



# Exploring Causality Within Collaborative Problem Solving Using Eye-Tracking

Kshitij Sharma<sup>1(✉)</sup>, Jennifer K. Olsen<sup>2,3(✉)</sup>, Vincent Alevén<sup>3</sup>,  
and Nikol Rummel<sup>3,4</sup>

<sup>1</sup> Norwegian University of Science and Technology, Trondheim, Norway  
`kshitij.sharma@ntnu.no`

<sup>2</sup> École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland  
`jennifer.olsen@epfl.ch`

<sup>3</sup> Carnegie Mellon University, Pittsburgh, PA, USA  
`aleven@cs.cmu.edu`

<sup>4</sup> Ruhr University Bochum, Bochum, Germany  
`nikol.rummel@rub.de`

**Abstract.** When students are working collaboratively and communicating verbally in a technology enhanced environment, the system is not aware of what collaboration is happening outside of the technology, making it difficult to adapt the system to better support the collaboration of the students. In this paper, we analyze the causal relationships between collaborative and individual gaze measures and the influence that the students dialogue, prior knowledge, or success has on these relationships to find indicators that can be used within an adaptive system. We found that when students are discussing concrete aspects of the problem, the causal relationship between their eye gaze measures changes compared to other types of dialogue patterns. The results also show a clear difference in causal relations when the pairs with high prior knowledge or success are compared with the pairs with low prior knowledge or success. Collaborative gaze causes the individual gaze for pairs with high prior knowledge and the opposite for the pairs with low prior knowledge.

**Keywords:** Collaboration · Dual eye tracking · ITS  
Granger causality

## 1 Introduction

In technology supported collaborative settings, students not only benefit from the support of the technology, but also from the exchange of ideas and explanations within their group. Currently, many technologies that are developed to support learning focus on the support of the domain material with support for collaboration being an afterthought, if explicitly supported at all, as is the case with individual Intelligent Tutoring Systems (ITSs). Additionally, in the classroom, students are often collaborating face-to-face and communicating verbally

with these verbal interactions occurring outside of the system making it difficult for the system to have a complete picture of the collaboration. For this paper, we are interested in supporting the collaborative interactions that occur between students as they work on a collaborative technology where all of the interactions may not be captured through the system.

Adaptive collaborative learning support (ACLS) can be used to adapt to the collaborative learning environment to provide appropriate support for the students by assessing student interactions, comparing them to a set of productive interactions, and providing interventions that will guide students closer to a productive interaction [49, 53]. Because verbal communication is still difficult to assess in real-time and students may not always be providing input to the learning technology. We propose using eye-tracking to assess student collaboration behaviors by investigating the different causal relationships of different process variables to find indicators that can be tracked and measured in real-time within a collaborative setting.

In this paper, we investigate the causal relationships between students' individual and collaborative gaze patterns (i.e., focus and similarity) for elementary school students working on a collaborative fractions ITS and examine how their dialogue plays a role in this relationship. For this analysis, we used time series data from the students working on the tutor. In the following sections, we will present an overview of the literature, study context, the analysis process, and causal inference results. These results provide insights into how eye-tracking measures can be used within collaborative learning environments to assess the level of collaboration and adapt to the current collaboration state. Specifically, we will address the following research question for this contribution: 1. What is the nature of causality between the collaborative and individual gaze patterns and 2. How do dialogue, prior knowledge and success alter this causality?

## 2 Related Work

Current implementations of ACLS often use either attributes of the student dialogue or interactions with the learning technology to assess the current collaboration state of the group. Many previous ACLS systems have used shallow indicators from dialogue to support student collaborations such as the number of student utterances [17, 41], used sentence openers [5, 34], or tracked particular sequences of dialogue actions (e.g., use of a question mark or dialogue talk moves) [1]. Often this analysis has been done on students who are communicating through chat where the features are easier to extract. Additionally, by including features of the learning environment in the assessment of the collaboration, such as the classification of the dialogue in relation to the actions the students are taking in the learning environment, often the intervention can be more impactful [33, 52, 54]. ACLS systems have also used interactions in the learning technology to gauge the collaboration, such as the request of hints and error patterns [54]. However, these interactions are not as useful in understanding what is happening outside of the system if there are long pauses between interactions when students

may be having discussions. Eye-tracking may be able to be used to make this link between the information provided in the learning technology and the group discussions that occur outside of them.

Eye-tracking may be a promising method to use to assess student collaboration as research has shown that eye gaze is tied to communication [35]. Previous research has shown a link between speech and eye gaze when people are working together on a task. There is a coupling of the collaborators' eye gaze around a reference [40], meaning that the collaborators' gaze may fixate, at approximately the same point in time, at the object referenced in the dialogue, for example just before mentioning it and just after hearing about it. The eye gaze has a closer coupling when each of the collaborators has the same initial information and when collaborators can visually share important objects that they are referencing in speech [26, 40], suggesting that concrete references may have more of an impact on eye gaze compared to abstract references.

Over the past few years, eye-tracking has become a key source of process data in educational research. Research using eye-tracking covers a wide range of educational ecosystems. Eye-tracking has not only been used to understand the learning processes in various contexts [38, 39, 46], but it also has been used to provide students appropriate, real-time, and adaptive feedback on their learning processes [14, 45]. In terms of collaborative learning scenarios, eye-tracking has most often been used with collaborating partners dialogues. Research has shown that there is a time lag between looking at an object and referring to the same object (eye-voice span) [21] and a time lag between a speaker's reference and a listener's gaze on the referred object (voice-eye span) [3]. Additionally, in terms of dual eye gaze, there is a lag in the eye-eye (speakers eye listeners eye) span (i.e., the time difference between the moment a speaker looks at an object and the moment the listener looks at the same object) [40]. Most of the dual eye-tracking studies have shown that the amount of time that the collaborating partners spend while looking at the same objects at the same time (cross-recurrence) is predictive of several collaborative constructs (e.g., collaboration quality [26]; misunderstandings [11]; learning gains [42]). In this paper, we go beyond correlational links to explore where there may be causal links between eye gaze measures and how they change during different forms of dialogue.

In this contribution, we propose a shift from correlation to causality. We borrow methods from finance and environmental studies to understand the causal relation between the different gaze-based variables. The key idea is to use the "cause" to "forecast" the effect and prepare for "adaptation" in ITSs. This has been a traditional practice in finance and environmental studies to use the causality to forecast [10, 12, 22, 28] and to use forecasting for adaptation requirements [7, 8, 32, 55]. We propose to use the causal relationship between the individual and collaborative gaze patterns to be able to forecast the behavior and provide adaptive feedback in a proactive manner.

For understanding the behavioral relation between the individual and collaborative gaze, we will use the Granger causality [20], a method that has been used in a multitude of domains to understand the relationship between

observable variables. For example, neuro-science [16, 19], user-consumption [36], stock-market [24] and economics [27, 50]. We will also explore the nature of this causal relationship using co-variables such as: pairs' dialogue, their prior knowledge and success levels.

In our work, we used a fractions ITS as a platform for our research. ITSs have been shown to be beneficial for student learning [30, 31] and are effective by providing cognitive support for students as they work through problem-solving activities. This cognitive support comes in the form of step-level guidance, namely, an interface that makes all steps visible, error feedback, and on-demand hints, which allow the system to adapt to the students current level of knowledge [51]. The cognitive support provided through the system can provide support for the student learning of the domain but does not provide support for the student collaboration when they are working in groups.

### 3 Methods

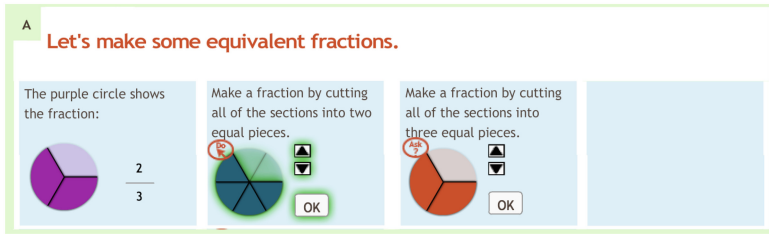
#### 3.1 Experimental Design and Procedure

Our data set involves 14 4th and 14 5th grade dyads from a larger study that investigated the benefits of collaborative versus individual learning [6, 37]. Each teacher paired the students participating in the study based on students who would work well together and had similar, but not equivalent, math abilities. The dyads were engaged in a problem-solving activity using a networked collaborative ITS, which allowed them to synchronously work in a shared problem space where they could see each others actions while sitting at their own computers. The students were able to communicate verbally through a Skype connection. Each dyad worked with the tutor for 45 min in a pull-out study design at their school. The morning before working with the tutor and the morning after working with the tutor, students were given 25 min to complete a pretest or posttest individually on the computer to assess their learning. During the experiment, dual eye tracking data, dialogue data, and tutor log data in addition to the pretest and posttest measures were collected. We collected eye-tracking data using two SMI Red 250 Hz infrared eye-tracking cameras.

#### 3.2 Intelligent Tutoring System

During the study, the dyads engaged with an ITS oriented towards supporting the acquisition of knowledge about fraction equivalence. Within each problem, the tutor provided standard ITS support, such as prompts for steps (i.e., revealing steps sequentially), next-step hints, and step-level feedback (i.e., correct or incorrect feedback) that allows the problem to adapt to the students problem-solving strategy [51]. Each of these different supports were displayed as actions on the screen that could guide the students actions and gaze.

For the collaboration, the ITS support mentioned above was combined with embedded collaboration scripts, which allowed students to take slightly different actions and see different information. The embedded collaboration scripts



**Fig. 1.** Example of a fractions interface showing incremental step reveals, feedback, and hint requests. Students had roles assigned that were displayed through their icon.

included three theoretically proven types of collaboration support: roles, cognitive group awareness, and individual accountability. First, for many steps, the students were assigned roles [29]. In the tutors, on steps with roles, one student was responsible for entering the answer and the other was responsible for asking questions of their partner and providing help with the answer. The tutor indicated the current role for the students through the use of icons on the screen. A second way in collaboration was supported was by providing students with information their partner did not have that they were responsible for sharing for the problem to be completed causing individual accountability [48]. The final feature was cognitive group awareness, where knowledge that each student has in the group is made known to the group [25]. On steps where this feature was implemented, each student was given an opportunity to answer a question individually before the students were shown each others answers and asked to provide a consensus answer.

### 3.3 Variables

For our analysis, we investigated a combination of data streams from eye gaze measures, dialogue, and test scores. For our eye gaze measures, we used focus and similarity because these two variables have been used in the recent research work concerning collaborative eye-tracking [43,46,47] to combine and analyse gaze behaviour at individual and collaborative levels. We used dialogue abstract as it can indicate how grounded the speech of the students is to what is occurring on the problem. Finally, the pretest and posttest scores allowed us to understand the relation of the causality to student knowledge.

**Individual Focus.** This is computed in terms of the entropy of the gaze. To compute the entropy, we divided the screen in 50-by-50 pixels grid. We also divided the whole problem-solving session into 10 seconds time windows. We then computed the proportion of the time spent in each block in the spatial grid for each 10-second time window. This resulted in a series of 2-dimensional proportionality vectors. Finally, we computed the Shannon Entropy for each of the vectors. A low entropy value (the minimum possible value is zero) depicts that the student was looking at only a few elements on the screen, which we called focused gaze. On the other hand, a high value of entropy indicates more

elements being looked at in a given time window, which we called unfocused gaze. Although focus and attention are related concepts, focus, as we defined here, does not contain the idea of processing the stimulus, as is required in the definition of attention. Attentive gaze indicates a certain level of processing of the sensory input. Focused gaze simply indicates a small number of elements looked over a fixed time period.

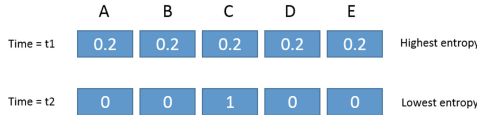


Fig. 2. Entropy computation

**Collaborative Gaze.** In order to compute the similarity between the gaze patterns of the collaborating students, we divided the screen space and the interaction time in the same manner as we did for entropy computation. We computed the similarity between the two proportionality vectors by using the reverse function  $(1/(1+x))$  of the correlation matrix of the two vectors. A similarity value of one will show no similarity between the two gaze patterns during a given time window. On the other hand, a higher value of similarity will show that the two participants spent time looking at the similar set of object on the screen during the same time window. Gaze similarity is an alternative measure of gaze convergence, the only difference between gaze similarity and gaze convergence comes from the mathematical formulation.

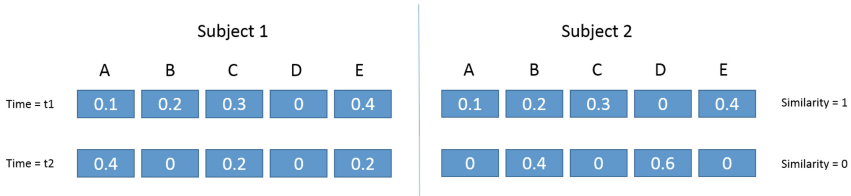


Fig. 3. A typical Similarity computation example

**Dialogue Abstraction.** Each of the student dialogues were transcribed and coded for abstraction levels. Abstraction is how grounded within the concrete aspects of the problem solving and communication the students utterance is. The level of abstraction is fully dependent on what occurs in the dialogue and is not intended to infer all mental processes. Within our transcripts, we coded for abstraction at the utterance level. This allowed us to have a fine-grained coding for each second of the dialogue without losing the context of the words. The abstraction codes consisted of five different levels: acknowledgement, read

out loud, interface, problem solving, and metacognitive (See examples below). The levels of abstraction followed an ordering with acknowledgments being the least abstract and metacognitive being the most abstract. For the coding, all statements that were off-task or were with a researcher were marked as “not applicable” and were discarded from the analysis. An inter-rater reliability analysis was performed to determine consistency among raters ( $Kappa = 0.78$ ).

1. Not applicable (NA): The student engages in off-task behavior, converses with the experimenter, or vocalizations without any context.
2. Acknowledgement (ACK): The student acknowledges their partner, or they request acknowledgment or a repeat of what the partner has said.
3. Read-out-loud (ROL): The student is reading information provided within the problem and presented on the screen.
4. Interface (INT): The student discusses actions that can be taken in the interface or engage in work coordination.
5. Problem solving (PRO): The student is providing an answer to the problem or showing evidence of think aloud as they solve the problem.
6. Metacognitive (META): The student verbally expressing their understanding of their current knowledge/problem solving state.

**Pretest and Posttest Scores.** To measure learning, we administered pretest and posttests to the students. The tests were computer-based and developed to closely align with the target knowledge covered in the tutors. The test comprised of 5 procedural and 6 conceptual test items. Two isomorphic sets of questions were developed, and there were no differences in performance on the test forms across all participants in the original study,  $t(79) = 0.96$ ,  $p = 0.34$ . The presentation of these forms as pretests and posttests was counterbalanced.

### 3.4 Data Analysis

We used Granger causality [20] test to examine the causality between the focus and similarity. The basic definition of Granger causality has two assumptions [20]. First, that cause occurs before effect and that the cause has information about the effect that is more important than the history of the effect. Although Granger causality is defined for linear and stationary time-series contexts, the variations for non-linear [4, 9, 18] and non-stationary [15, 23] contexts exist. The basic principle of Granger causality is to compare two models to test if  $x$  causes  $y$ . The first model predicts the value of  $y$  at time  $t$  using the previous  $n$  values of  $y$ . The second model predicts the value of  $y$  at time  $t$  using the previous  $n$  values of both  $x$  and  $y$ . Mathematically, following is a bivariate linear auto-regressive model for two variables  $x$  and  $y$ :

$$y(t) = \sum_{j=1}^p \alpha_{11j}x(t-j) + \sum_{j=1}^p \alpha_{12j}y(t-j) + \varepsilon_1(t) \quad (1)$$

$$x(t) = \sum_{j=1}^p \alpha_{21j} x(t-j) + \sum_{j=1}^p \alpha_{22j} y(t-j) + \varepsilon_2(t) \quad (2)$$

Where,

$p$  = model order, maximum lag included in the model

$\alpha$  = coefficients matrix, contribution of each lag value to the predicted value

$\varepsilon$  = residual, prediction error

We can conclude that  $x$  *granger-causes*  $y$  if coefficients in  $\alpha_{12}$  are jointly significant from zero. Statistically, this can be tested using F-test with the null hypothesis  $\alpha_{12} = 0$ . Also, the value of  $p$  can be decided based on the AIC [2] or BIC [44] model estimation values.

## 4 Results

In this section, we will provide the different analyses to arrive at a causal relationship between the variables mentioned in the Sect. 3.3. First, we would give an example about how to determine the granger causality between two variables to make the method explained in the Sect. 3.4.

Let us take the case of “focus” (the probability that both the participants have low gaze entropy) and “similarity” (the extent to which the peers looked at a similar set of objects in the a given time window). Table 1, comparison 1 shows the granger causality results for the overall data. The order of the model (Table 1, column 2) denotes how much lag was used to compute the causal relationship ( $p$  in Eqs. 1 and 2). In the case of Table 1, comparison 1, the lags used are 4 time windows (each time window corresponds to 10 s). To check if similarity granger causes focus, we create two models given by Eqs. 1 and 2 and compare them using F-test. The F and p values denote the effect size and significance of the model (Table 1, columns 3 and 4, respectively). We repeat the same process for checking if focus granger causes similarity. As we can see in Table 1 comparison 1, that “similarity granger causes focus” have a higher F (2.51) and lower (and significant) p value (.03) than “similarity granger causes focus” (F = 2.04, p = .09). Thus, we can conclude that “similarity granger causes focus”.

The remainder of this section presents the main results for this contribution. We observe that *similarity Granger causes focus* during the whole interaction (Table 1, Comparison 1). This causality also holds up when the dyads are engaged in a dialogue (Table 1, Comparison 2). Considering the data from the individual dialogue categories, The same causality holds when the peers are talking about interface issues (INTF, Table 1, Comparison 3). However, the causality changes the polarity (that is *focus Granger causes similarity*) while the peers are talking about problem solving (Table 1, Comparison 4); And there is no conclusive causality for ACK and META.



**Table 1.** The Granger causality model, across different data types, for collaborative similarity and probability that both participants have high focus. The direction of causality is denoted with a \*.

| Model   | Order | F-value | p-value |
|---|-------|---------|---------|
| Overall Data (1)  |       |         |         |
| Focus <- Similarity   | 4     | 2.51    | .03*    |
| Similarity <- Focus   | 4     | 2.04    | .09     |
| Participants Engaged in Dialogue (2)                          |       |         |         |
| Focus <- Similarity   | 8     | 2.12    | .03*    |
| Similarity <- Focus   | 8     | 0.93    | .47     |
| Participants Engaged in Dialogue w/INTF Abstraction (3)       |       |         |         |
| Focus <- Similarity   | 6     | 2.83    | .009*   |
| Similarity <- Focus   | 6     | 1.01    | .41     |
| Participants Engaged in Dialogue w/PRO Abstraction (4)        |       |         |         |
| Focus <- Similarity   | 5     | 0.21    | .95     |
| Similarity <- Focus   | 5     | 2.52    | .02*    |
| Dyads with High Average Posttest Scores (5)                   |       |         |         |
| Focus <- Similarity   | 2     | 3.91    | .02*    |
| Similarity <- Focus   | 2     | 1.70    | .18     |
| Dyads with Low Average Posttest Scores (6)                    |       |         |         |
| Focus <- Similarity   | 3     | 7.04    | .00001* |
| Similarity <- Focus   | 3     | 2.04    | .11     |
| Dyads with High Average Posttest Scores w/PRO Abstraction (7) |       |         |         |
| Focus <- Similarity   | 2     | 2.81    | .05*    |
| Similarity <- Focus   | 2     | 1.01    | .31     |
| Dyads with Low Average Posttest Scores w/PRO Abstraction (8)  |       |         |         |
| Focus <- Similarity   | 3     | 0.54    | .44     |
| Similarity <- Focus   | 3     | 2.74    | .05*    |
| Dyads with High Average Pretest Scores (9)                    |       |         |         |
| Focus <- Similarity   | 3     | 6.49    | .0002*  |
| Similarity <- Focus   | 3     | .04     | .98     |
| Dyads with Low Average Pretest Scores (10)                    |       |         |         |
| Focus <- Similarity   | 3     | 0.11    | .95     |
| Similarity <- Focus   | 3     | 4.42    | .004*   |

However, when we divide the data into pairs with high and low average posttest scores, we observe a few different relations. For the pairs with high posttest average *similarity Granger causes focus* (Table 1, Comparison 5) This polarity does not change for “PRO” abstraction (Table 1, Comparison 7). For the pairs with low

posttest average *focus Granger causes similarity* (Table 1, Comparison 6) and the polarity changes for “PRO” abstraction (Table 1, Comparison 8).

This result shows that there is some kind of interaction between the focus, similarity and performance. There is also an interaction between the focus, similarity and dialogue. Finally, we considered the relation between the pre and the post test scores. There is a positive significant correlation between the average pretest and the posttest scores for the pairs ( $r(27) = 0.57, p = .001$ ), indicating that prior knowledge also contributes in the success. Therefore, we divided the dataset into dyads with low and high average pretest scores and found that *similarity granger causes focus* for the pairs with high average pretest scores (Table 1, Comparison 9); whereas, *focus granger causes similarity* for the pairs with low average pretest scores (Table 1, Comparison 10).

## 5 Discussion and Conclusions

Granger causality is useful for forecasting the caused variable. In this paper, we examined the causal relation between individual and collaborative gaze-patterns, and used the dialogue, pretest and posttest scores as co-variables to explain the observed causality in detail. By understanding the causality, we can better use these measures to assess the collaborative state of students and develop interventions to guide the collaborative process.

In our analysis, we found that overall the collaborative similarity is causing the individual focus. This causality switches, that is individual gaze causes collaborative gaze, when the pairs are talking about “how to solve the problem?” One plausible explanation for this is that when two peers are talking about ways to solve problems, they both are individually focused on the problem description areas and hence start looking at the same section of the screen. Moreover, there is no conclusive causality during the episodes when the peers are in “ACK” or “META” abstraction. This may be explained by the fact that there is no need for the stimulus support when acknowledging a partner’s dialogue or a requirement to reflect upon a peer’s own state of understanding.

The key difference between the two causalities “looking at the same place hence focused” and “focused hence looking at the same place” might explain the fact whether collaboration is driving the individual gaze or the other way. In the case of successful pairs the collaboration seems to drive the individual behaviour, while in the case of unsuccessful pairs the relationship seems reversed. The same difference is there for the pairs with high and low prior knowledge. That is “similarity causes focus” for the pairs with high prior knowledge and “focus causes similarity” for the pairs with low prior knowledge. This difference could be a guiding factor about “how to provide adaptive feedback to the students?”

Additionally, the different causal relations for pairs with different levels of prior knowledge and success show that collaborative gaze causing the individual gaze is indicative of a “top-down” approach while individual gaze causing the collaborative gaze points to a “bottom-up” approach. Having coordinated gaze is a result of deeper socio-cognitive mechanisms [26, 40, 42, 43, 46] than just looking at a few elements on the screen (high focused gaze, by definition). In this

way, one can hypothesize that individual focus is similar to gaze reacting to the stimulus (screen or partner's dialogue) that is bottom-up behaviour [13]. On the other hand, the coordinated gaze is similar to cognition-driven gaze (referential gestures or familiarity with the interface or prior knowledge) that is top-down behaviour [13]. Our results show that examining the causality between collaborative and individual gaze patterns can unveil intriguing cognitive mechanisms underlying the collaborative learning with tutoring systems.

By forecasting the focus of the peers, we can take suitable actions for keeping the focus size for students in check. Using our results, when the focus size is large, given the similarity of the students, we can provide appropriate gaze-aware cues to the students, which would increase their similarity. From our results, this increase in similarity should increase the student focus, which can lead to more effective collaboration.

Additionally, we can provide feedback to the students based upon their eye gaze patterns. For example, whenever we detect that the focus is causing similarity, which tells us that they are not talking "PRO" then we can provide prompts to the students to guide their discussion back to the problem. We can test the impact of the prompts if we see that the *similarity is causing focus*, indicating the students' dialogue is discussing the problem.

Another opportunity for the personalized and adaptive feedback arises from the different causal relations based on the prior knowledge of the pairs. We found that for the pairs with high prior knowledge (high average pretest score) similarity causes focus, while for the pairs with low prior knowledge (low average pretest score) this is the focus that causes similarity. For such pairs (low prior knowledge), one can start giving feedback about where the partner is looking at, from the beginning of the session so that the high levels of similarity could be initiated and maintained throughout the collaboration and hence high levels of individual focus.

This work contributes to adaptive learning by revealing causality relations between individual and collaborative eye gaze measures that can be used to assess the collaboration of a group so that interventions can be applied at the correct moments. In future work, we would like to both extend our analysis to account for how features of the tutoring environment impact the findings as well as apply our findings to an adaptive environment to investigate if an adaptive system developed using these indicators is effective. A limitation of our work is that we had a small sample size and this may have impacted the results, which should be addressed in future work. Overall, our results indicate that student dialogue can impact the eye gaze relations as well as student prior knowledge. Understanding these relations allow us to adapt the system to better support student collaboration.

## References

1. Adamson, D., Rosé, C.P.: Coordinating multi-dimensional support in collaborative conversational agents. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 346–351. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-30950-2\\_45](https://doi.org/10.1007/978-3-642-30950-2_45)
2. Akaike, H.: A new look at the statistical model identification. *IEEE Trans. Autom. control* **19**(6), 716–723 (1974)
3. Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K.: Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *J. Memory Lang.* **38**(4), 419–439 (1998)
4. Ancona, N., Marinazzo, D., Stramaglia, S.: Radial basis function approach to non-linear granger causality of time series. *Phys. Rev. E* **70**(5), 056221 (2004)
5. Baker, M., Lund, K.: Promoting reflective interactions in a CSCL environment. *J. Comput. Assist. Learn.* **13**(3), 175–193 (1997)
6. Belenky, D., Ringenberg, M., Olsen, J., Alevan, V., Rummel, N.: Using dual eye-tracking to evaluate students' collaboration with an intelligent tutoring system for elementary-level fractions. Grantee Submission (2014)
7. Bertolli, C., Buono, D., Mencagli, G., Vanneschi, M.: Expressing adaptivity and context awareness in the assistant programming model. In: International Conference on Autonomic Computing and Communications Systems, pp. 32–47 (2009)
8. Challinor, A.: Towards the development of adaptation options using climate and crop yield forecasting at seasonal to multi-decadal timescales. *Environ. Sci. Policy* **12**(4), 453–465 (2009)
9. Chen, Y., Rangarajan, G., Feng, J., Ding, M.: Analyzing multiple nonlinear time series with extended granger causality. *Phys. Lett. A* **324**(1), 26–35 (2004)
10. Cheng, C.H., Wei, L.Y., Chen, Y.S.: Fusion anfis models based on multi-stock volatility causality for taieX forecasting. *Neurocomputing* **72**(16–18), 3462–3468 (2009)
11. Cherubini, M., Nüssli, M.A., Dillenbourg, P.: Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In: Proceedings of the 2008 symposium on Eye tracking research & applications, pp. 173–180. ACM (2008)
12. Clements, M.P., Hendry, D.F.: An overview of economic forecasting. A companion to economic forecasting pp. 1–18 (2002)
13. Connor, C.E., Egeth, H.E., Yantis, S.: Visual attention: bottom-up versus top-down. *Curr. Biol.* **14**(19), R850–R852 (2004)
14. D'Angelo, S., Begel, A.: Improving communication between pair programmers using shared gaze awareness. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 6245–6290. ACM (2017)
15. Ding, M., Bressler, S.L., Yang, W., Liang, H.: Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biol. Cybern.* **83**(1), 35–45 (2000)
16. Ding, M., Chen, Y., Bressler, S.L.: 17 granger causality: basic theory and application to neuroscience. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, **437** (2006)

17. Dowell, N.M., Cade, W.L., Tausczik, Y., Pennebaker, J., Graesser, A.C.: What works: creating adaptive and intelligent systems for collaborative learning support. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) ITS 2014. LNCS, vol. 8474, pp. 124–133. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-07221-0\\_15](https://doi.org/10.1007/978-3-319-07221-0_15)
18. Freiwald, W.A., et al.: Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. *J. Neurosci. Methods* **94**(1), 105–119 (1999)
19. Goebel, R., Roebroeck, A., Kim, D.S., Formisano, E.: Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magn. Reson. Imaging* **21**(10), 1251–1261 (2003)
20. Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: J. Econom. Soc.* 424–438 (1969)
21. Griffin, Z.M., Bock, K.: What the eyes say about speaking. *Psychol. Sci.* **11**(4), 274–279 (2000)
22. Hafner, C.M.: Causality and forecasting in temporally aggregated multivariate garch processes. *Econom. J.* **12**(1), 127–146 (2009)
23. Hesse, W., Möller, E., Arnold, M., Schack, B.: The use of time-variant EEG granger causality for inspecting directed interdependencies of neural assemblies. *J. Neurosci. Methods* **124**(1), 27–44 (2003)
24. Hiemstra, C., Jones, J.D.: Testing for linear and nonlinear granger causality in the stock price-volume relation. *J. Financ.* **49**(5), 1639–1664 (1994)
25. Janssen, J., Bodemer, D.: Coordinated computer-supported collaborative learning: awareness and awareness tools. *Educ. Psychol.* **48**(1), 40–55 (2013)
26. Jermann, P., Nüssli, M.A.: Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, pp. 1125–1134. ACM (2012)
27. Joerding, W.: Economic growth and defense spending: granger causality. *J. Dev. Econ.* **21**(1), 35–40 (1986)
28. Kavussanos, M.G., Nomikos, N.K.: Price discovery, causality and forecasting in the freight futures market. *Rev. Deriv. Res.* **6**(3), 203–230 (2003)
29. King, A.: Discourse patterns for mediating peer learning (1999)
30. Kulik, J.A., Fletcher, J.: Effectiveness of intelligent tutoring systems: a meta-analytic review. *Rev. Educ. Res.* **86**(1), 42–78 (2016)
31. Ma, W., Adesope, O.O., Nesbit, J.C., Liu, Q.: Intelligent tutoring systems and learning outcomes: a meta-analysis. *J. Educ. Psychol.* **106**(4), 901 (2014)
32. Manevitz, L., Bitar, A., Givoli, D.: Neural network time series forecasting of finite-element mesh adaptation. *Neurocomputing* **63**, 447–463 (2005)
33. McLaren, B.M., Scheuer, O., Mikšátko, J.: Supporting collaborative learning and e-discussions using artificial intelligence techniques. *Int. J. Artif. Intell. Educ.* **20**(1), 1–46 (2010)
34. McManus, M.M., Aiken, R.M.: Supporting effective collaboration: using a rearview mirror to look forward. *Int. J. Artif. Intell. Educ.* **26**(1), 365–377 (2016)
35. Meyer, A.S., Sleiderink, A.M., Levelt, W.J.: Viewing and naming objects: eye movements during noun phrase production. *Cognition* **66**(2), B25–B33 (1998)
36. Narayan, P.K., Smyth, R.: Electricity consumption, employment and real income in australia evidence from multivariate granger causality tests. *Energy policy* **33**(9), 1109–1116 (2005)

37. Olsen, J.K., Belenky, D.M., Alevan, V., Rummel, N.: Using an intelligent tutoring system to support collaborative as well as individual learning. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) ITS 2014. LNCS, vol. 8474, pp. 134–143. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-07221-0\\_16](https://doi.org/10.1007/978-3-319-07221-0_16)
38. Prieto, L.P., Sharma, K., Dillenbourg, P.: Studying teacher orchestration load in technology-enhanced classrooms. In: Conole, G., Klobučar, T., Rensing, C., Konert, J., Lavoué, É. (eds.) EC-TEL 2015. LNCS, vol. 9307, pp. 268–281. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24258-3\\_20](https://doi.org/10.1007/978-3-319-24258-3_20)
39. Raca, M., Dillenbourg, P.: System for assessing classroom attention. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, pp. 265–269. ACM (2013)
40. Richardson, D.C., Dale, R., Kirkham, N.Z.: The art of conversation is coordination. *Psychol. Sci.* **18**(5), 407–413 (2007)
41. Rosatelli, M.C., Self, J.A.: A collaborative case study system for distance learning. *Int. J. Artif. Intell. Educ.* **14**(1), 97–125 (2004)
42. Sangin, M., Molinari, G., Nüssli, M.A., Dillenbourg, P.: Facilitating peer knowledge modeling: effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Comput. Hum. Behav.* **27**(3), 1059–1067 (2011)
43. Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., Pea, R.: Using mobile eye-trackers to unpack the perceptual benefits of a tangible user interface for collaborative learning. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **23**(6), 39 (2016)
44. Schwarz, G., et al.: Estimating the dimension of a model. *Ann. Stat.* **6**(2), 461–464 (1978)
45. Sharma, K., Alavi, H.S., Jermann, P., Dillenbourg, P.: A gaze-based learning analytics model: in-video visual feedback to improve learner’s attention in MOOCs. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, pp. 417–421. ACM (2016)
46. Sharma, K., Caballero, D., Verma, H., Jermann, P., Dillenbourg, P.: Looking at versus looking through: a dual eye-tracking study in MOOC context. In: International Society of the Learning Sciences, Inc. [ISLS] (2015)
47. Sharma, K., Jermann, P., Nüssli, M.A., Dillenbourg, P.: Understanding collaborative program comprehension: Interlacing gaze and dialogues. In: Proceedings of Computer Supported Collaborative Learning (CSCL 2013), vol. 1, pp. 430–437 (2013)
48. Slavin, R.E.: Research on cooperative learning and achievement: what we know, what we need to know. *Contemp. Educ. Psychol.* **21**(1), 43–69 (1996)
49. Soller, A., Martínez, A., Jermann, P., Muehlenbrock, M.: From mirroring to guiding: a review of state of the art technology for supporting collaborative learning. *Int. J. Artif. Intell. Educ.* **15**(4), 261–290 (2005)
50. Thornton, D.L., Batten, D.S.: Lag-length selection and tests of granger causality between money and income. *J. Money Credit Banking* **17**(2), 164–178 (1985)
51. Vanlehn, K.: The behavior of tutoring systems. *Int. J. Artif. Intell. Educ.* **16**(3), 227–265 (2006)
52. Viswanathan, S.A., VanLehn, K.: High accuracy detection of collaboration from log data and superficial speech features. International Society of the Learning Sciences, Philadelphia (2017)
53. Walker, E., Rummel, N., Koedinger, K.R.: Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *Int. J. Comput.-Support. Collab. Learn.* **6**(2), 279–306 (2011)

54. Walker, E., Rummel, N., Koedinger, K.R.: Adaptive intelligent support to improve peer tutoring in algebra. *Int. J. Artif. Intell. Educ.* **24**(1), 33–61 (2014)
55. Wilby, R.: Decadal climate forecasting techniques for adaptation and development planning. Report for DfID (2007)