10th International Conference on Computer Supported Collaborative Learning
Madison, Wisconsin, June 15-19, 2013

Using Dual Eye-Tracking Measures to Differentiate Between Collaboration on Procedural and Conceptual Learning Activities

Daniel Belenky[1], Michael Ringenberg[1], Jennifer Olsen[1], Vincent Aleven[1], Nikol Rummel[2]

[1]Human-Computer Interaction Institute
Carnegie Mellon University
dbelenky@andrew.cmu.edu, mringenb@cs.cmu.edu,
jkolsen@cs.cmu.edu, aleven@cs.cmu.edu

[2]Institute of Educational Research
Ruhr-Universität Bochum, Germany
nikol.rummel@rub.de

Dual eye-tracking measures enable novel ways to test predictions about collaborative learning. For example, the research project we are engaging in uses measures of gaze recurrence to help understand how collaboration may differ when students are completing various learning activities focused on different learning objectives. Specifically, we hypothesize that collaboration may be particularly beneficial for facilitating the development of conceptual knowledge, but perhaps less optimal for the development of procedural skills. As one test of this hypothesis, we anticipate that dyads working on a conceptual problem should show longer, more sustained periods of gaze convergence, as opposed to dyads working on procedural problems. We present preliminary data from one dyad that supports this hypothesis. Additionally, we discuss other potential uses for dual eye-tracking data in the context of our larger research project.

Technology can be a powerful tool for advancing both research and practice in education. For example, many studies have shown a benefit for student learning when using Intelligent Tutoring Systems (ITSs). These systems support students as they engage in complex problem-solving tasks by providing continuous and detailed guidance. ITSs have been shown to lead to significant improvement in student learning in a number of domains (Beal, Walles, Arroyo, & Woolf, 2007; Graesser, Chipman, Haynes, & Olney, 2005; Koedinger & Aleven, 2007; VanLehn, et al., 2005). The current study reported here expands upon a line of ITSs research on Cognitive Tutors, which are one type of ITS, developed with a firm grounding in cognitive theories of learning (Aleven, McLaren, Sewall, & Koedinger, 2009). Of particular relevance for this workshop are the ways in which dual eye-tracking measures are being used to evaluate hypotheses about how students engage with this system.

To date, Cognitive Tutors have been primarily focused on teaching students as they work individually. However, there are many reasons to believe that expanding this software to include opportunities for collaborative learning activities will be beneficial. A large number of studies have demonstrated the effectiveness of collaboration for improving students' learning and problem-solving, both in face-to-face (e.g., Slavin, 1996) and computer-based settings (for an overview, see Lou, Abrami, & d'Apollonia, 2001). In the present study, we are particularly interested in understanding how different types of learning outcomes may be more readily supported by collaboration. That is, collaboration may particularly useful for developing conceptual knowledge, as prior research indicates that this type of knowledge is developed through sense-making activities such as explanation and comparison, which can occur when students work with each other. This can be contrasted with individual instruction, which, by allowing multiple opportunities for problem-solving practice, may help develop procedural fluency with problem-solving steps. We know of only one study that has directly tested the hypothesis that individual and collaborative learning may help develop different types of knowledge (Mullins, Rummel, & Spada, 2011). The results of that study did indicate that individual learning activities enhanced learning of problem-solving procedures more strongly, while collaborative learning activities supported the development of conceptual knowledge.

Given this prior theoretical and empirical evidence, the larger research agenda for the project is to construct a learning platform that utilizes both individual and collaborative learning modes to maximize both conceptual and procedural learning. However, for the purposes of this workshop submission, we will focus exclusively on the collaborative side of this research agenda, and, in particular, on how dual eye-tracking measures may allow us unique opportunities to evaluate a number of interesting hypotheses.

As one example of the sort of hypotheses dual eye-tracking will allow us to investigate, we predict that students working on the conceptual problems will have more intense (i.e., longer in duration and with a higher proportion of shared fixations) periods of gaze convergence, as they closely align their attention to discuss conceptual features of the problems. In contrast, the procedural problems may have more a diffuse pattern of gaze convergence, as students may generally be working on similar steps, but, as only one is tasked with completing each of the steps, the students' attention may wander to other parts of the interface. As a preliminary first step, we compare the gaze recurrence of one dyad working collaboratively on a problem aimed at fostering conceptual knowledge versus one aimed at developing procedural knowledge.

The type of gaze recurrence analysis we use builds upon prior research, which has been used to highlight the degree to which two people's attention is focused on the same visual area. For example, Richardson and Dale (2005) showed individual participants images of the characters from the television show "Friends," and asked them to talk about things that happened on the show while recording their fixations across these images. The researchers subsequently played back the audio that was generated this way to other participants, and recorded their gaze behavior on the same visual display as the speaker had originally seen. Using gaze recurrence analysis, they found that the listeners' gazes matched up very closely with the speaker's gaze, showing that people's visual attention can align quite naturally when relevant visual markers are present. Additionally, they found that more convergence with the speaker's gaze was related to better comprehension. More closely related to our study, Nüssli (2011) recorded gaze behavior dyads worked collaboratively on a programming task. The gaze recurrence of different dyads was analyzed, and more overlap in visual attention was found for participants who performed better at the task. In particular, dyads who went on to demonstrate a higher level of understanding were found to spend more of their time jointly fixating on more meaningful elements of the display, rather than simple user interface aspects.

These studies have shown that collaborators do indeed orient their visual attention to similar parts of a visual display, and, perhaps more critically, that the degree to which they do so is predictive of subsequent performance. Put another way, there is strong evidence that one can tell the "good" dyads from the "bad" (as judged on post-test metrics collected after the collaboration) by using gaze recurrence analysis. In the current work, we hope to extend this past research in a few ways. First, we will collect collaboration data from students working on a real-world ITS. While this offers an opportunity to test the validity of such measures, there are some elements of this ITS that may complicate the collection of eye-tracking measurements. For one thing, the interface of this ITS is highly interactive with many different elements appearing on the screen at different points in the problem-solving progression. For another, the two collaborators are not always receiving the exact same information at the same time, since one way in which collaboration is supported in this system is by giving students different pieces of information that they must share with their partners. While this may complicate gaze recurrence analysis, it provides an opportunity to test the applicability of such measures outside of the more artificial settings it has been examined in so far.

Another contribution of this line of research will be in testing how much additional information gaze convergence provides above the other data streams we will be collecting. Specifically, we will be collecting pre and post-test data, as well as tutor log data and audio data of the collaboration. While gaze convergence is an interesting analysis tool, it is unclear whether it provides additional explanatory power, beyond these other measures. While we are still at a preliminary stage of this project and cannot fully address this question in this submission, we believe that our project provides a unique opportunity to explore this issue.

Collaborative Fractions Tutor

For the current research study, we have created an adapted version of the Fractions Tutor (Rau, Aleven & Rummel, 2013), a successful Cognitive Tutor that has been developed in our lab to help teach fractions to 4th and 5th graders. The Fractions Tutor covers a comprehensive set of 4th and 5th grade fractions topics and activities, such as interpreting graphical representations of fractions, creating graphical representations, ordering fractions, determining if fractions are equivalent, and more. For the current study, this web-based ITS has been adapted in such a way that it can now be used collaboratively by a pair of students working at two different machines. These students can communicate to one another in different ways, depending on the set-up of the study or the classroom; the machines may be located in the same room, or they may have a voice-based connection using a program like Skype. In the present study, collaborating students see slightly different views of the same problem and problem state. Additionally, they have different responsibilities for moving the problem forward. These different views and actions have been designed so as to support a particular kind of collaboration script, where each student is responsible for different elements of the problem. This kind of remote and synchronous computer-supported collaborative system has been successfully implemented in both laboratory and classroom studies (e.g., Olsen, Belenky, Aleven, & Rummel, in press; Walker, Rummel & Koedinger, 2011).

Additionally, we have created a set of conceptually oriented and procedurally oriented learning activities. In constructing these, we have differentiated between these two types of knowledge by following theoretical definitions laid out in prior research. For example, conceptual knowledge deals with "the principles that govern a domain and of the interrelations between pieces of knowledge in a domain" (Rittle-Johnson & Alibali, 1999, p. 175). Procedural knowledge, in contrast, is defined as the ability to execute actions to generate the correct solution to a problem (Rittle-Johnson & Alibali, 1999). As discussed earlier, different types of instructional activities (e.g., sense-making for conceptual knowledge, practice for procedural) are expected to promote the development of each type of knowledge (Koedinger, Corbett, & Perfetti, 2012), and the activities we have designed emphasize these different learning activities. Specific examples of how we have designed these different instructional activities will be discussed next, along with additional details of the pilot study we have been conducting.

Methods

We present pilot data collected in our laboratory from one dyad, comprised of two 4th grade students, who completed a set of four procedural and four conceptual learning activities (see Figures 1 and 2 for representative examples of the procedural and conceptual activities, respectively). While different collaboration scripts were developed for each type of problem, the scripts for both the procedural and conceptual problems shared a number of features that were designed to foster productive collaboration. Specifically, on each problem, students are assigned either a "problem solver" or "helper" role, and each role is given responsibility for solving different aspects of the problem. Additionally, each student is given different information that they must share with their partner, giving each student an opportunity to elaborate and explain to their partner. Students alternate between the problem solver and helper roles across problems.

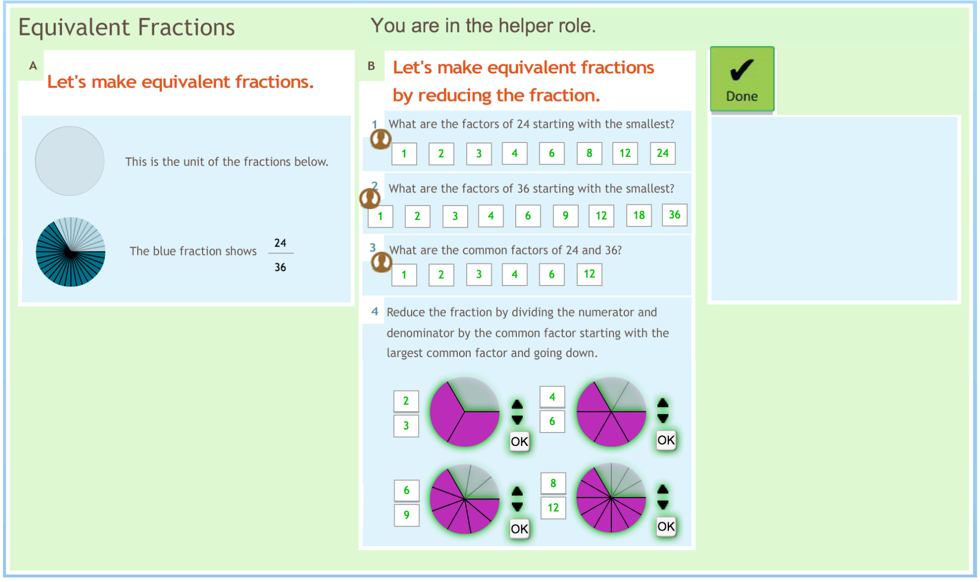

Figure 1. Procedural collaborative problem.

In the procedural problem, the participants completed a number of problem-solving steps to generate a set of fractions equivalent to a given fraction. Figure 1 shows the interface at a point when the entire problem has been solved, but students working through the interface are only shown one piece of the problem at a time, with each next step being revealed when they have successfully completed the prior one. The tutor guides the students through these steps, beginning with one student listing all of the factors of the numerator (see panel B1 in the figure). When they have done so correctly, they move on to finding the factors of the denominator, before moving on to listing the common factors of the numerator and denominator. The student knows that it is his job to complete these steps because they are marked with a silhouette icon, indicating that this is a step he must complete (his partner does not have the ability to interact with the interface on those steps). After this is completed, his partner completes panel B4, where he must divide by the common factors to create equivalent fractions, first writing them out numerically, and then creating a graphical representation of that fraction. Throughout the problem, students are given particular actions to complete based on their assigned role; either problem-solver or helper. In this problem, the helper was responsible for generating the lists of factors, and the problem-solver was responsible for using that information to create the equivalent fractions. These distributed roles help keep both participants actively engaged, and students alternate between the problem-solver and helper roles across problems.

The conceptual problem (see Figure 2) begins with each student seeing the same two fractions presented as circle representations and symbolically (e.g., 3/5 and 9/15 shown in Panel A), and being asked to compare them (i.e., decide whether or not they are equivalent). They are also shown the B1 panel, which provides a sample explanation from another student who describes whether or not the fractions are equivalent and why. Each student received one particular explanation and had to share it with their partner, as denoted by the star icon, which indicates that this part is information that only this student has, and it is the student's responsibility to share this information with their partner. The silhouette icon again indicates that this is a step that only the student can complete. As in the procedural problem, the subsequent parts of the problem are only revealed when students have finished the current part of the problem. In panel B2, each student has a space where they can mark whether they agree that a certain comparison between the two fractions is true or not, and both students need to select the correct options before they are allowed to continue. That is, only after the student in the helper role chooses a particular option can the problem-solver choose that option, and receive feedback from the ITS on whether or not they are correct. Finally, in panel B3, they need to decide which of the two student sample responses they read was correct.

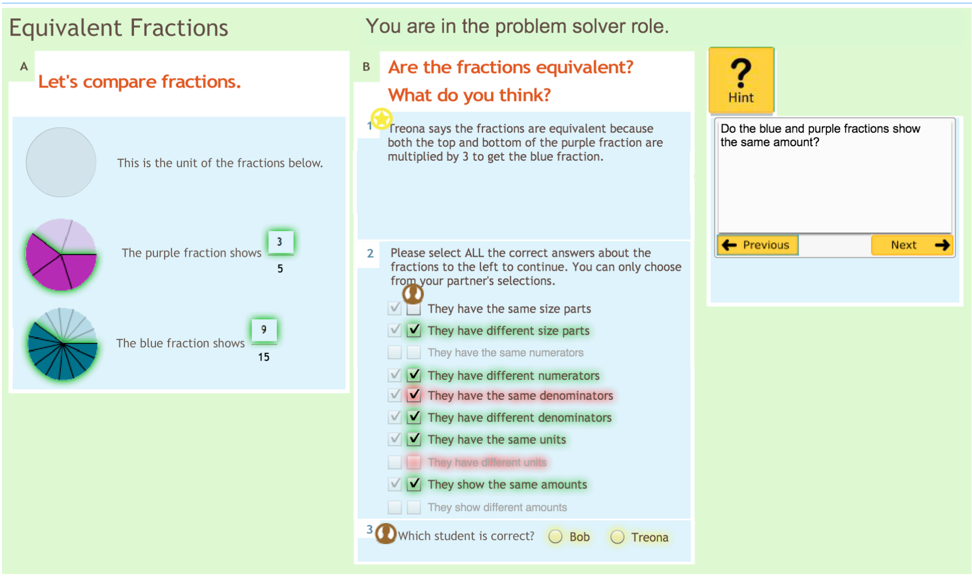

Figure 2. Conceptual collaborative problem.

Figure 2 also illustrates a few other elements of the tutoring interface that are important to note. One is the visual feedback that students receive. When they successfully complete a step, it turns green; when they are incorrect, it turns red. Additionally, hints are available on demand via the hint button. When a student clicks on this button, text appears in the hint window that guides them towards successfully completing that step. Asking for a hint also has the tutor highlight the step that the student should be currently working on (in Figure 2, the names “Bob” and “Treona” are highlighted in B3, at the bottom of the screen). In terms of collecting eye-tracking data, it is important to note how the different sections of the screen are separated and take up a significant amount of visual angle, allowing for analyses, which break up the interface into distinct Areas of Interest (AOIs) to be conducted.

The two students in the pilot dyad presented here worked on two different machines located in different rooms. They communicated with one another via Skype. In addition to collecting log data from the tutor, each machine had a SMI Red 250 Hz infrared eye-tracking camera attached to the monitor that the student worked on, which recorded their eye-movement behavior.

While many different measures are possible, our initial analyses have focused on gaze recurrence, following other research on understanding collaboration via eye-movement data (e.g., Jermann, Mullins, Nüssli, & Dillenbourg, 2011). Gaze recurrence plots were generated for the dyad for each of the two problems they completed, using the CRP toolbox for MatLab (available at http://tocsy.pik-potsdam.de/CRPtoolbox/index.html; Marwan & Kurths, 2002). These recurrence plots show the degree to which two collaborators have similar gaze patterns at any given point in time of the collaboration, with darker areas indicating points in time where participants’ gaze were in agreement more closely. Dark areas along the diagonal (and closely around it) indicate that participants were looking at the same areas at the same time, while points either above or below this line indicate that one participant led and the other followed the other’s gaze.

More specifically, gaze recurrence was calculated by binning the data into two-second slices. For each of these two-second slices, data pertaining to the gaze location during fixations (non-fixation based data was removed) was collected, described in x,y coordinates relative to the pixels on the screen. As the eye-tracker was sampling at 250 Hz, this provides a maximum of 500 such data points for each two-second slice for each participant. Each of these data points, representing where the student’s fixation was located for that particular sample, was compared to each of the other student’s fixation locations for a corresponding two second slice. This comparison was made using a simple linear distance comparison; that is, in terms of pixels, how far away were the two students’ gazes? The proportion of these data points in which the students’ gazes were located in the same area (defined, in this analysis, as being less than 80 pixels, or 22.6 millimeters, away) was calculated for each two second slice, and graphed according to a color scale, with darker colors indicating a larger proportion of fixation-based data points being located in the same area. Eighty pixels was chosen as the criterion for agreement because it is similar to the criterion used in prior research (i.e., 70 pixels in Jermann et al., 2011) and because the interface elements are each about 80 pixels large.

Results

The first problem we address is the *procedural problem* (see Figure 3). Students working on this problem seemed to pass through distinctive stages of gaze convergences, although these periods were not very focused and had a relatively small percentage of convergent gazes. In particular, the students begin with relatively little convergence until about two minutes into the problem. Analysis of the interaction indicates that the two students first spent a bit of time discussing the different aspects of the problem and what exactly the steps were asking them to do. Additionally, the student in the helper role, who is responsible for entering all of the factors of the numerator and denominator (see panel B1 in Figure 2), relied on her partner to tell her how to complete this step.

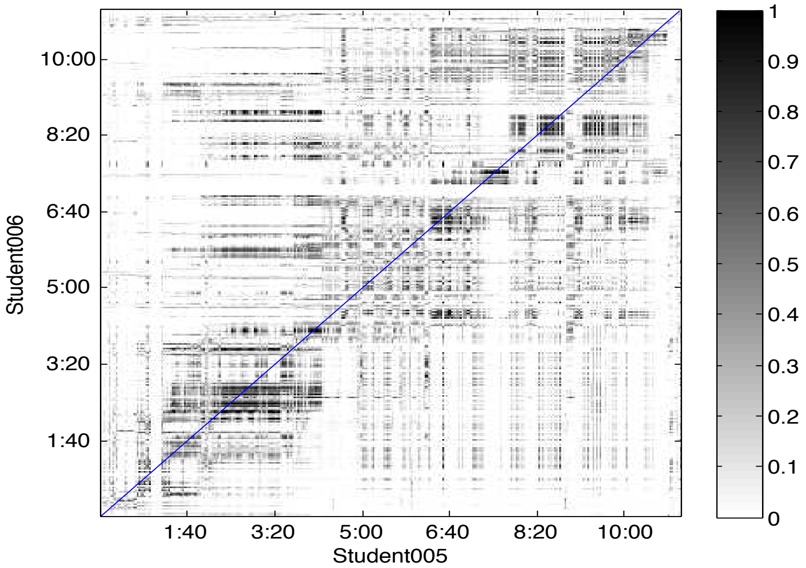

Figure 3. Gaze recurrence for the procedural problem.

Around two minutes into the problem, the students enter an incorrect factor. At this point, the gaze convergence strengthens, as they begin engaging in a discussion about what the error is and how to correct it. The helper student, on the x-axis, continue looking in that area for a little while longer as she continues entering the next factors as well, while the problem solver lets his eyes wander over other aspects of the interface.

Next, in the middle of the graph, a period of weak gaze convergence is visible from about 4.5 minutes until about 6 minutes. This is the period where they move on to creating a series of fractions that are equivalent to the original one they were given, by dividing the numerator and denominator by the common factors (see panel B4 in Figure 1). This is also the step on which the active partner switches; up until now, the helper has been entering values into the interface, but now the problem solver does. As this is a new step, the partners begin by first discussing what the problem is asking them to do. They do this for about a minute, before they begin trying to enter values for the numerator of the reduced fraction.

Beginning around the 6 minute and 40 second mark, the dyad alternates between periods of intense gaze convergence and very little convergence. This appears to be due to the problem solver focusing intently on the problem, with his gaze occasionally going off screen as he looks as the keyboard. The helper, in contrast, is not focusing as intently on this part of the interface. Additionally, when the problem solver received negative feedback, such as when the graphical representation he created was misaligned, or, more substantively, when he entered a wrong numerator and was given feedback that he was incorrect, this sparked a brief period of intense gaze convergence, as the two discussed what exactly was wrong and how to proceed.

This can be contrasted to the gaze recurrence generated during the *conceptual problem* (see Figure 4). The gaze recurrence plot generated by this problem can be described as revealing three main periods of interactions, which are related to the different parts of the interface that are progressively revealed as the students successfully solved each step (note that students took less time on this problem, so the axes are different than in Figure 3). In the first, period, which lasted for the first 40 or so seconds of the problem, there was a moderate degree of gaze convergence between the two students. This corresponded to the period where both of the students saw the two fractions that are being evaluated in panel A (see the tutor interface in Figure 2) and each received a different sample response that they were supposed to share with their partner. The participants tended to be looking at either the two fractions or the text, but, as they alternated between these, they did not strongly seem to keep their gazes focused on the same area as their partner, producing somewhat weaker recurrence. In this period, each of the participants was supposed to click on different elements of the interface at different times before they could proceed, so the fact the recurrence was somewhat lower here was to be expected. Next was a period of about 25 seconds where one student (on the Y-axis) continued looking at some of those same regions, while the other did not. Overall, recurrence was quite low during this period, partly because the students became confused about how to proceed, and frequently looked off screen while discussing or asking the experimenters for help. Finally, the period between 80-115 seconds shows a high degree of gaze recurrence, as the two students worked together to decide which characteristics were true of the two fractions that were being compared for equivalence. As both students needed to act in concert, with both selecting a correct choice before it was accepted, this period generated a lot of discussion and a lot of focused attention on the same areas of the screen.

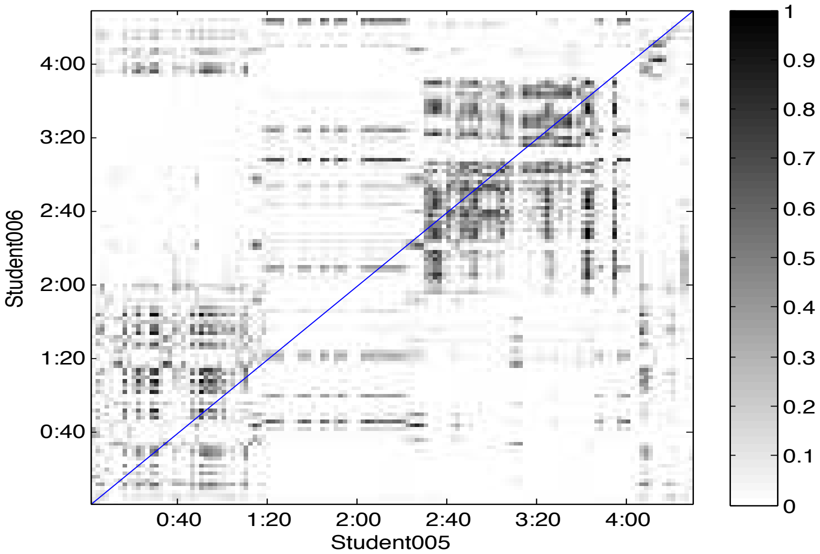

Figure 4. Gaze recurrence for the conceptual problem.

Conclusions and Future Directions

While we anticipate that our larger program of research will produce a number of contributions, of particular note for this workshop is the use of the eye-tracking to examine the differences between individual and collaborative learning activities in ITSs, and, as discussed here, comparing gaze behavior for procedural and conceptual problems. While still quite preliminary, this research evidences how dual eye-tracking can be used to answer important research questions about collaborative learning in computer-supported settings.

So far, we have begun to explore gaze recurrence measures, to test the hypothesis that these provide a useful index for the quality of collaboration between students, and to examine how they can complement the other data we are collecting, such as speech data, tutor log data, and pre/post test data. In the limited pilot data presented here, we observed that the conceptual problem led to discussions that were accompanied by a high degree of gaze recurrence. The procedural problem, in contrast, also generated some periods of high gaze convergence, but these were almost always in response to receiving tutor feedback about an incorrect step. If this pattern is observed with a more robust set of data, it may provide interesting evidence that conceptual problems can lead to fruitful discussions where both members of the dyad both focus their attention on the same idea. In contrast, procedural problems may be less likely to spark such discussions unless the students make errors, in which case the need to re-assess their knowledge may spark both students to focus their attention in the same place. Alternatively, it is also possible that gaze recurrence measures are useful for distinguishing between "good" and "bad" collaborations, but that they do not help differentiate between the different types of collaboration that may emerge for conceptual and procedural learning activities. As only pilot data is available for now, strong statements cannot be made either way about this possibility.

A related issue is whether gaze recurrence measures contribute additional explanatory power beyond the more traditional data streams we are collecting. For example, it may be possible that elements of collaboration captured in speech may provide additional paths for assessing the same aspects of collaboration. As such, it will be interesting to analyze the correlations (or lack thereof) between measurements of gaze convergence with measures of interactivity in speech, such as the number of turns taken. Gaze convergence, as a measure of quality of interaction, may echo other measures, which are simpler to collect and analyze. On the other hand, gaze convergence may yet provide additional information that is missing from other data streams, like the audio log of the collaboration. For example, there may be times where partners have high degrees of gaze convergence but low turn taking, such as when one partner explains something to their partner at length. Conversely, there may be episodes where partners are engaged in fruitful discussion with a high amount of turn-taking, but still end up with relatively low levels of gaze convergence. Exploring situations like these will provide a robust test of the utility of dual eye-tracking, given the other analysis options that are available.

In particular, we are interested in understanding how multiple data streams may be combined to provide a more complete description of the collaborative learning process. We have been able to insert time-stamped log messages from the Collaborative Fractions Tutor directly into the data file that the eye-trackers generate. This will aid in analysis, allowing us to investigate in a fine-grained way the connections between particular steps in the tutor and gaze behavior. If gaze behavior can be used as reliable proxy for cognitive processing, this type of analysis may inform us about the conceptual and procedural nature of the mental work the student is engaging in at any given time. Additionally, we will be able to examine interactions in a more in-depth manner. For example, it is possible that gaze convergence at particular moments, such as immediately following a mistake, is an important predictor of learning outcomes. As seen in the procedural problem presented here, student-generated errors may make for particularly fruitful episodes of collaboration. If this is the case, incorporating opportunities for such discussions into an ITS could lead to significant learning gains. By adapting systems such as AdaptErrEx, an ITS that support learning through erroneous examples (McLaren et al., 2012) to allow for collaboration and the collection of dual eye-tracking data, this hypothesis could be more fully explored in future work.

An additional path to explore will be to understand what eye-movement behaviors of dyads indicate about the cognitive processing they are engaging in. It is possible that different scan paths reveal different levels of understanding, or reflect different underlying computations. Linking gaze data with the tutor log data will allow for more in-depth testing of such hypotheses, as the tutor has been designed such that each step represents the use of a discrete knowledge component from the larger set of domain knowledge. Anderson and Gluck (2001) were able to document an eye-movement signature that indicated an error was about to occur. It may be possible to create similar *collaborative* eye-movement detectors that indicate whether the dyad is not engaging with the ITS, and to create some way to focus the dyad's attention back to where it could most productively be used. A similar approach has proven successful for individuals working with an ITS with a relatively simple interface (D'Mello, Olney, Williams & Hays, 2012), and extending such support to collaborators working with a more complex ITS would be a very valuable contribution.

Finally, the research project presented here is being conducted both in laboratory and in school settings, with the eye-trackers being used in studies directly in elementary schools. Students will participate at their own schools, spending some of their class time in a lab room we setup in their school. This will allow a greater amount of data collection opportunities, and also provide a degree of external validity to this research that cannot be achieved when only running participants in laboratory settings, as students will work with their normal classmates in a more natural setting.

The use of dual eye-tracking contributes to our abilities as researchers to both design and evaluate our hypotheses regarding the unique benefits of individual and collaborative learning on procedural and conceptual knowledge. This additional data stream can help elucidate patterns of interaction that may not be visible solely in the tutor logs or in the recorded audio, providing another avenue for understanding how collaboration helps learning, and how to best structure instruction to maximize these benefits.

References

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education, 19*(2), 105-154.

Anderson, J. R. & Gluck, K. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. M. Carver (Eds.) *Cognition & Instruction: Twenty-five years of progress*, 227-262. Erlbaum.

Beal, C. R., Walles, R., Arroyo, I., & Woolf, B. P. (2007). Online tutoring for math achievement: A controlled evaluation. *Journal of Interactive Online Learning, 6*, 43-55.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies, 70,* 377-398.

Graesser, A. C., Chipman, P., Haynes, B.C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education, 48,* 612-618.

Jermann, P., Mullins, D., Nüssli, M.-A., & Dillenbourg, P. (2011). Collaborative Gaze Footprints: Correlates of Interaction Quality. In CSCL '11: *Proceedings of the 2011 conference on Computer support for collaborative learning.* International Society of the Learning Sciences, 2011.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19,* 239-264.

Koedinger, K. R., Corbett, A. C., & Perfetti, C. (2012). The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36 (5),* 757-798.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449-521

Marwan, N, & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots, *Physics Letters A, 5-6,* 299–307.

McLaren, B.M., Adams, D., Durkin, K., Goguadze, G. Mayer, R.E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S., & Van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In A. Ravenscroft, S. Lindstaedt, C. Delgado Kloos, & D. Hernándex-Leo (Eds*.), Proceedings of EC-TEL 2012: Seventh European Conference on Technology Enhanced Learning,* LNCS 7563 (pp. 222-235). Springer, Berlin.

Mullins, D., Rummel, N. & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer Supported Collaborative Learning, 6*(3), 421-443, DOI: 10.1007/s11412-011-9122-z.

Nüssli, M.A. (2011). *Eye-tracking methods for the study of remote collaborative problem solving.* (Doctoral Dissertation). doi: 10.5075/epfl-thesis-5232.

Olsen, J.K., Belenky, D.M., Aleven, A., & Rummel, N. (in press). Intelligent Tutoring Systems for Collaborative Learning: Enhancements to Authoring Tools. Paper to appear in *Proceedings of the 2013 Conference on Artificial Intelligence in Education*

Rau, M. A., Aleven, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks: which dimension should we interleave? *Learning and Instruction, 23,* 98-114.

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,* 1045–1060.

Rittle-Johnson, B., & Alibali, M.W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*(1), 175-189.

Slavin, R. E. (1996). Research on Cooperative Learning and Achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43-69.

Walker, E., Rummel, N. & Koedinger, K. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer Supported Collaborative Learning.*

VanLehn, K., Lynch, C., Schultz, K., Shapiro, J. A., Shelby, R. H., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education, 15*(3), 147-204.

Acknowledgements

This research is funded by a grant from the Institute of Education Sciences, # R305A120734, and supported in part by Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education, #R305B090023.