An Interactional Analysis of Gaze Coordination during Online Collaborative Problem Solving Activities

Murat Perit Çakır and Selin Deniz Uzunosmanoğlu
Informatics Institute, Middle East Technical University, Ankara, Turkey
perit@metu.edu.tr, sdeniz@metu.edu.tr

Abstract: Simultaneous tracking of eye gaze patterns of two or more students while they are engaged in a collaborative learning activity has recently attracted increasing interest in the learning sciences literature. The dual eye tracking paradigm allows researchers to quantify the degree of gaze overlap over the course of a learning activity, which is often treated as an estimate of the degree of joint attention among peers. In this paper we employ the dual eye tracking paradigm in the context of online collaborative geometry problem solving to quantify the degree of joint attention evidenced in gaze patterns, and explore what qualitative factors facilitate gaze coordination by focusing on the interactional organization of sequences in which gaze recurrence takes place.

Introduction

Investigating ocular correlates of learning with eye tracking methods has become an important focus of research in the learning sciences, particularly in the area of multi-media learning (van Gog & Scheiter, 2010). The majority of eye tracking studies in this field consider the individual as the primary unit of analysis. Such studies often control for the content and the pace of the instructional materials presented to the learners, measure the eye fixation patterns of the learners to identify what parts of the interface they allocate their attention, and finally relate these measures to learning outcomes to identify which presentation strategies were more effective. Since fixation sequences by themselves do not reveal what subjects are thinking or whether they are really paying attention, think-aloud protocols are frequently employed to interpret the eye-tracking data. However, since learning to see relevant visual structures at a scene and to associate them with appropriate linguistic terminology are primarily socially shaped processes (Goodwin, 1994), think-aloud alone may not be adequate to understand how students learn to allocate their visual attention to relevant objects in a particular domain.

The dual eye tracking paradigm, where the eye gaze of multiple subjects are monitored simultaneously while they are collaborating on a shared task, provides important opportunities to investigate the mechanisms underlying joint attention (Nüssli & Jermann, 2012). The *degree of overlap* or *cross-recurrence* among fixation sequences of interlocutors provides important information regarding to what extent the participants mutually orient to each other and to the objects in the shared scene (Richardson, Dale & Kirkham, 2007). A general finding in this emerging literature is that there is a positive correlation between the degree of gaze recurrence and collaborative learning outcomes (Jermann et al., 2011). In the context of well-defined tasks, it may even be possible to predict whether a dyad will succeed in solving a puzzle based on the patterns of synchronization detected in their raw gaze and speech data via machine learning algorithms (Nüssli, 2011). These findings also motivated the design of learning environments where users' gaze information is visualized on the screen in real-time as an awareness mechanism. A recent study on such a joint learning environment reported significant gains in gaze coordination and learning outcomes (Schneider & Pea, 2013).

Current studies that employ the dual eye tracking paradigm in the context of collaborative learning focus on devising quantitative metrics that reveal the type and the quality of collaboration, without particularly focusing on the role sequential organization of actions/utterances play in the achievement of joint attention. In this paper we focus on the relationship between changes in gaze recurrence and the sequential organization of interaction in a chat environment called Virtual Math Teams (VMT), where we asked dyads to collaboratively work on dynamic geometry problems. In this short paper we present our preliminary findings regarding (a) how measures of gaze recurrence relate to the overall success of collaboration in the context of collaborative math problem solving online, and (b) how variation in gaze recurrence relate to specific interactional events.

Methodology

This study was conducted with 18 volunteered university students at METU in Ankara, Turkey. Participants were grouped into 9 pairs and asked to collaboratively work on 10 geometry problems by using the VMT environment (Figure 1). Before the experiments participants received a short training that illustrated basic features of VMT. VMT is a CSCL environment that provides chat, shared whiteboard and wiki features to support discussions on math topics online (Stahl, 2009). Most recent version of VMT supports GeoGebra-based dynamic geometry constructions. The tasks were selected among activities presented in Stahl (2013), which included ruler and compass constructions such as constructing a perfect equilateral triangle, an isosceles triangle, and a perfect hexagon in GeoGebra. Participants coordinated their work and discussed strategies via the chat, and took turns to construct the desired dynamic geometry representations. Dyads were also asked to come up with an informal proof detailing why they thought that their construction was correct upon completion of each problem. Teams posted their proofs on a separate summary tab. During the online sessions eye movements of the participants were recorded with two desktop eye trackers (Tobii T1750 and Tobii T120) at a resolution of 50Hz.

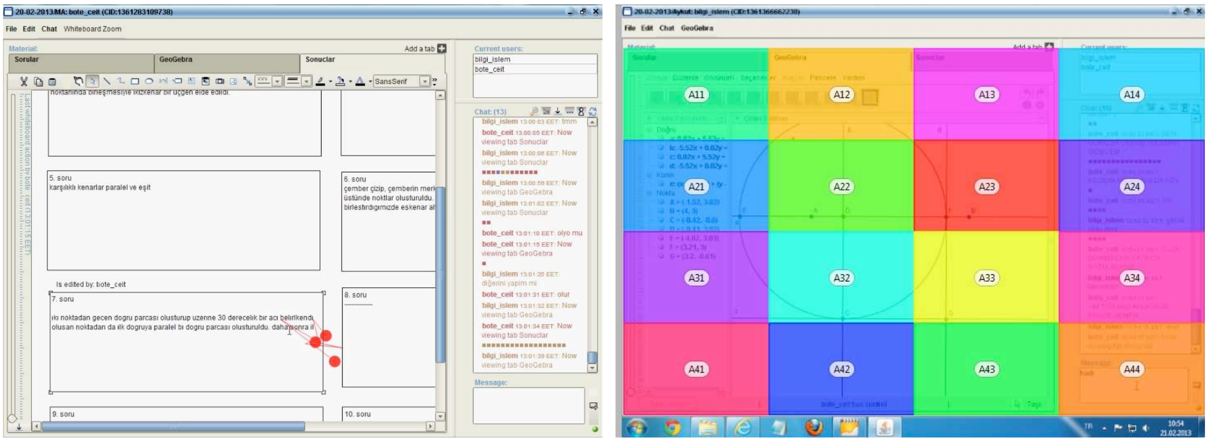

Figure 1. A screen shot of the VMT environment is presented on the left. The screen shot on the right shows the 16 areas used for partitioning the screens into areas of interest.

Eye-tracker data were analyzed both quantitatively and qualitatively. Gaze recordings were split into synchronized excerpts in the Tobii Studio Software. The screen recording of each participant was then split into a 4x4 matrix of non-overlapping areas of interests (AOIs) as shown in Figure 1. 12 of the AOIs covered the area where the dynamic geometry representations were constructed and 4 AOIs covered the chat component on the right. The time course of gaze duration falling over each AOI was then visualized by a custom software written in Java. The visualization provides a scarf plot that represents which AOI participants looked at, when their eye gazes overlapped, and over which AOI (Figure 4). The third row in the visualization indicates gray bars that highlight those time points where both participants' gaze fall on the same AOI. The whitespace indicate those instances where the participant is not looking at any of the AOIs (e.g. during typing).

Each of these 16 regions was considered as an approximation of the part of the screen over which the participants were attending to at any given time. While monitoring the shared scene users either moved their eye gaze with saccadic movements or fixated on specific locations by keeping their eyes still over a location. During a fixation event the fovea, which is the part of the retina that has the highest concentration of light sensitive cells, is oriented towards the fixated location. The fovea covers approximately 1-2 degrees of visual field, and at a distance of 65 cm from a screen, 1-2 degrees of visual field corresponds to a circular area with a diameter of 2.2 cm on the screen (Duchowski, 2007). The visual attention span is considered to cover a larger area covered by the foveal projection, as evidenced in dual-task experiments (Holmqvist et al., 2011). Since 17 inch displays with 4:3 aspect ratios were used during our experiments, the width and the length of the screen was 35 cm and 26 cm respectively. Splitting this area into 16 equal non-overlapping rectangular AOIs covers an area approximately 9 cm wide and 7 cm long. In this study this rectangle was considered as a rough approximation of where the person is attending to at any given time. Using the same AOI definitions on both screens over synchronized eye-tracking data allowed us to quantify gaze overlaps. Since the screens were divided equally, the probability that one of the participants allocate their attention on a given AOI is 1/16. Assuming independence of random gaze events, the possibility that two people allocate their attention on the same AOI is $1/16 \times 1/16 = 1/256$. So, the likelihood of having systematic gaze overlap among 2 people by chance is 0.004, despite the low spatial resolution provided by 16 AOIs.

In addition to the scarf plot, the software also computes a gaze recurrence plot for each problem attempted by every pair. The plots computed for each pair were then combined into a single recurrence percentage plot by taking the average of all corresponding data points coming from the plots for individual problem solving segments (see Figure 2). This yields a recurrence plot that summarizes the gaze patterns over all problem solving segments of the particular pair. In the combined summary plot, data points range from -4000 msec to +4000 msec with a 100 msec resolution. Point 0 indicates the recurrence percentage for synchronous gaze overlaps, -200 indicates the recurrence percentage in which B's gaze follows A's with a 200 msec delay with respect to A, and vice versa. The blue line shows the baseline recurrence level, which is computed over randomly shuffled gaze data. The vertical lines are the standard error bars, which indicate the amount of deviation in the data for the corresponding time.

Results and Discussion

In order to test the relationship between group performance and the degree of gaze overlap, 9 pairs were split into 3 achievement groups in terms of the number of problems they could solve correctly, whether they worked on tasks in a coordinated way and whether both participants contributed to the discussion. The degree of gaze overlap observed during each session was used as an indicator of the level of joint attention achieved by each group. Previous studies conducted with voice enabled computer-mediated communication systems found that participants took on average approximately 2 seconds to focus their attention on an object after it was mentioned by her partner (Richardson & Dale, 2005). In the present study the communication among partners is mediated by a chat and a shared drawing tool. One may expect that reading a chat utterance and then allocating one's attention to the referred object on the drawing board take more than two seconds. However, the gaze recurrence plots with various lag combinations indicated that the highest degree of gaze overlap occurs within a similar time interval. Therefore, those instances in which one subject looks at the same area of the screen that his partner looked at within two seconds were treated as instances of gaze overlap.

The high achievement group exhibited on average 31% gaze overlap, which is followed by the medium and low achievement groups with 24% and 13% gaze overlap respectively. A one-way ANOVA conducted over gaze overlap values indicated that this difference is statistically significant, $F(2,8) = 11.917$, $p<0.001$, $\eta^2 = 0.341$. Games-Howell post hoc tests found a significant difference between low and medium achievement groups ($MD=-12.32$, $p<0.05$), as well as between low and high achievement groups ($MD=-20.19$, $p<0.01$). The difference between medium and high achievement groups was not significant. In other words, higher achieving groups exhibited significantly higher gaze coordination during collaborative problem solving sessions.

Figure 3 shows the overall gaze recurrence plots corresponding to the entire session of two different dyads. The plot on the left of Figure 3 corresponds to the best performing dyad in our sample, whereas the plot on the right corresponds to the worst performing team. The better performing team's recurrence plot differ from the other team's plot in several ways. First, the mean percent recurrence values are much higher for the better team. Second, the better team's recurrence plot significantly deviates from the random baseline, whereas the other team's recurrence plot cannot be distinguished from the baseline. The recurrence plot for the better team has two peaks, one around -1500 msec, whereas the other at 1500 msec. This suggests that both partners equally followed each other's gaze, indicating a high degree of coordinated behavior. Such a pattern is not visible for the other team.

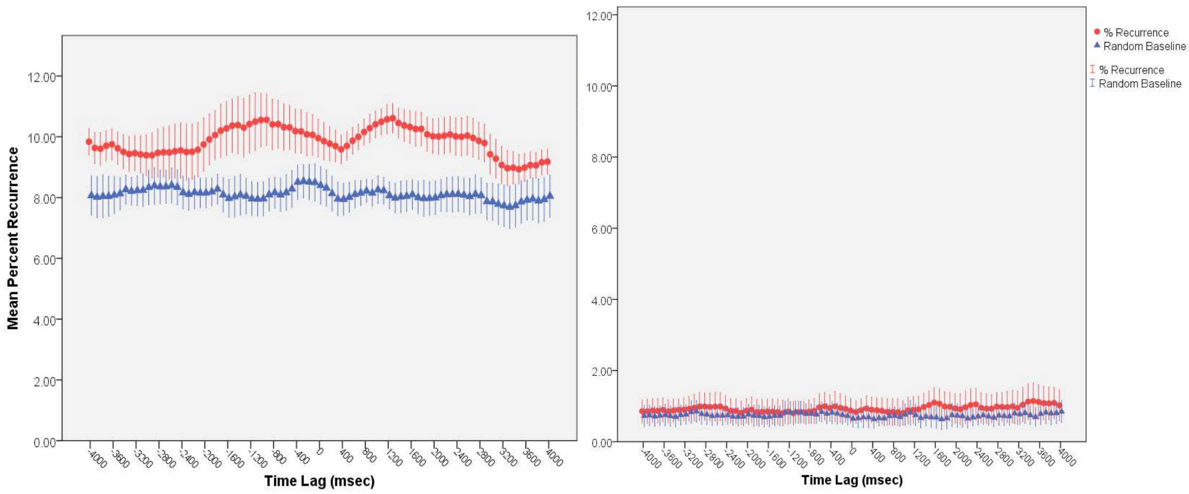

Figure 2. Overall gaze recurrence plots for the highest achieving (left) and the lowest achieving (right) dyads.

In an effort to explore some of the factors underlying these global gaze recurrence patterns, we conducted an interaction analysis of a sample excerpt obtained from both groups' chat sessions. We used the scarf plot in Figure 4 as a guide to index those moments in collaboration with greater degree of gaze overlap. The excerpt shown in table 1 shows the high achieving dyad's work on question 6, which asked them to construct a perfect hexagon by using points, circles and line segments only. In the first three lines, A orients his partner to the new problem. In line 4, after reading the question (as indicated by fixation patterns in the video recording) B solicits A's help for ideas. Next A raises a question which proposes a possible way to characterize a hexagon in terms of 6 equilateral triangles. B's responses in the next two lines suggest that he is not sure about the proposed relationship. Next A takes control of the drawing area and adds a circle. Then, he posts a chat message stating that he does not know how to proceed. He also mentions that he thought about using angles but that did not help him proceed any further.

While A was constructing the circle, B was looking at the unfolding construction on the GeoGebra tab. Since he was following these steps right after they were committed into the system and broadcasted to all clients, there is a slight lag between A's and B's gaze patterns. Once A seems to be done with his drawing, B makes a suggestion that A should consider drawing an equilateral triangle. As A is drawing the triangle, B looks at places where he expects the next line segment to appear. Such *anticipatory gaze patterns* highlight a stronger degree of common ground or indexical symmetry between the two partners as compared to a situation where one partner merely follows the drawing actions of the other partner.

Table 1. Excerpt from the high achieving dyad

#	User	Time	Chat Message / Drawing Action
1	A	23:43.2	6ya geçelim *(Lets move to 6)*
2	B	23:46.0	her şey silindi bende *(everything is erased)*
3	A	23:46.0	ben sildim 6ya bakalım :) *(I did that lets take a look at 6:))*
4	B	23:57.0	nasıl yaparız? *(how can we do it?)*
5	A	24:32.6	Şimdi altıgenin içinde 6 tane eşkenar üçgen var di mi *(now there are 6 equilateral triangles inside the hexagon, right?)*
6	B	24:46.3	6 mı *(6?)*
7	B	25:16.9	bilmem *(I don't know)*
8	A	25:19.7	dur bi başlayalım *(let's just start)*
9	A		***constructs a circle with center A and a point B***
10	A	26:58.1	bilemedim nasıl yaparız *(I don't know how we should do it)*
11	A	27:04.5	açı kullancaktım ama yok *(I would use angles, but there are none)*
12	B	27:10.0	düşünelim *(let's think about it)*
13	B	27:20.0	yukarda 3 kenar olcak aşağıda da *(There should be 3 edges at the top and also at the bottom)*
14	B	27:31.9	nasıl buluruz *(How can we find it)*
15	A		***adds point C on circle then draws the diameter connecting B and C (see Figure 3, left)***
16	B	29:29.8	eşkenar üçgen oluştursan *(how about you draw an equilateral triangle)*
17	A		***removes diameter, adds another circle sharing the radius with existing circle, draws an equilateral triangle by using the radius (see Figure 3, middle)***
18	B	30:46.4	aynısını diğer tarafa da yap *(do the same to the other side)*
19	A		***draws lıne segments DB and BC, adds a new circle passing through A and B, marks the intersection E (see Figure 3, right)***
20	B	31:37.2	aferim *(good job)*
21	B	31:41.0	tebrkler *(congrats)*

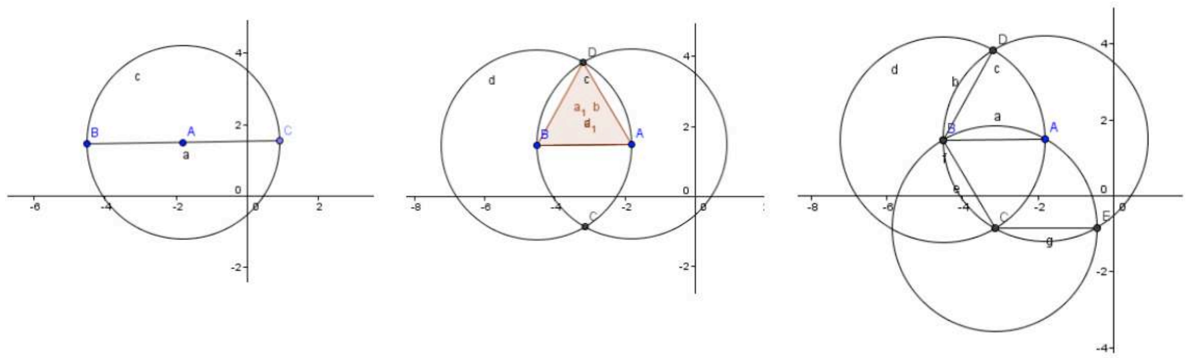

Figure 3. The evolution of A's drawing actions for question 6

Once A completes the triangle (see Figure 3, middle) B suggests him to draw the same triangle on the other side. This is followed by the addition of line segment BC, which is monitored by B. Next, A adds another circle centered at C and sharing a radius with the two other circles. When he defines point E at the intersection of the new circle and the circle on the top right, B recognizes this as the solution. At that moment his eye gaze traverses over the projected new intersections that will become the hexagon once connected by line segments.

Analysis of excerpts from the other team reveals a very different form of interactional organization. As part of their task description, each team was asked to summarize their findings on a separate tab called Summary. Starting with question 2, this team switched to a cooperative mode of operation, where one partner aimed to solve the next question on the GeoGebra tab, while the other partner was busy summarizing the previous solution on the Summary tab. This led to very little transactive interaction among the partners and hence very small degree of gaze overlap. The scarf plot in Figure 4 summarizes the gaze overlap distribution of this team while they were working on one of the questions. The whitespace in the scarf plot also reveals that the first user was busy typing the textboxes in which the team was summarizing their solution for the first question. Occasional gaze overlaps occurred at the chat component when participants exchanged occasional messages.

Overall, the preliminary results of the interaction analysis were consistent with the gaze recurrence analysis. Gaze recurrence analysis is successful for giving a global view of gaze patterns, such as if there is a significant amount of gaze cross recurrence and if there is an asymmetry among peers (e.g. one partner tended to follow the eye gaze of the other). For instance, the cross recurrence map for the successful dyad was symmetric with respect to the origin, meaning that both partners almost equally followed each other's gaze, which signals that there is a stronger level of coordination among their actions. As far as qualitative aspects of collaboration is concerned, the quality of collaboration increased together with the increase in gaze overlap if both participants actively contributed to the problem solving process by proposing ideas, constructing geometric diagrams and monitoring each other's actions.

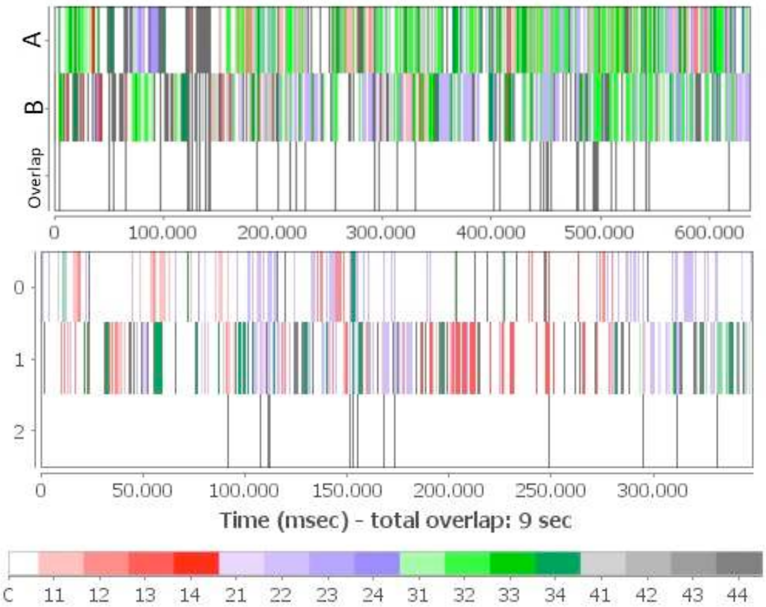

Figure 4. Scarf plot showing the distribution of eye gaze of A and B while they were working on the hexagon problem. The lower graph shows the scarf plot for the lower achieving group.

The degree of gaze overlap was closely related to the achievement of indexical symmetry among partners that renders indexical terms such as "the other side", "draw an equilateral" etc. meaningful to the participants. In the sequentially unfolding context of their interaction, such terms attain their specificity through the actions of the interlocutors. The presence of indexical symmetry seems to be best evidenced in the case of anticipatory gaze patterns, where one partner fixates on a location where he/she expects the next relevant action to happen. In future work, we are planning to do a more fine-grained analysis of the temporal course of gaze recurrence in an effort to distinguish anticipatory gaze patterns from simple forms of action following. Such features may lead to more fine grained ocular correlates of joint attention and transactive reasoning, which are important factors on the overall success of collaborative learning.

References

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. New York, NY: Springer.

Goodwin, C. (1994). Professional Vision. *American Anthropologist, 96*(3), 606-633.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York, NY: Oxford University Press.

Nüssli, M., & Jermann, P. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task *Proceedings of CSCW'12,* 1125-1134. New York, NY: ACM.

Nüssli, M. A. (2011). Dual eye-tracking methods for the study of remote collaborative problem-solving. Unpublished PhD thesis, École Polytechnique Fédérale de Lausanne.

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.

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *Int Journal of Computer-Supported Collaborative Learning*, *8*(4), 375-397.

Stahl, G. (2013). Topics in Dynamic Geometry for Virtual Math Teams. Retrieved from http://gerrystahl.net/elibrary/topics/topics.pdf

Stahl, G. (Ed.). (2009). *Studying virtual math teams*. New York, NY: Springer.

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