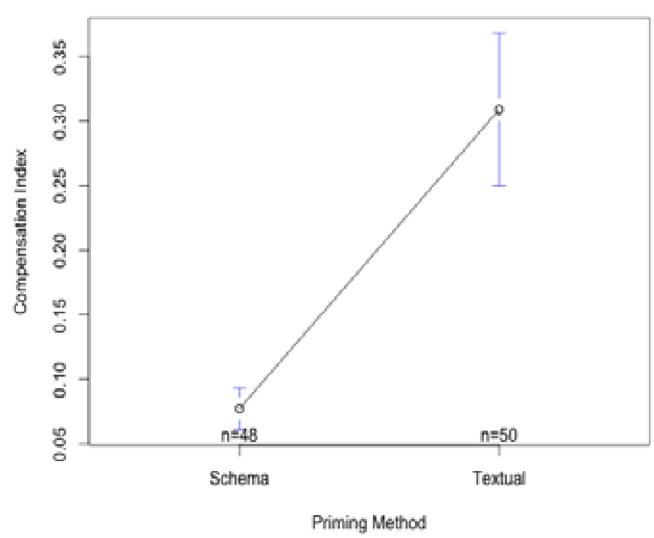

Figure 6: Gaze compensation index for the two priming conditions.

Discussion

The research questions we addressed in the present study were about the relationships between priming, learning gain, students' gaze patterns and time taken during the video-watching phase. In this section, we present the plausible explanations for the results presented in the previous section.

The first question concerns the effect of priming on the learning gain of the participants (figure 3). The learning gain of the participants in the textual priming condition group is significantly higher than that of the participants in the schema priming condition group. Our explanation for this effect is based on the theory of Tormey & LeDuc (2014) about Activating Student Knowledge (ASK) using priming methods. Tormey & LeDuc (2014) compared students' learning gain with and without priming in a history lecture. The priming method used in the study was a pre-test. We expand upon the concept by using two different versions of pre-test (textual and schema based). The textual method for ASK gave better results than the schema method. One plausible reason for the improved effect on learning gain is that the textual version provides more exact terms to look for in the lecture than that of the schema version of the pre-test. We will use this as the basis for future research.

The second question explores the effect of priming on individual gaze patterns during the video lecture. We found that the participants in the textual priming condition group looked more at the schema elements of the video, and that the participants in the schema priming condition looked more at the textual elements of the video (figure 5). This is a compensation effect of the priming. We also calculated the gaze compensation effect based on the ratio of textual and schema elements present on the screen and the ratio of the gaze on them respectively (figure 6). The participants in the schema priming condition undercompensate for the priming they received in the video-watching phase and so they missed some of the key concepts. One plausible reason for the participants in the schema priming conditions to miss the key concepts is that they missed the key connections between the schema and textual elements of the video, which could have had a detrimental effect on their learning gains.

The second question also considers the action patterns during the video-watching phase. We found that the students in the schema priming condition spent significantly more time on the video-watching phase than the students in textual priming condition (figure 4). Spending more time on the video-watching phase without high gaze compensation on the different elements in the video also affected the understanding of the students from the schema priming condition.

The experiment was designed in a MOOC context, and we found that the type of priming has an effect on the learning gain and the gaze patterns of the students during a video lecture. However, more experimentation is required to comment on the long-term effects of priming, and to generate design guidelines for MOOC teachers to shape the attention of their students using priming methods.

Conclusion

We presented a study on dual eye-tracking in a MOOC context. There are two salient features of the study. First, we indicated key knowledge points in the lecture beforehand using two different priming methods (schema and textual priming). Second, we showed that priming could help in shaping the attention of students during MOOC video lectures.

We proposed a new gaze measure as a gaze compensation index in order to compare gaze patterns during the video lecture. This measure tells us how much the students compensate for their activated knowledge through priming. We found that the students in the textual priming condition compensated more than the students from the schema priming. This had a detrimental effect on the learning gains of the MOOC students.

The results from the study are encouraging enough to continue research on prior knowledge activation of MOOC students, and having add-on activities (collaborative concept map). Moreover, a gaze-aware feedback system can also help students better compensate for their prior knowledge activation in a video lecture. The future work for the authors is to compare the results with those of a study in the same setting involving no priming and no add-on activity.

References

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

Alkan, S., & Cagiltay, K. (2007). Studying computer game learning experience through eye tracking. *British Journal of Educational Technology*, *38*(3), 538-542. doi: 10.1111/j.1467-8535.2007.00721.x Retrieved from http://www.researchgate.net

Amadieu, F., van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction*, *19*(5), 376-386. doi:10.1016/j.learninstruc.2009.02.005 Retrieved from http://www.researchgate.net

Gergle, D., & Clark, A. T. (2011). See what I'm saying?: using dyadic mobile eye-tracking to study collaborative reference. In P. Hinds, J. C. Tang, J. Wang, J. Bardram, & N. Ducheneaut (Eds.), *Proceedings of the ACM 2011 conference on Computer Supported Cooperative Work* (pp. 435-444). doi:10.1145/1958824.1958892 Retrieved from http://www.researchgate.net

Jermann, P., & Nüssli, M.-A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In S. Poltrock, C. Simone, J. Grudin, G. Mark, & J. Reidl (Eds.), *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1125-1134). doi:10.1145/2145204.2145371

Jermann, P., Nüssli, M.-A., & Li, W. (2010). Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. In T. McEwan, & L. M. McKinnon, *Proceedings of the 24th BCS Interaction Specialist Group Conference* (pp. 36-44). Retrieved from Electronic Workshops in Computing website: http://ewic.bcs.org

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 C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Proceedings of the 9th International conference on Computer Supported Collaborative Learning* (vol. 1, pp. 78-82). doi:10.3115/1600053.1600065 Retrieved from http://www.researchgate.net

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. doi:10.1207/s15516709cog0000_29

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. doi:10.1111/j.1467-9280.2007.01914.x Retrieved from http://www.researchgate.net

Sharma, K., Jermann, P., & Dillenbourg, P. (2014a, February). *How students learn using MOOCs: An eye-tracking insight*. Paper presented at EMOOCs 2014, the Second MOOC European Stakeholders Summit, Lausanne, Switzerland. Retrieved from Infoscience of the École Polytechnique Fédérale de Lausanne website: http://infoscience.epfl.ch

Sharma, K., Jermann, P., & Dillenbourg, P. (2014b, June). *"With-me-ness": A gaze-measure for students' attention in MOOCs*. Paper presented at the International Conference of Learning Sciences 2014, Boulder, CO. Retrieved from Infoscience of the École Polytechnique Fédérale de Lausanne website: http://infoscience.epfl.ch

Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2012, November). *Gaze evidence for different activities in program understanding*. Paper presented at the 24th Annual conference of Psychology of Programming Interest Group, London, UK. Retrieved from Infoscience of the École Polytechnique Fédérale de Lausanne website: http://infoscience.epfl.ch

Sharma, K., Jermann, P., Nüssli, M.-A., & Dillenbourg, P. (2013, June). *Understanding collaborative program comprehension: Interlacing gaze and dialogues*. Paper presented at the Computer Supported Collaborative Learning Conference, Madison, WI. Retrieved from Infoscience of the École Polytechnique Fédérale de Lausanne website: http://infoscience.epfl.ch

Slykhuis, D. A., Wiebe, E. N., & Annetta, L. A. (2005). Eye-tracking students' attention to PowerPoint photographs in a science education setting. *Journal of Science Education and Technology*, *14*(5-6), 509-520. doi:10.1007/s10956-005-0225-z

Tormey, R., & LeDuc, I. (2014, June). *The activating student knowledge (ASK) method in lectures*. Paper presented at Educational development in a changing world conference (ICED 2014), Stockholm, Sweden.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior*, *25*(3), 785-791. doi:10.1016/j.chb.2009.02.007

Van Gog, T., Paas, F., & van Merriënboer, J. J. (2005). Uncovering expertise-related differences in troubleshooting performance: combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*(2), 205-221. doi:10.1002/acp.1112 Retrieved from Miwa Lab website: http://miwalab.cog.human.nagoya-u.ac.jp

Van Gog, T., Paas, F., van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237. doi:10.1037/1076-898x.11.4.237

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*(2), 95-99.

Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146-154.