Toward Collaboration Sensing: Applying Network Analysis Techniques to Collaborative Eye-tracking Data

Bertrand Schneider[1,2], Sami Abu-El-Haija[2], Jim Reesman[2], Roy Pea[1]
Stanford University
Graduate School of Education[1], Computer Science department[2]
schneibe@stanford.edu, haija@stanford.edu, jreesman@cs.stanford.edu, roypea@stanford.edu

ABSTRACT
In this paper we describe preliminary applications of network analysis techniques to eye-tracking data. In a previous study, the first author conducted a collaborative learning experiment in which subjects had access (or not) to a gaze-awareness tool: their task was to learn from neuroscience diagrams in a remote collaboration. In the treatment group, they could see the gaze of their partner displayed on the screen in real-time. In the control group, they could not. Dyads in the treatment group achieved a higher quality of collaboration and a higher learning gain. In this paper, we describe how network analysis techniques can further illuminate these results, and contribute to the development of 'collaboration sensing'. More specifically, we describe two contributions: first, one can use networks to visualize and explore eye-tracking data. Second, network metrics can be computed to interpret the properties of the graph. We conclude with comments on implementing this approach for formal learning environments.

Categories and Subject Descriptors
K.3.1 [**Computer Uses in Education**]: Collaborative Learning

General Terms
Algorithms, Experimentation, Human Factors.

Keywords
Network Analysis, Eye-tracking, Computer-Supported Collaborative Learning, Awareness Tools.

1. INTRODUCTION
Nowadays, massive datasets are becoming available for a wide range of applications. Education is no exception to this phenomenon: cheap sensors can now detect every movement and utterance of a student. On the web, Massive Open Online Courses (MOOCs) collect every click of users taking classes online. This information can provide crucial insights into how learning processes unfold *in situ* or in a remote situation. However, researchers often lack the tools to make sense of those giant datasets; our contribution is to propose additional ways to explore massive log files, such as eye-tracking data.

2. RELATED LITERATURE
Our work lies in the intersection between basic network analysis and studies of the effects of gaze awareness on collaborative learning. While there is literature in both of these areas, there appears to be none squarely in the intersection of those two domains; as such, we believe the proposed work is novel and relevant to generating insights. We discuss the literature from related areas in order to justify our proposed work. In this section we look briefly at eye-tracking studies on collaborative learning, basic network analysis techniques, and at examples employing simple network analysis of eye tracking data.

Previous studies in CSCL (Computer-Supported Collaborative Learning) have used eye-trackers to study joint attention in collaborative learning situations. For instance, Richardson & Dale [8] found that the number of times gazes are aligned between individual speaker–listener pairs is correlated with the listeners' accuracy on comprehension questions. Jermann, Nuessli, Mullins and Dillenbourg [4] used synchronized eye-trackers to assess how programmers collaboratively worked on a segment of code; they contrasted a 'good' and a 'bad' dyad, and their results suggest that a productive collaboration is associated with high joint visual recurrence. In another study, Liu [6] used machine-learning techniques to analyze users' gaze patterns, and was able to predict the level of expertise of each subject as fast as one minute into the collaboration (96% of accuracy). Finally, Cherubini, Nuessli and Dillenbourg [2] designed an algorithm which detected misunderstanding in a remote collaboration by using the distance between the gaze of the emitter and the receiver. They found that if there is more dispersion, the likelihood of misunderstandings is increased. In all those studies, however, no data-mining techniques were used to uncover more complicated patterns. We thus propose to build large networks based on eye-tracking data. Our work deals mainly with basic graph property determination, since it is an exploratory attempt at building networks on top of gaze movements. This includes but is not limited to network size, degree distribution, clustering coefficient, and so forth [5]. By analyzing these basic attributes of the networks of eye tracking data, we lay the foundation for future research, which can control for various network properties in order to determine their effect on study outcomes.

We are unaware of existing studies that have tried to apply network analysis tools on eye-tracking data We believe that analyzing networks based on subjects' gaze behavior can shed a new light on collaborative learning processes. In the next section we describe our dataset and describe our work.

3. THE CURRENT STUDY
The first author previously conducted an experiment [9] where dyads of students remotely worked on a set of contrasting cases [1]. The students worked in pairs, each in a different room, both looking at the same diagram on their computer screen. Dyads were able to communicate through voice. Their goal was to learn how the human brain processes visual information (Fig. 2). In one condition, members of the dyads saw the gaze of their partner on the screen; in a control group, they did not have access to this information. This intervention helped students achieve a higher

LAK '13, April 08 - 12 2013, Leuven, Belgium

quality of collaboration and a higher learning gain compared to the control group. To our knowledge, this is the first study that was able to highlight the learning benefits of this kind of "gaze-awareness" tool.

Participants in the "visible-gaze" group outperformed the dyads in the "no-gaze" condition for the total learning gain: $F(1,40) = 7.81$, $p < 0.01$. For the sub-dimensions, they also scored higher on the transfer questions $F(1,40) = 4.47$, $p < 0.05$. With a larger sample, the difference is likely to be also significant for the terminology questions $F(1,40) = 3.59$, $p = 0.065$ and for the conceptual questions $F(1,40) = 2.11$, $p = 0.154$, since the effect sizes are between medium and large (Cohen's d are 0.62 and 0.5, respectively). Additionally, in our previous submission [9] we manually categorized each member of the dyad as "leader" and "follower". The leader was chosen to be the member that starts more conversations, and leads the dyad's problem-solving processes during the experiment. After the collaboration on the diagram, all participants then individually completed a short test that examined their learning gain about the topic. Their performance (number of correct answers) on this test was taken as their "total learning score". Interestingly (see Fig. 1), we found an interaction effect between those two factors (experimental conditions and individuals' status) on the total learning score: $F(1,38) = 5.29$, $p < 0.05$. Followers learned significantly more when they could see the gaze of the leader on the screen. We also rated the quality of collaboration of each dyad using Meier, Spada and Rummel's [7] rating scheme. Our results show that Dyads in the treatment group ("visible gaze") had a higher quality of collaboration: $F(1,10) = 24.68$, $p < 0.001$ (mean for the treatment group = 7.18, SD = 3.75; mean for the control group = 0.6, SD = 6.2). More specifically, those dyads were better at sustaining mutual understanding ($F(1,10) = 21.78$, $p < 0.01$), pooling information ($F(1,10) = 15.94$, $p < 0.01$), reaching consensus ($F(1,10) = 31.25$, $p < 0.001$) and managing time ($F(1,10) = 27.50$, $p < 0.001$).

The eye-tracking data of this study, however, has been largely unexploited so far. In the current work, we use network analysis techniques to describe how subjects in the "visible-gaze" condition outperformed subjects in the "no-gaze" condition.

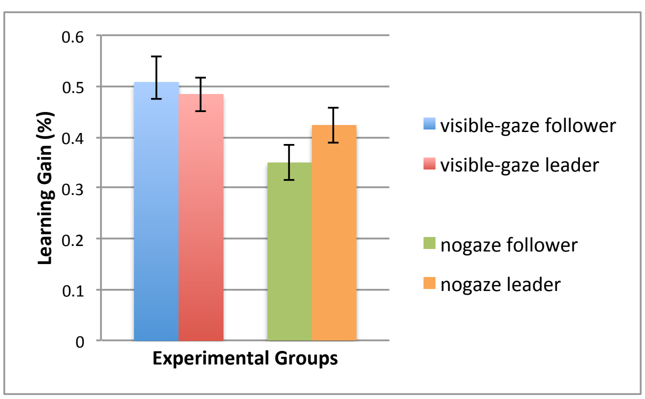

Figure 1: results of the learning test in the previous study. We found that students using the gaze-awareness tool outperformed the students who did not have access to it on the learning test. "Follower" in particular benefited from this intervention.

4. CONSTRUCTING GRAPHS WITH EYE-TRACKING DATA

4.1 Goals

Our goal is twofold: first, we want to provide an alternative way to explore eye-tracking data. This approach involves data visualization techniques, such as force-directed graphs. We believe that uses of visualization techniques for representing massive datasets can provide interesting insights to researchers. Second, we want to compute network measures based on those graphs; our goal is to examine whether some metrics are statistically different across our two experimental groups. Additionally, those metrics can provide interesting proxies for estimating dyads' quality of collaboration.

4.2 Deriving Nodes and Edges from Gaze Movements

Nodes can be created in various ways from this dataset. One choice would be to convert every pixel on the screen to a node. For obvious performance reasons, and to make computations more meaningful, nodes should contain an area larger than a pixel. A better choice would be dividing the screen into a uniform grid of nodes. Other methods involve building/detecting nodes empirically. For example, it is possible to sum the total gaze time for each pixel (by all dyads), and then cluster those pixels into nodes by the Mean-Shift algorithm. To limit the scope of this paper we limit ourselves to the first approach. For all the subsequent analyses, we divided the screen into 44 different areas based on the configuration of the diagrams (Fig. 2). Students had to analyze 5 contrasting cases; the answer to the top left and top right cases were given. Possible answers were given on the right. Thus, we have 30 areas that cover the diagrams of the human brain and 8 areas that cover the answer keys.

Edges can represent many aspects of the subject's behavior. For example, one can do path analysis and convert a dyad's gaze into a "path" on the nodes. Here, it is possible to analyze the steps in the gaze processes that dyads went through (e.g., by counting common sub-paths), how fast they switch between nodes, and the average number of times that they visit a node.

In one analysis, we constructed undirected weighted graphs, where a time-sliding window (e.g., of 2 seconds width) detects when a given dyad gazes at two screen areas within that time-frame. Then one can increment the edge weight by a certain quantity. Unlike the path analysis, this undirected graph has no notion of chronology because the temporal relationship between nodes is expressed as edge weights.

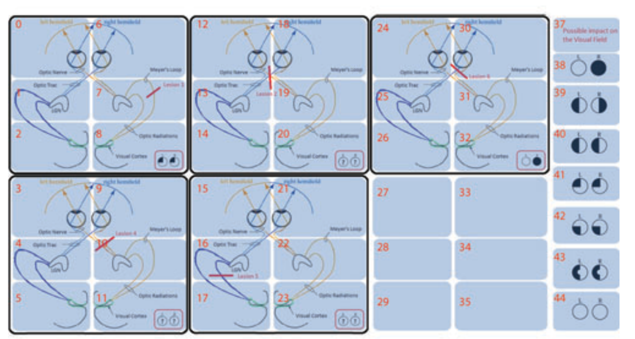

Figure 2: To create the nodes, we choose to divide the screen in 44 different areas based on its visual configuration.

For our first attempt, we focused on the simplest solution: each time a participant gazed between two regions, we created (or incremented the weight of) an edge. Future work should explore alternative ways of creating both nodes and edges based on eye-tracking data. However this simple method generates quite interesting results, described in the following section.

4.3 At the Individual Level

In this section we describe graphs created with individuals as the unit of analysis: each network is built by using the eye-tracking data of one subject (Fig. 3). The label on each node corresponds to a screen region as defined in Fig 2. The size of a node shows the number of fixations on this area. Node colors correspond to screen section. Blue nodes correspond to a diagram region (major/left side of the screen). Orange nodes correspond to answer keys (right column of the screen). An edge represents a saccade between two regions. The width of an edge shows the number of times a subject compared those two regions. Those graphs were implemented with a force-directed layout and can be directly manipulated on a web page[1].

This approach already shows interesting results: we can observe that subject 1 (on the top) spent a lot of time understanding the diagram on the top right corner of the screen; however (s)he mostly neglected the answers on the right. Subject 2 (on the bottom), had a completely different strategy: (s)he intensively compared answers and diagrams. Thus, with this visualization one can quickly identify patterns and build hypotheses to investigate.

One limitation of this approach is known as the “hair ball” problem in data visualization: since the graph is quite dense, every node is connected to a lot of other nodes and thus makes interpretations difficult. This problem is inherent to eye-tracking dataset: since an edge is a saccade, each node is going to be connected to at least two other nodes. Moreover, due to the limited amount of potential nodes, our graphs are bound to be highly connected and highly clustered. Another limitation is the fact that this visualization totally ignores the collaborative aspect of the study [9]. In previous results, dyadic synchronization was found to be a critical factor for a positive learning experience.

In the next section, we describe how we circumvented those issues. We sought to create smaller and more informative graphs by focusing on dyads instead of individuals. Those graphs provide a different window into our dataset.

4.4 At the Dyad Level (Joint Attention)

Our next attempt involved building one graph for each dyad. Here, we want to capture the moments in which dyad members were cooperating. The nodes correspond to the screen areas, as previously defined. At the dyad level, however, a node will only appear in the dyad graph if both dyad members gazed at the respective screen area within a 2 seconds time-window of each other.

Small graphs with few nodes are characteristic of poor collaboration, and large graphs with highly connected nodes show a strong flow of communication across members of the dyad. Figure 4 provides an example of this kind of contrast.

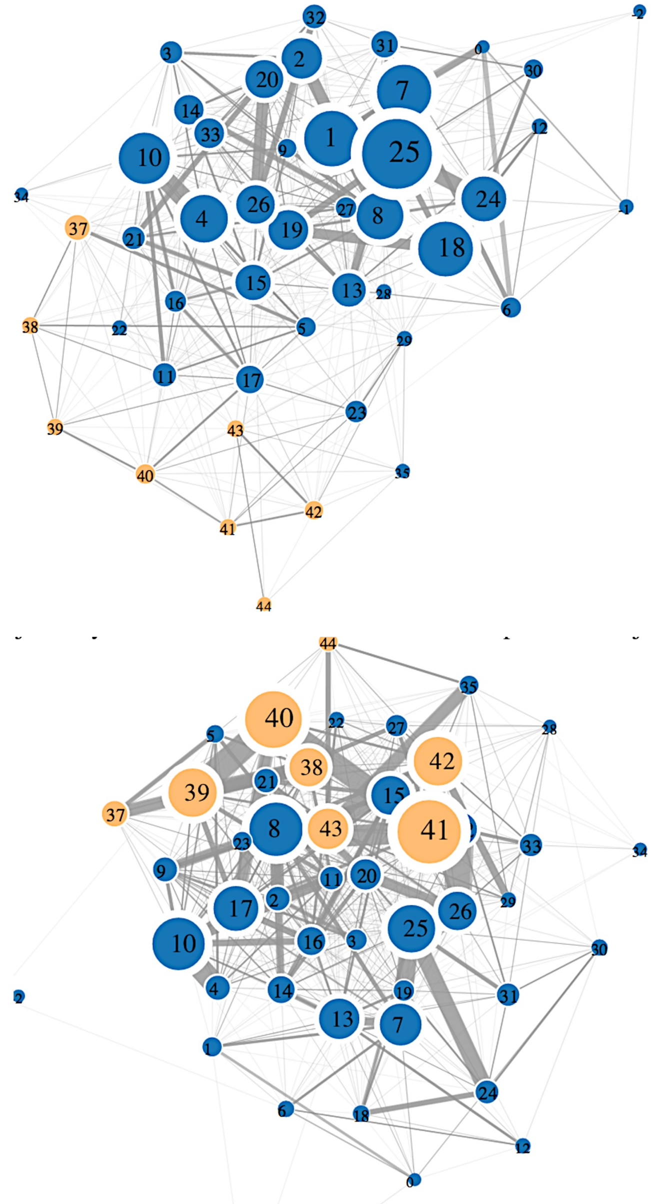

Figure 3: Two graphs created based on individuals' data. Blue means brain diagram, orange means answer key on the right of the screen. Both graph suffer from the “hair ball” problem since they contain a lot of edges.

The color scheme of the nodes is identical to the individual graphs. However, the node size in the dyad graphs is proportional to the number of times dyad members looked at the respective screen area within a 2 second time window. The choice of “2 seconds” is based on the work done by Richardson & Dale [8], where they find that it takes a *follower* about 2 seconds to look at the screen area that the *leader* is discussing. Edges are defined as previously (i.e., number of saccades between two areas of the screen).

[1] The visualizations described in this paper can be accessed via stanford.edu/~schneibe/cgi-bin/d3/examples/force/force.php

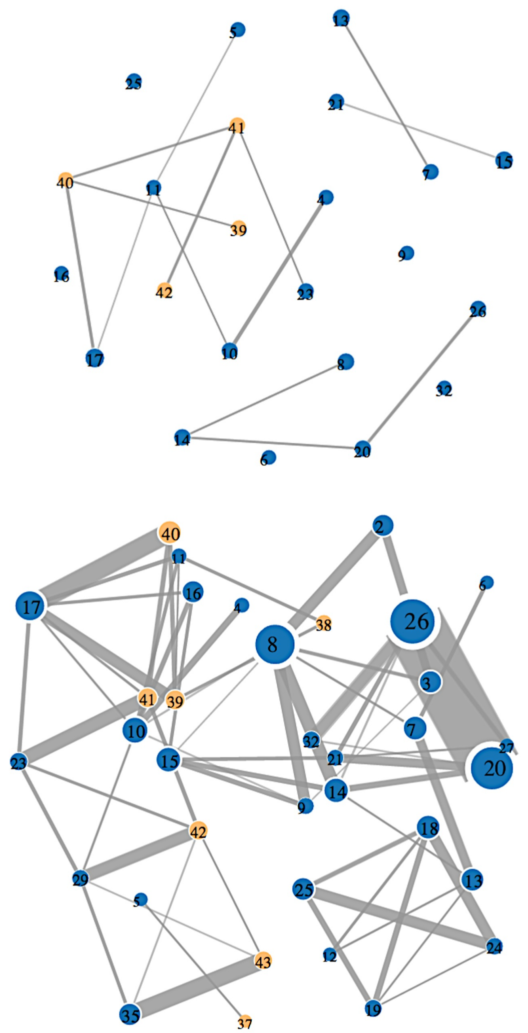

Figure 4: Graph build on dyads' data. The size of each node reflects the number of moments of joint attention members of the group shared on one area of the screen. Graph on the top is from a dyad in the "no-gaze" condition; graph on the bottom is from a dyad in the "visible-gaze" condition.

Again, from a data visualization perspective, this approach conveys key patterns in collaborative learning situations. The top graph in Fig. 4 shows a dyad in the "no-gaze" condition; one can immediately see that students did not share a common attentional focus very often. Nodes are small and poorly connected. The graph on the bottom represents a dyad in the "visible-gaze" condition and is a strong contrast to the previous example: here students are looking at common things much more frequently and those moments of joint attention provide opportunities to compare different diagrams or answers. Nodes are bigger and better connected.

Based on this new dataset, we computed various network metrics. We found that the average size of the nodes was significantly bigger in the "visible-gaze" condition: $F(1,12) = 7.19$, $p = 0.02$ and there were more edges: $F(1,12) = 4.9$, $p = 0.047$. As expected, the average betweenness centrality (a measure of a node's centrality in a network, computed by counting the number of shortest paths from all vertices to all others that pass through that node) was also higher for this group: $F(1,12) = 6.44$, $p = 0.026$. Those results indicate that we can potentially separate our two experimental conditions solely based on network characteristics.

More interestingly, several measures were significantly correlated with the groups' quality of collaboration (as defined by the rating scheme described in the methods' section [7]). The average size of a node was correlated with the overall quality of collaboration ($r(16) = .52$, $p = 0.039$), students' orientation toward the task ("each participant actively engages in finding a good solution to the problem"): $r(16) = .54$, $p < 0.001$, and students' reciprocal interaction ("partners treat each other with respect and encourage one another to contribute their opinions and perspectives"): $r(16) = 59$, $p < 0.001$. The number of the nodes in the graph was correlated with the sub-dimension "Reaching Consensus" ("Decisions for alternatives on the way to a final solution (i.e., parts of the diagnosis) stand at the end of a critical discussion in which partners have collected and evaluated"): $r(16) = .65$, $p < 0.001$ and the sub-dimension "Information pooling ("Partners try to gather as many solution-relevant pieces of information as possible"): $r(16) = .52$, $p = 0.002$. Betweenness centrality was correlated with all the sub-dimensions above - but also with the sub-dimension "Sustaining Mutual Understanding" ("Speakers make their contributions understandable for their collaboration partner, e.g., by avoiding or explaining technical terms from their domain of expertise"): $r(16) = .37$, $p = 0.037$.

5. DISCUSSION

Our preliminary results show the relevance of using network analysis techniques for eye-tracking data. In particular, we found this approach fruitful when applied to social eye-tracking data (i.e., a collaborative task where the gaze of two or more participants are recorded at the same time).

More specifically, we found that different aspects of collaborative learning were associated with different network measures. The average size of a graph's nodes appeared to be a good proxy for students' orientation toward the task and their level of reciprocity toward their partner; the number of nodes can be used to estimate to what extent dyads try to reach a consensus and pool information to find a good solution to the problem at hand. Finally, the betweenness centrality of a graph appears to be an indicator of how well students try to sustain mutual understanding between one another. Of course, more work is needed to replicate those results. But overall, we find that network analysis techniques can be used advantageously to further our understanding of group collaboration processes.

In terms of future work, we are currently exploring several aspects of the research described above. The most direct extension is to apply machine-learning techniques to predict dyads' quality of collaboration using network metrics. Early results make us believe that networks' characteristics can be used as relatively good input features for a Support Vector Machine (SVM) algorithm. Another interesting line of work is to explore the way graphs evolve over time: for instance, we would expect the betweenness centrality to increase as the team develop a higher mutual understanding an agreement over the time. Finally, more work is needed to interpret the meaning of the correlations described above: for instance, it is not entirely clear why betweenness centrality is strongly associated with developing mutual understanding in the group. At this stage it is probably necessary to conduct a more fined-grained analysis of those results, for instance by focusing on one particular group and defining key events associated with a higher betweenness centrality (e.g. students jointly revisiting a particular node, and making connections with hypotheses previously developed during the activity).

Our work has limitations. First, we studied only one particular kind of collaboration (i.e., remote collaboration). It is likely that *in situ* interactions are different from online collaborative work. Secondly, we only had 9 dyads in each experimental group; with more subjects we would probably find more statistically significant patterns. Thirdly, it is possible that the two experimental conditions are interfering with our results: collaboration traits may be exacerbated by our gaze-awareness tool and thus not reflect natural patterns of social interaction. Finally, there are other interesting network metrics that we did not use for this preliminary analysis. Future work shoulde replicate those results in other settings and look at more complex properties of graphs.

6. CONCLUSION

This work provides two significant contributions. First, we developed new visualizations to explore social eye-tracking data. We believe that researchers can take advantage of this approach to discover new patterns in existing datasets. Second, we showed that simple network metrics can serve as proxies for evaluating the quality of group collaboration. As eye-trackers become cheaper and widely available, we can develop automatic measures for assessing people's collaboration. Such instrumentation would enable researchers to spend less time coding videos and more time exploring patterns in their data. In formal learning environments, such measures could be computed in real time; teachers could employ such metrics of 'collaboration sensing' to target specific interventions while students are at work on a task. In informal networked learning, collaboration sensor metrics could trigger hints or provide other scaffolds for guiding collaborators to more productive coordination of their attention and action. We also envision the extension of such network analyses as these for eye-tracking during collaboration to other interaction data related to interpersonal coordination and learning, such as gestures and bodily orientation.

7. ACKNOWLEDGMENTS

We would like to thank the teaching staff of CS224W ("Social and Information Network Analysis" taught by Assist. Prof. Jure Leskovec at Stanford) for their help and support during this project.

8. REFERENCES

[1] Bransford, J.D. and Schwartz, D.L. Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, (1999), 61–100.

[2] Cherubini, M., Nüssli, M.-A., and Dillenbourg, P. Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. *Proceedings of the 2008 symposium on Eye tracking research & applications*, ACM (2008), 173–180.

[3] Erdős, P. and Rényi, A. On the Evolution of Random Graphs. *Publication Of The Mathematical Institute Of The Hungarian Academy Of Sciences,* (1960), 17–61.

[4] Jermann, P., Mullins, D., Nüssli, M.A., and Dillenbourg, P. Collaborative Gaze Footprints: Correlates of Interaction Quality. *Proceedings of CSCL*, (2011), 184–191.

[5] Kleinberg, J. The small-world phenomenon: an algorithm perspective. *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, ACM (2000), 163–170.

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

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

[8] Richardson, D.C. and Dale, R. Looking To Understand: The Coupling Between Speakers' and Listeners' Eye Movements and Its Relationship to Discourse Comprehension. *Cognitive Science 29*, 6 (2005), 1045–1060.

[9] Schneider, B., & Pea, R. (submitted). Using Eye-Tracking Technology to Support Visual Coordination in Collaborative Problem-Solving Groups. *International Conference on Computer-Supported Collaborative Learning,* 2013.