Using gaze tracking technology to study student visual attention during teacher's presentation on board

Enrique García Moreno-Esteva, Markku S. Hannula

Using gaze tracking technology to study student visual attention during teacher's presentation on board

Enrique García Moreno-Esteva and Markku S. Hannula

University of Helsinki, Department of Teacher Education, Helsinki, Finland, Enrique.garciamoreno-esteva@helsinki.fi

We present some initial findings obtained from a study on student attention in class using mobile gaze tracking technology. In this descriptive case study, we use a teacher's verbal, gaze, and gestural cues to identify the area on the board she wants her students to focus on, and analyse how one student's gaze location is shifting in relation to this. We found out that the student was actively following the cues during most of the time. However, we also observed moments when the student's gaze location was not in synchrony with the teacher's cues. The gaze tracking methodology seems to be a promising tool for a fine grained analysis of classroom communication and student attention.

Keywords: Mobile gaze tracking, student attention in mathematics classes, teacher gestures and indications.

INTRODUCTION

An important aspect of teaching is the multimodal communication (Radford, 2009; Arzarello, Paola, Robutti, & Sabena, 2009) between teachers and their students. Using mobile gaze tracking technology we will take a close look at the way a student's visual attention responds to cues from the teacher when the teacher is presenting new content on the board.

The teaching act is twofold. There is a pre-planned element of what the teacher does, which may be well planned and rehearsed several times over the years. There is also an improvisational element, reaction to unexpected student questions or comments, which requires the teacher to think 'on her feet'. Both of these modes of action include consciously chosen words, prosody, gestures, and facial expressions. However, some of the visible and audible messages are not consciously chosen, but are enacted more or less automatically, or even unconsciously.

Roth (2012) gives an example of a university professor whose dis-fluency he sees as an indication of the communicative act not being completed 'in mind' beforehand. Rather, it is "an unfolding event of communicating and thinking, which are not ready-made but develop in real time" (p. 237). Radford (2009) claims that "thinking [...] does not occur solely *in* the head but *in* and *through* language, body and tools" (p. 114).

The student side of classroom interaction is more complex than the simple receiving of messages from the teacher. Student behaviour in class is based on the student's personal agency, which is determined by his or her needs, goals, and identity. At the same time, student behaviour is largely reactive to changes in the environment, especially to what the teacher and the student's peers do. There is no research method that can provide a full account of meanings of this behaviour and reasons for it. Clinical interviews and think-aloud protocols distort the social interaction in class and thus lack ecological validity. Interviews done afterwards (including stimulated recall) can only have access to the student's post hoc reconstructions. Moreover, all self-reports are subject to be biased towards socially acceptable responses. Observations of facial expressions, brain imaging, and other physiological measures fail to capture the meanings students associate with their behaviour. Yet, each new methodology has shed light on some new aspects of the complexity of student cognition.

In this research report we shall present results from a pilot study using a mobile gaze tracking device to record students' visual attention during mathematics lessons. Gaze tracking is an established method for the study of attention at the automatic level of processing, in which the person is not consciously aware of fixations and shifts of attention. However, until recently, the technology had only been applicable in laboratory settings. Only now is it becoming available for use in ecologically valid situations such as outdoors (Baschnagel, 2013; Foulsham, Walker, & Kingstone, 2011) or in classes (Yang, Chang, Chien, Chien, & Tseng, 2013; Rosengrant, 2013).

THEORETICAL FRAMEWORK

Human communication consists of more than just words and diagrams; gestures, glances, body movement, prosody are also important aspects of it. Roth (2012) and Radford (2009) claim that there is a debate on the relationship between different communicative modalities. Some contend that both speech and gesture originate from the same psychological structure, while others claim that speech and gesture originate from different psychological structures. Still others claim that speech and gesture are different communicative channels, and that gesture serves a subordinate function.

Goldin-Meadow (1999) points out that the importance of non-verbal aspects of communication such as gestures has been recognized for a long time, at least two thousand years, in theatre, rhetoric, philosophy, and language. She identifies two different types of gesturing: gestures that substitute speech (e.g., sign language) and gestures that accompany speech, often unconsciously. She also points out that gesture enriches communication by providing a different representational format. For a speaker, gesture reduces cognitive burden, helps retrieval from memory, and is a tool for thinking. Gestures have also been observed to be important in the process of forming new concepts (Goldin-Meadow, 1999; Arzarello et al., 2009; Radford, 2009)

In gesture studies, McNeill's (1992) categorisation has been used frequently. He identified four different types of gestures: 1) beats, that do not have content information, but give rhythm and emphasis for talk, focus attention, and coordinate taking turns; 2) deictic gestures (pointing), that point to something concrete or abstract and typically have a verbal counterpart such as 'here', 'there', 'that', 'me', etc.; 3) iconic gestures that pictorially represent the target, for example, by drawing in the air; and 4) metaphoric gestures, that also create an image, but such that the image refers to an abstract concept metaphorically.

When teachers teach mathematical concepts, gesturing – and especially pointing – is common (Alibali & Nathan, 2012). Pointing gestures reflect the grounding of cognition in the physical environment, and pointing can be used to highlight connections between related representations. This exemplifies the need to interpret communication in context, which considers the material and graphical structure of the interaction (Goodwin, 2003; Arzarello et al., 2009).

Sweller and his colleagues have studied the influence of split attention on cognitive load (Yeung, Jin, & Sweller, 1997). Their studies show that when attention is split between two sources of information, a higher cognitive load may impede learning. This effect can be ameliorated when the sources of information are physically integrated. However, the effect is dependent on the student expertise. For more advanced learners additional information may be nonessential and impede learning.

Our research project will study student attention to mathematical diagrams and script on the board during mathematics lessons. Our focus in this paper is how well teacher talk and gestures direct student attention. In our study, we are specifically interested in student navigation when information is presented in two distinct areas, and in how effective is the teacher use of gesturing to help students integrate two sources of information. To our knowledge, no previous study has analysed students' visual attention in classroom situations using gaze tracking.

METHODOLOGY

In order for the researchers to monitor a student's attention in class, the student wears a gaze tracking device, which consists of a glass frame equipped with miniature cameras which produce a video scene and keep track of the direction that the eyes are pointing at. This device allows the software calculation of the direction of the gaze in class, producing a video scene with a dot indicating the locus of visual attention of the person wearing the glass frame (see Figure 3). The gaze tracking glasses are connected to a laptop with two cords, which prevents the student from getting up, but does not restrict movement while seated. The device was developed at the Finnish Institute of Occupational Health in collaboration with Aalto University (Lukander, Jagadeesan, Chi, & Müller, 2013). The prototype of this device is used in this study (see Figure 1). In addition to gaze tracking, there is another video recording of the class from behind, focusing on the teacher and the board, and there is a third video focussed on the subject student wearing the glass frame.

The data is obtained from a Finnish eighth grade classroom in a school where the language of instruction is English and where our subject student is a native English speaker (in this school, the students are bilingual). The subject wearing the glass frame was chosen on a voluntary basis. Informed consent has been obtained from the teacher, the focus student, all other students seen in any of the pictures, and their legal guardians.

The video recording we analyse shows a teacher explaining the topic of linear equations and relating

Figure 1: The gaze tracking device

them to the geometry of the lines using GeoGebra on a Smartboard. The dynamic functions of GeoGebra are not used, but the teacher draws additional lines and symbols on the Smartboard as shown in Figure 2. Specifically, throughout the approximately five-minute clip, a couple of lines appear on a Cartesian coordinate system. The lines intersect at a point (point P1 in Figure 2). Then the following sequence of events ensues. The teacher explains how to find the y-intercept of the first line (line L1, point P2, equation E2), and then, how to find the gradient of the same line by locating a second point (in this case P1) on the line and identifying the rise (segment S1) and its length (N1) and the run (S2) and its length (N2) from the lower point to the higher point. The gradient is calculated (E1). She then identifies the y-intercept and the gradient of the line in the equation of the line (in region Q). The same procedure is repeated for the second line (L2). The y-intercept (P3) is located and its value noted (E3), and a second point (P4) chosen in order to find the rise (S3) and its value (N3) and the run (S4) and its value (N4). The value of the gradient is calculated (E4). Finally, these values are identified in the equation corresponding to the line in region Q. In Figure 2, we identify all the areas of interest to which the teacher makes reference.

We made two segmentations of the video clip in the following way. First, we segmented the clip according to what we think are the areas of the board to which the teacher wants to bring attention, as indicated by the teacher's verbal cues, hand movements, and gaze. We then segmented the clip according to the student's

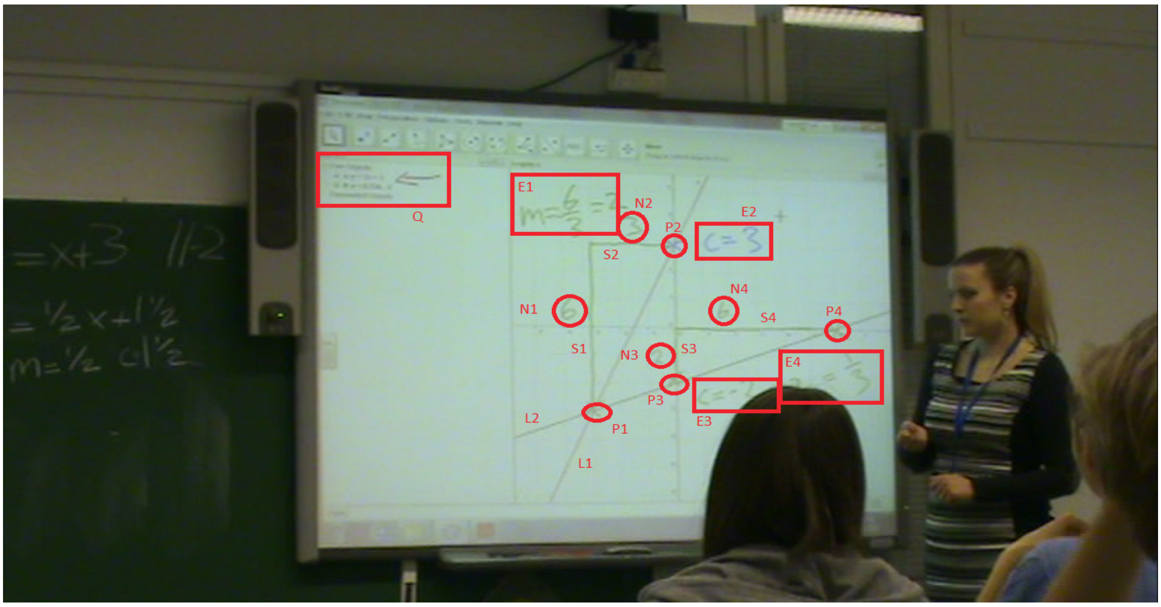

Figure 2: The board with the various regions of attention

gaze location – gaze is typically located for a short period of time at a certain position on the board before moving on to another location. Occasionally, there are glances which move away from a position and back to the original position, or somewhat rapid, but not too rapid shifts back and forth between two spots. We interpret these to indicate that the student attention is split between two areas and in these cases we coded both locations for the gaze. We then analysed the segments in order to find regions where gaze location follows the teacher's cues that direct attention distinctly, and regions where this is not the case.

DESCRIPTION AND ANALYSIS OF DATA

For most of the time during the teacher's presentation, the student gaze location follows teacher's cues quite well. However, there are also moments when this is not the case. We present a detailed analysis of two sections of the clip. Specifically, we shall analyse moments when the student's attention is not matching the intended area of attention. These sections occur after the teacher has finished explaining how to calculate the *y*-intercept and the gradient, and proceeds to link this information to the equations of the lines.

In the first instance (for line 1) the student follows the teacher's indications closely. In the second instance, the student gaze does not follow the pattern indicated by the teacher. The transcriptions follow. We have provided the graphic in Figure 4 to facilitate visualization of the ongoing processes. In the graphic we show how the area of intended attention (AIA) segments interlace in time with the gaze location (GL) segments for each of the two lines discussed by the teacher, line 1 (Line 1) and line 2 (Line 2). The bars representing each segment have been colour-coded according to the area of intended attention or the area where gaze is located: red is the code for the Q area, blue for the teacher's face or hands (T), white for unspecified locations, and other colours for other areas.

For the discussion pertaining to each line we provide the following information. First, we provide the beginning time of the segment and the transcript of what the teacher is saying. We then indicate if there are verbal, hand, or gaze cues that might be used to focus the student's attention onto an intended area of attention and what we perceive to be the intended area of attention. To the right of the transcript of the teacher intervention, we give the same information

Teacher			Student	
Time (ms)	Utterances [actions and non-verbal cues that indicate the area of attention]	Intended area of attention	Time of gaze shift (ms)	Gaze location
742006	and as you can see [glances briefly at equations in area Q]	Not specified	742014	Teacher face
742966	actually [takes another pen]	Not specified	743341	Q
743611	the equation of this line is here [draws an arrow pointing at the equation; gaze and gesture] (Figure 3c)	Q		Q Teacher face Teacher hands Q
748697	you can see that yes, [gaze and gesture]	P2, E2	 749137	P2, E2 Q
	the y-intercept [gaze and gesture]	E2, Q	750496 751959	Teacher face Q
752007	is the number that stands alone [gesture, glance]	Q	753940	Teacher face
754704	and the gradient [gaze and gesture]	E1	755372	E1
756229	is the number [gaze and gesture]	Q	756602	Q
757503	that is the coefficient of x [gesture only]	Q	757743 759268	Teacher face Q
760510	End of sequence		764774	

Table 1: Teacher behaviour and student gaze when the teacher explains the connection between gradient and y-intercept values and the equation for line 1

for the gaze location, except for a transcript. Notice that in some instances, there might be more than one intended area of attention, and in some instances the gaze location information might include two regions if there is a sequence of rapid shifts of gaze location between two spots. First we provide the data that concerns the discussion of line 1.

In the beginning of the segment we see the student focusing on the equation of the line before the teacher gives any explicit indication. Upon closer analysis, we see that when the teacher has finished the previous stage and before she picks a new pen, she casts a brief glance towards the equations. This was very difficult to observe and we noticed it only as we were carefully trying to find the reason why the student moved his gaze into the area where the teacher would move a fraction of a second later. We believe that the student observed the brief glance of the teacher as he was looking at the teacher's face, and therefore was able to react to the teacher's utterance, "as you can see" with a foresight as to where the teacher was likely to focus next.

Next, the teacher is trying to connect the explanation of the *y*-intercept and gradient to the equation of line 1, and the student's gaze location follows the teacher's indications to attend to area Q very distinctly (Figure 3c), and indeed to other areas as well, as can be seen from the top two bars in Figure 4 (in Figure 4, red codes the Q area, and blue codes the teacher's face or hands). What we see in this segment is how the student's gaze shift to the area of intended attention is in slight delay with respect to the teacher's cues. Also note how attention in this section is split between the equations of the line on the left (Q) and the values of *y*-intercept (E2) and gradient (E1) on the left. The teacher uses pointing gestures to successfully guide the student's attention across these two areas of interest.

Now we provide the data that pertains to the discussion of line 2. The student again anticipates the teacher movement to discuss the equations, but he does so substantially earlier (Figure 3a, first red bars in the upper and lower tiers of Line 2 in Figure 4). Moreover, he dismisses the teacher's direction to attend area Q at the end of the intervention (Figure 3b). We speculate as to whether this is partially due to the fact that she moves and turns around and is no longer facing the student; it would seem as if the student wants to see what he thinks the teacher is looking at. In any case, the teacher's pointing gestures are well visible in the student gaze video, but he does not direct his gaze as a reaction to the teacher's actions. This could be an example of self-initiated shifts of attention based on the student's active construction of his own understanding. Also, this might indicate a lack of attention, a blank stare.

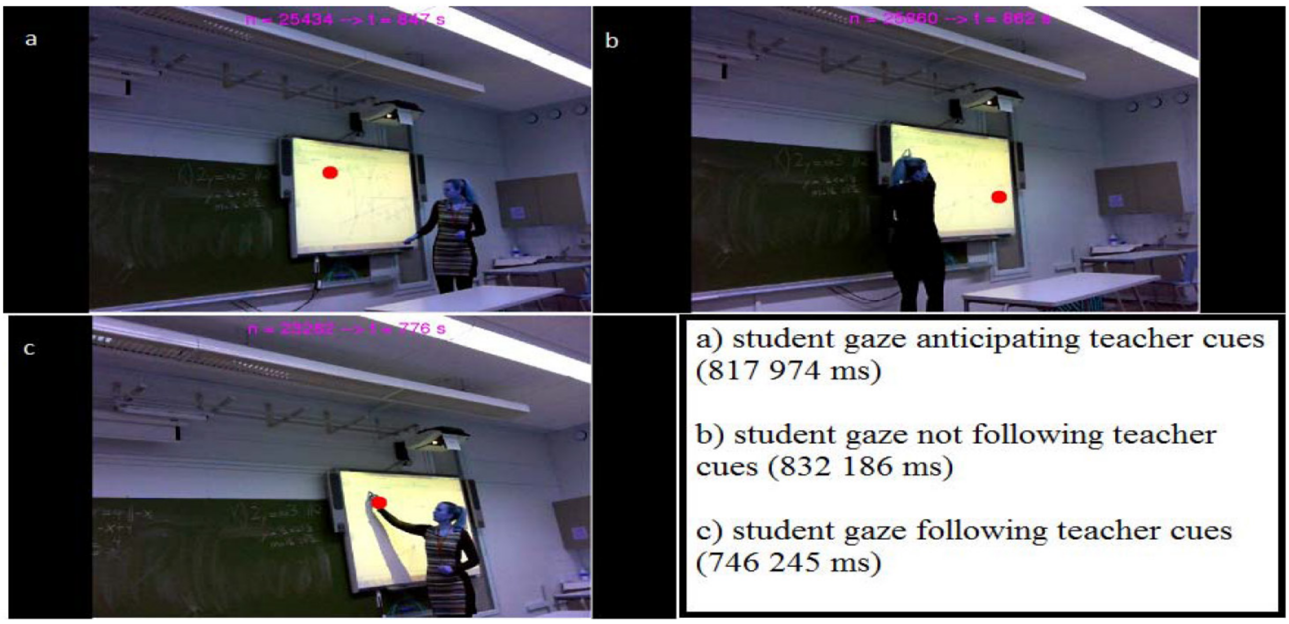

Figure 3: Pictures showing the situations described in the data analysis (Due to calibration error for this distance, the red dot indicating the student gaze location is systematically about 20 cm too much to the right)

Teacher			Student	
Time (ms)	Utterances [actions and non-verbal cues that indicate the area of attention]	Intended area of attention	Time of gaze shift (ms)	Gaze location
817115	pause, changes pen [no gaze, no gesture] (Figure 3a)	not specified	817105 818231	Q teacher hands
818733	this one uses approximate values [gaze, gesture]	Q	819104 819657	teacher face Q
821333	the correct answer there would be one over three (mumble)	Q	821621 822716	teacher face E4
826416	and the same thing again, can I find the [gaze, gesture]	E3		E4
829920	y-intercept [gaze, gesture]	P3	830243	P3
830920	here, the constant [gaze, gesture, teacher turns her back to the student] (Figure 3b)	Q	834352	Not Q, L2
835021	the eeh [gaze, gesture]	E3		Not Q, L2
837102	coefficient of x	Q	837078	S2, N2
838424	is the gradient	Q, E4		S2, N2
840141	End of section		840601	

Table 2: Teacher behaviour and student gaze when the teacher explains the connection between gradient and y-intercept values and the equation for line 2

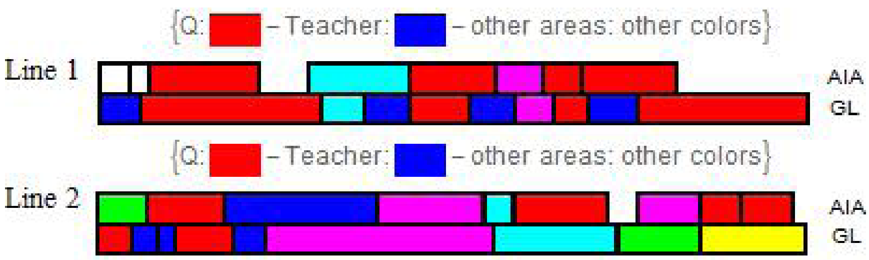

Figure 4: Segment interlacing for the transcripts above

DISCUSSION OF RESULTS

In previous research, the reliability of coding for gestures, speech, and their relationship has typically been high (85%-94%) (Goldin-Meadow, 1999). It seems likely that observing explicit gestures is very natural for human observers, which allows high accuracy. However, when we first watched the video, we did not notice the teacher's glances which, nevertheless, were used as cues by our focus student. This highlights the importance of paying close attention to the teacher's glances in future analysis on student visual attention.

The sections analysed are examples of a well preplanned teacher explanation. Yet, it includes unconscious cues (gaze cues such as glances). The student can follow the explicit cues very well, but is also observing subtle cues. Throughout the whole five minute sequence, the student splits his attention between the board and the teacher's face. In the sequences of closer analysis, there was additional split attention when the teacher connected inscriptions on two separate areas of the board. In both cases of split attention, the student seemed to have no problem following the teacher's cues. In summary, the student seems to be following the teacher's explicit gestural cues and subtle gaze cues quite closely.

However, when the teacher repeats the process for the second line, the student's gaze does not follow the teacher's cues. There are several hypotheses for why this happens. One option is that the student is processing the situation independently and is able to move ahead of the teacher, as indicated by his movement to the equations for the second line well before the teacher's. Another explanation is that the teacher's gaze is such an important cue for the student that when he loses it, his attention begins to drift. There is also the possibility that the student loses interest, due to the repetitive nature of the activity. Although the student's gaze continues to be on the board, it is possible that he is not attending to what he is looking at. We note that just observing the student does not provide information which we gather from gaze tracking data. We see from our data that when the teacher is explaining the first of two examples on the board, the student follows closely, but when she explains the second example, the student is looking towards the board, but is no longer following her explanations closely, lending credence to the usefulness of our technique.

REFERENCES

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: evidence from students' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247–286.

Arzarello, F., Paola, D., Robutti, O., & Sabena, C. (2009). Gestures as semiotic resources in the mathematics classroom. *Educational Studies in Mathematics, 70*, 97–109.

Baschnagel, J. S. (2013). Using mobile eye-tracking to assess attention to smoking cues in a naturalized environment. *Addictive Behaviors*, *38*, 2837–2840.

Foulsham T., Walker E., & Kingstone A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research, 51*, 1920–1931.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences, 3*, 419–429.

Goodwin, C. (2003). The semiotic body in its environment. In J. Coupland & R. Gwyn (Eds.), *Discourses of the body* (pp. 19–42). New York: Palgrave/Macmillan.

Lukander, K., Jagadeesan, S., Chi, H., & Müller, K. (2013). OMG!: a new robust, wearable, and affordable open source mobile gaze tracker. In M. Rohs et al. (Eds.), *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 408–411). Munich: MobileHCI.

McNeill, D. (1992). *Hand and mind. What gestures reveal about thought.* Chicago, IL: The University of Chicago Press.

Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics, 70,* 111–126.

Rosengrant, D. (2013). Using eye-trackers to study student attention in physical science classes. *Bulletin of the American Physical Society, 58*, 1–20.

Roth, W.-M. (2012).Tracking the origin of signs in mathematical activity: A material phenomenological approach. In M. Bockarova, M. Danesi, & R. Nuñez (Eds.), *Cognitive science and interdisciplinary approaches to mathematical cognition* (pp. 209–247). Munich: Lincom Europa.

Yang, F.-Y., Chang, C.-Y., Chien, W.-R., Chien, Y.-T., & Tseng, Y.-H. (2013). Tracking learners' visual attention during a multimedia presentation in a real classroom. *Computers and Education 62*, 208–220.

Yeung, A. S., Jin, P., & Sweller, J. (1997). Cognitive load and learner expertise: split attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology 23*(1), 1–21.