Through the (Thin-Slice) Looking Glass: An Initial Look at Rapport and Co-Construction Within Peer Collaboration

Jennifer K. Olsen, Human-Computer Interaction Institute, Carnegie Mellon University, jkolsen@cs.cmu.edu
Samantha Finkelstein, Human-Computer Interaction Institute, Carnegie Mellon University, slfink@cs.cmu.edu

Abstract: Within peer collaboration, both cognitive and social phenomena have been identified as important components for success, though little is known about the relationship between these factors. In this work, we examined math collaboration discourse between 11 4th grade dyads in 30-second slices to investigate the relationship between *rapport state* and *reasoning state*. Prior to collaboration, students watched one of three instructional videos modeling either domain knowledge, collaborate reasoning, or both. There was no impact of video type on student talk behaviors, nor posttest scores. However, we found a correlation between high rapport states and strong reasoning states, as well as a marginal effect of more co-constructive reasoning leading to improved posttest scores. This work demonstrates that students' rapport states may play a role in students' reasoning states, and thus calls for a deeper investigation within the CSCL community about the role of rapport in peer learning.

Introduction

Today's classrooms recognize the role of *talk* as not just a medium for conveying ideas, but also a process that creates *new* knowledge through the sharing of these ideas (Lee, Quinn, & Valdes, 2013). Peer collaboration can result in notable learning benefits for students (Lou, Abrami, & d'Apollonia, 2001), though prior work demonstrates certain conditions must be in place for these effects to be realized (Kollar, Fischer, & Hesse, 2006). Numerous conditions for success have been proposed, including both cognitive factors, such as transactively constructing new ideas with a partner (Chi & Wylie, 2014), and social factors, such as friendship status (Azmitia & Montgomery, 1993). Although both cognitive and social factors have been found to have an impact on learning, the problem of how these separate phenomena relate to each other has still been largely unaddressed. One notable exception is recent work demonstrating that friend-status impacts how peer tutors help their partners solve problems, with friend-tutors doing more question asking, while stranger-tutors do more knowledge-telling (Madaio, Ogan, & Cassell, 2016). By understanding how social and cognitive factors vary together, designers of learning experiences gain the ability to use one set of factors to leverage the other. Similarly, it is important to understand how each of these student factors is impacted by the design of the learning environment. In this paper, we investigate two primary questions: (1) what is the relationship between students' rapport states and reasoning states throughout a session, and (2) how does the way task activities are modeled to students impact either of these factors and their relationship? We analyzed the math collaboration discourse of 11 4th grade student dyads in 30-second slices over the course of one thirty-minute session. Dyads were presented with one of three instructional videos that either modeled domain knowledge, strong peer collaboration, or both so we could assess the impact on student behavior (Rummel, Spada, & Hauser, 2009).

From a cognitive perspective, collaborative learning is productive due to the reasoning that occurs through dialogue and provides students with deeper conceptual understandings of a domain (Webb, 2013). In the interactive-constructive-active-passive (ICAP) framework, Chi and Whylie (2014) argue that some of the success of peer collaboration is due to the type of cognitive engagement that the students voluntarily express together during the interaction. In ICAP, there are four levels of cognitive engagement, ranging from passive engagement (e.g., back-channel, or agreeing without providing new information) to interactive engagement (e.g., building off of the partner's ideas.) The more time a dyad spends engaging in talk that is closer to interactive co-construction rather than passive engagement, the more students will learn. We use this framework, discussed in more detail in the methods section below, as the inspiration for the *reasoning states* scheme we used to annotate students' on-task talk.

There are additionally a myriad of social factors that have been posited to impact student performance. For example, students' perception of how much they *belong* in a learning environment may impact the level to which they participate, and how *well* they participate. For example, students are far less likely to engage in school-ratified science reasoning when they feel that the expected behaviors within the learning environment are counter to their personal identities (Brown, 2006). On the other hand, there is evidence that by helping students feel like they are part of a learning community, we may be able to positively impact both their identity as a learner and their willingness to learn domain information (Gee, 2000). This connection between personal identity and the learning environment may be one explanation as to why student learners perform better when working with friends rather than acquaintances (Azmitia & Montgomery, 1993). Similarly, the rapport between students as they collaborate can have an influence on student learning. As students build a relationship, the social behaviors that are engaged in during the learning process may need to gradually change (Ogan et al., 2012). Although both cognitive and social factors have been found to influence group learning, there is not much work into the relationship between these factors, nor how we may be able to leverage students' social states in the design of our technologies so that they could be used to improve student reasoning.

In collaboration settings, students do not spontaneously produce the sorts of talk associated with success without appropriate support (Fischer et al., 2006). Collaboration scripts, in which students are given specific instructions for what to say and what to work on, have been shown to successfully facilitate the collaborative process (Fischer et al., 2013). However, if the script is too restrictive for the student's current level of knowledge, it can lead to *over-scripting* (Dillenbourg, 2002). Additionally, scripting may be difficult to implement on a large-scale since it may need to be adapted to each new domain. To address these issues, researchers have examined the potential benefits of *modeling* collaboration (Rummel, Spada, & Hauser, 2009). Modeling allows students to observe the behavior of others successfully completing a task and then integrates these same behaviors into their own interactions through vicarious learning (Decker & Nathan, 1985). When compared against scripting, modeling has been shown to demonstrate higher positive learning gains for students (Rummel, Spada, & Hauser, 2009). Furthermore, vicarious learning through modeling has been associated with social effects – specifically with students feeling like they belong to a learning community (Stenning, 1999). Finally, while *epistemic* and *social scripts* have been compared for their impact on student learning behaviors (Weinberger et al., 2005), such studies have not been done with modeling.

This paper builds upon prior work to explore how *modeling type* impacts students' rapport and reasoning states, and how these behaviors impact subsequent student learning. We break down this goal into five hypotheses we aimed to address in our study presented below: (1) When students have more instances of high reasoning states co-occurring with high rapport states, they will demonstrate higher posttest learning gains, (2) Students will be less likely to demonstrate low reasoning states during high rapport instances, (3) Students will be more likely to demonstrate high reasoning states during higher rapport states, (4) Students who see one of the two models that include collaboration will more high rapport and high reasoning states than students who see a domain-only model, and (5) students who see models that include strong collaboration will demonstrate more learning gains than students who see a domain-only model.

Methods

Instructional videos and problem-solving environment

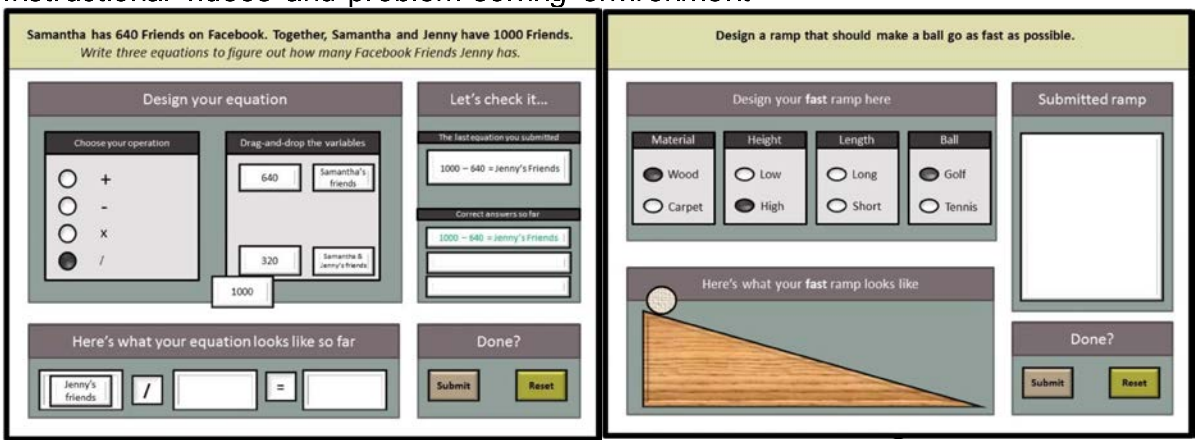

Figure 1. Before doing collaborative problem-solving, the students watched short videos of peers modeling the collaboration/problem-solving process in the domain (left) or out of the domain (right).

To model the collaboration for the students, we designed three different 4-minute vicarious learning videos for collaboration within the domain (Mixed), collaboration outside of the domain (Collaboration), and individual work within the domain (Domain). Each video was an animated version of a computer screen with students taking actions on the interface to solve the problem (see Figure 1) and audio of student(s) talking through the problem solving. For the videos that included modeling of the domain problem-solving skills (i.e., Domain and Mixed), students saw an interface that resembled the interface that was used in the intervention for the word problems. For the video that did not model domain skills (i.e., Collaboration), the students saw an interface for solving a ramp physics problem. Within the videos, students demonstrated productive talk around the domain material by demonstrating good problem-solving steps either through a student thinking out loud (Domain) or dialogue with a partner (Collaboration and Mixed). For the collaboration, students demonstrated four types of talk adapted from the academically-productive talk framework (Michaels, O'Connor, & Resnick, 2008), including explicit reasoning (e.g., "we could try x, because y"), eliciting partner reasoning (e.g., "wait, why?"), transactivity (building off of partner ideas), and unity (referring to the collaboration with 'we' words). For the domain-related videos, we kept the scripts as identical as possible. For example, instead of cognitive conflict and disagreement in the Mixed video, the Domain video showed the student self-correcting a mistake.

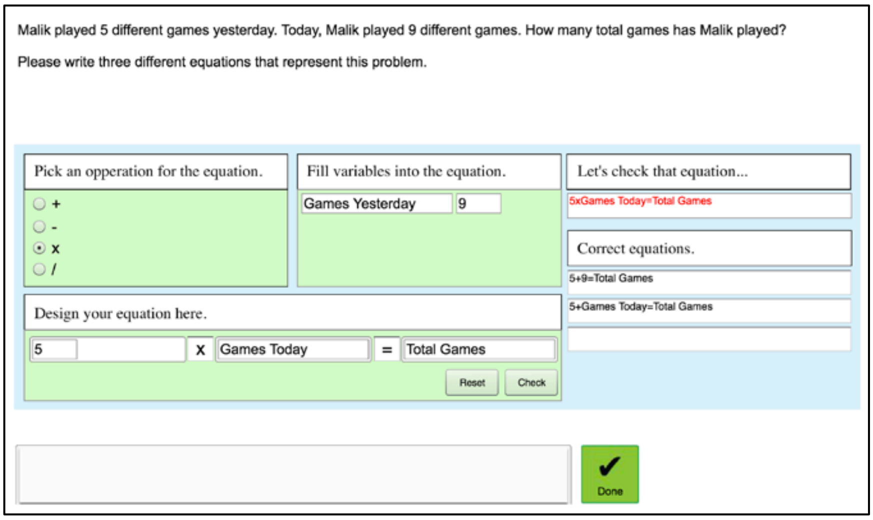

Figure 2. An example of the translation task for the word problem. On the bottom left of the screen the students can construct equations based on the word problem at the top. On the bottom right of the screen, the students are provided with feedback on their solution and can see their correct solutions.

After watching the video of modeled collaboration/domain skills as a pair, the students worked on a word translation problem set with their partner. We focused on this domain as it would be age-appropriate for 4th graders and offers multiple solution paths. To identify problems that were appropriately difficult, we piloted our interface with three 4th grade student dyads and iterated wording to promote comprehension as necessary. The problem set contained a total of nine problems, all of which were focused on translating word problems into one-step equations. Five word problems contained addition/subtraction operations, and four problems contained multiplication/division operations. Each word problem contained one unknown value, half with the total being the unknown value. For each problem, the students were asked to write three different equations that represented the word problem out of 16 possibilities. The system would accept any combination of given variable names (three options) and values (two options) with the correct operation that represented the problem; the equation did not have to be in an order that would allow the student to solve the problem. To prevent typing errors in the system, the students were provided with a drag-and-drop interface (see Figure 2) for the values and radio buttons to select the operators. The students received correctness feedback on the entire equation after pressing the 'Check" button. The open-ended nature of this domain allowed the students to have discussions around the solutions rather than just looking for the one correct answer (Webb 2013).

Experimental design and procedure

The study was conducted in a classroom setting with 22 4th grade students using a pull-out design. The students participated in the experiment during a free period in their regular school day for three days. Students completed pretest and posttest individually during 15-minute blocks on days 1 and 3, and on day 2, students worked with a partner on a math computer program that we designed for the experiment. We asked our partner teacher to pair students based on their perception of the students' friendship (i.e., could work together) and ability levels (i.e., do not have drastically different ability levels).

Before beginning the problem solving, students watched one of the three vicarious learning videos described above. Students were told that this was an example video "completed by students just like you" to "give them an example of what using this program might be like." Student dyads were randomly assigned to one of the three conditions: Domain, Collaborative, or Mixed (see Table 1). The students watched the videos in pairs as collaboratively watching a worked example video may result in students learning as much as one-on-one tutoring (Chi, Roy, & Hausmann, 2008). During the study, partners shared a single laptop but both the trackpad and an external mouse were provided allowing students to negotiate control of the interface. During the intervention, the students were video and audio recorded.

Table 1. Descriptions of the three different vicarious learning models used in the study

Domain	Collaboration	Mixed
One student performing a think-aloud around the math domain content targeted within this task. This student clearly explains the domain content through worked examples.	Two students demonstrating strong collaboration behaviors on a science task that is unrelated to the math domain. The students demonstrate interactive engagement through co-construction of ideas.	Two students demonstrating strong collaboration behaviors on the math domain content targeted within the task. This model shows both relevant math worked examples, as well as co-construction.

Hypotheses and dependent measures

We use this study design to address five primary hypotheses regarding the relationship between students' rapport states, reasoning states, and subsequent learning gains. We collected pretest and posttest measures on paper, and recorded students' dialogue as they worked through the intervention.

For analysis, files of each dyad dialogue were divided into 30-second segments for thin-sliced coding (Murphy, 2005). We chose thin-slice methodologies so we could rate students' talk holistically within a segment, rather than more traditional measures of annotation that mark individual utterances. Each segment was rated for two different types of behaviors that previous research has identified as being a potential mechanism explaining the relationship between collaboration and performance: reasoning state (Chi & Maneske, 2015) and rapport (Tickle-Degnen & Rosenthal, 1990). These annotations allowed us to address questions about how our vicarious learning conditions impacted dyad talk along these dimensions, as well as how students' talk behaviors impacted their learning results.

For reasoning state, a rating scale was developed based upon the ICAP framework (Chi & Whylie, 2014) with categories from zero to four (see Table 2). The rating scale represents an ordinal categorization of five different reasoning states that captures the four different *levels* of the ICAP framework (i.e., passive, active, constructive, interactive) and follows the same hierarchy that interactive talk is more effective than constructive talk, which is more effective than active talk, which is more effective than passive (Chi & Maneske, 2015). For our rating scale, an inter-rater reliability analysis using the Kappa statistic was performed to determine consistency among raters within the 30-second thin-slice audio clips (Kappa = 0.89).

Table 2: Reasoning states aligning with overt cognitive actions seen in the student's dialogue and ICAP

Rating	Overt Actions	ICAP Framework
0	Neither student is talking about the problem or the solution (or no talk is present at all). Talk around work coordination, or is off-task	Passive-Passive
1	Only one student is talking and either repeats information already provided or suggests a new solution	Passive-Active, Passive-Constructive
2	Both students are talking but are only repeating information already provided	Active-Active
3	Both students are talking and at least one new solution is provided, however, the solutions are not related	Active-Constructive, Constructive-Constructive
4	Giving and receiving of questions/answers, referencing a partner's previous solution in a new suggestion, co-construction of ideas	Interactive

The rapport rating scale used was adapted from a rapport thin-slice coding scale developed by Sinha & Cassell (2015). It captures aspects of mutual attention, coordination, and positivity within the peer collaboration context (Tickle-Degnen & Rosenthal, 1990). The rapport rating scale consists of five rating categories from zero to four (see Table 3). Similar to the reasoning state rating scale, the categorizations are ordinal. An inter-rater reliability analysis was performed to determine consistency among raters (Kappa = 0.69).

Although both rating scales are measuring types of talk that may play a role in collaboration, they are analyzing different aspects of the student behavior and are not dependent on one another. The reasoning state rating scale takes into consideration how the students present solutions to the problems and build upon each other's prior work while the rapport rating scale takes into consideration the social aspects of the student interactions and how well the students are engaging with their partner, regardless of the content. To ensure that these rating scales were independent in practice, we looked at cell distributions after annotation was complete to ensure that all cell combinations were present, which they were. In addition, to reduce the influence that one rating scale had on the category assignment from the other rating scale, after the inter-rater reliabilities were established, different researchers coded the reasoning state scale and the rapport rating scale.

Table 3: Rapport rating scale aligning with overt social actions seen in the student's dialogue

Rating	Overt Actions
0	Silence, or only talk to the experimenter or other students in the classroom, but not to each other. (These segments were considered categorically different than the rest of the ordinal scale, and not very low rapport. They were removed for some analyses presented below.)
1	Low rapport. This code is marked by lack of synchrony between students, low positivity, or lack of attention (e.g., "ugh, can you stop it? I'm working on this. I said stop.")
2	Neutral rapport. This code marked segments where students were demonstrating the 'bare minimum' for dyadic interaction to be successful (e.g, "Should we add here?" "yep." "okay.")
3	Positive rapport. Students' dialogue flowed smoothly, and students demonstrated some active interest in their partners' contributions (e.g., "Ooh, should we add here?" "okay yeah I thought so too!" "okay"
4	Very high rapport. These segments were marked by student social behaviors that appeared to proxy high levels of coordination. On the surface, these interactions may have been marked by strong levels of positivity (e.g., "we got it!" "yes!" we're soooo good at this!") or, conversely, positively-received teasing (e.g., "ooooh you think you're going to get this one now? No way." [partner laughs]).

Finally, we collected students' pretest and posttest data with a paper assessment using two equivalent, counterbalanced test forms. The tests contained a total of eight problems. Four of the problems were isomorphic to the intervention and four of the problems were transfer problems. For two of the transfer problems, the students were asked to solve a word problem (not just write an equation) and received one point for the correct answer. For the other two transfer problems and the isomorphic problems, the students were asked to write three different equations that represented the given word problem. An equation was counted as correct if it used either the numbers or variables presented in the problem with the correct operators to make an accurate translation of the problem. The isomorphic problems were one-operation equations while the transfer problems were two-operation equations. For each correct equation (no credit for solving), the students got one point for a possible three points for each word problem. On the tests there were 20 possible points for the 8 questions.

Results

Out of the 22 students in the study, 20 students were included in the analysis because of technical errors. There were four pairs assigned to the Mixed condition, three pairs assigned to the Collaborative condition, and three pairs assigned to the Domain condition. To check the distribution of knowledge across conditions, we compared student pretest scores and found no significant difference between conditions, $F(17,2) = 1.42, p = .27$.

What is the impact of vicarious learning model on students' talk behaviors?

We hypothesized that students who heard a model of strong collaborative talk (Mixed or Collaboration) would subsequently demonstrate higher rapport and more instances of strong collaborative talk than those who were shown a video that exclusively modeled strong domain reasoning. To investigate the impact that condition had on student dialogue during the collaboration, we conducted a MANOVA analysis. Using the Pillia's trace, there was not a significant effect of condition on students' reasoning states nor rapport states, $V = 0.23, F(4,14) = 0.46, p = .76$. This analysis demonstrated that there was no impact of vicarious learning model on students' rapport or collaborative talk as measured by our thin-slice annotations.

What is the impact of rapport-level on the likelihood of strong cognitive talk?

To test our hypotheses that (a) low rapport states would be *less likely* to have higher reasoning states, and (b) high levels of rapport will be *more likely* to have higher reasoning states, we conducted a Person's Chi-squared. Because there were not very many instances of category 2 for the cognitive rating scale (less than five), we combined categories 2 and 3 for a combined category that reflected when both students were talking but were not being interactive. In addition, we removed all segments with a '0' in the rapport scale as this marked utterances where students were not talking. This left us with four categories for both the cognitive rating scale (i.e., 0,1,3,4) and the rapport rating scale (i.e., 1,2,3,4) and a total of 540 segments.

We identified a significant association between the cognitive rating category of a talk segment and the rapport rating category of a talk segment, $\chi^2(9) = 41.11$, $p < .05$ (see Figure 3). Based on the standardized residuals of the different cells, we found that there were significantly more talk segments that had a 0 category for cognitive talk and a 1 category for rapport talk than would be expected ($p < .05$). In contrast, we found that there were significantly fewer talk segments that had a 4 category for cognitive talk and a 1 category for rapport talk than would be expected ($p < .05$). These results indicate that when students are having interactions marked by notably low rapport (such as lack of coordination or notable negativity), they are less likely to be co-generating interactive math reasoning. In addition, there was a marginally significant effect for more talk segments that had interactive talk (reasoning state rating = 4) and moderately high rapport (rapport state = 3) than would be expected ($p = .09$). These results indicate that while very high rapport doesn't co-occur with strong cognitive talk, better-than-average coordination indicated by a '3' on this scale may facilitate strong collaborative reasoning.

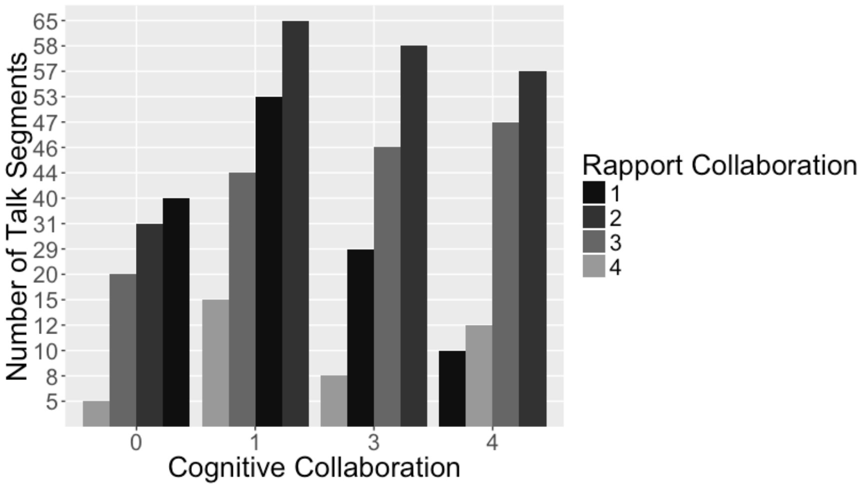

Figure 3. The distribution of reasoning state categories and rapport categories across talk segments.

What is the relationship between student talk annotations and vicarious learning model on students' learning gains?

To investigate the impact that the modeling conditions and annotated student talk variables had on learning gains, we used a multilevel approach to take into account the repeated measures of the pretest and posttest and differences between dyads. For this analysis, we were interested in a particular combination of annotated talk moves – specifically the co-occurrence of interactive cognitive talk (rating 4) with high rapport (segments rated as a 3 or 4). We hypothesized that more instances of this combination would lead to greater learning gains. To test this, we included the percentage of talk instances for each dyad where the students had a cognitive score of 4 and a rapport score of at least 3 into the analysis. This percentage ranged from 0 to 0.41 across pairs. We used a hierarchical linear model (HLM) with student at the first level and dyad at the second level. At level 1, we modeled the pretest and posttest scores; at level 2, we accounted for individual differences, condition, and the pair's percentage of high rapport/collaborative talk. Within our model, we chose pretest and the Collaborative modeling condition as the baselines. For each variable, the model includes a term for each comparison between the baseline and other factors of the variable. We measured the effect size with Pearson's correlation coefficient (r) where 0.1 is a small effect size, 0.3 is a medium effect size, and 0.5 is a large effect size.

Table 4: Means and (standard deviations) for the pretest and posttest scores across conditions

Condition	Pretest Score	Posttest Score
Collaborative Modeling Only	4.67 (3.78)	8.17 (4.11)
Domain Modeling Only	2.17 (2.14)	4.17 (2.48)
Collaborative and Domain Modeling	2.50 (2.45)	3.88 (2.95)

The results of the learning gains by condition are displayed in Table 4 and pretest and posttest analysis are shown in Figure 3. There was a significant difference between pretest and posttest scores, $t(17) = 2.76$, $p < .05$, $r = 0.56$, with the posttest scores being higher across all conditions. For the condition differences, there was a marginally significant difference between Collaborative and Mixed models, $t(6) = -2.12$, $p = .07$, $r = 0.65$, with the Collaborative condition having the higher test scores and a non-significant difference between Domain and Collaborative models, $t(6) = -1.73$, $p = .13$. There was not a significant interaction between pretest/posttest and Collaborative/Mixed conditions, $t(17) = -1.27$, $p = .22$, and pretest/posttest and Domain/Collaborative conditions, $t(17) = -0.84$, $p = .41$. Regarding the impact of the percentage of high rapport/collaborative talk, there was a marginally significant main effect, $t(6) = 2.27$, $p = .06$, $r = 0.68$, with dyads with more high rapport/collaborative talk co-occurrence segments having higher posttest scores.

Discussion and conclusion

In this work, we addressed two primary questions: (1) what is the relationship between students' rapport state and their corresponding reasoning state, and (2) how do vicarious learning models impact students' overall rapport and reasoning behaviors? By looking at how students' rapport states impact the likelihood of interactive reasoning, we gain a better understanding of how rapport plays a role in peer collaboration. Once this relationship is better understood, we may be able to design collaborative learning systems that productively leverage students' relationships during the learning process.

Our first result demonstrates that students' rapport state does in fact impact the likelihood of interactive reasoning. Low rapport (rating = 1) time slices are significantly less likely than expected to co-occur with interactive reasoning, and significantly more likely than expected to co-occur with talk the lowest reasoning states. In this work, time slices were deemed to have very low rapport (rating = 1) when an interaction seemed emotionally tense – for example, students did not seem aligned in their task goals, were attending in different directions, or were criticizing their partners' performance without positive response from the targeted student. We also found marginal significant results indicating that high (but not extremely high) rapport states (rating = 3) may be more likely than expected to co-occur with interactive reasoning, and that interactive reasoning was more likely to occur when rapport was *high* than when it was *neutral*. The results of this work demonstrate that students' relationship dynamic *at a given moment* is associated with the likelihood that high-quality co-constructive reasoning is also being produced. In other words, it may not be that *friends learn better than strangers,* which some prior work suggests (Azmitia & Montgomery, 1993), but that *friends may have fewer instances of low rapport than strangers,* and that this improved alignment results in more instances of idea co-construction when collaborating.

We additionally found that students demonstrated increased learning gains after receiving the vicarious learning model that focused on collaboration in a domain-separate task, rather than either a model that focused on domain-relevant information, or even collaboration within this domain-relevant task. This result is similar to the result by Weinberger et al. (2005), who found that students demonstrated greater learning gains after following a social script, rather than a domain script. Interestingly, the type of vicarious learning model did not impact students' average rapport-states over the course of the session, nor the overall number of time slices where interactive reasoning occurred. This indicates that while learning model type may have an impact on students' posttest scores, it did not have an impact on how the students spoke to each other while completing the task (as captured by the features that were annotated for in the data.) While we caveat these results due to both our small sample size and the marginal nature of the finding, we report it due to its large effect size ($r = .6$), and encourage other members of the research community to investigate the impacts of collaborative over domain modeling within their own work.

These analyses serve as a first investigation into the relationship between students' rapport states and the corresponding reasoning states by looking at co-occurrence. It demonstrates that *social relationship in a given moment* plays a role in the likelihood of co-constructive idea generation, and that ignoring rapport between dyad members would be ignoring part of the equation. Moving forward, we will investigate more complex patterns in students' rapport states and how they may predict *future* reasoning states. We will also aim to investigate how different *patterns of reasoning* over the course of a session may predict learning gains in ways that may be more sophisticated than the "more is better" hypothesis that was tested in this work. Like learning itself, human relationships are complex, and uncovering the patterns of rapport that are predictive of success will give the CSCL and education communities a greater understanding of how to best leverage this often ignored phenomenon in students' learning experiences.

References

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social development*, *2*(3), 202-221.

Brown, Bryan A. "“It isn't no slang that can be said about this stuff”: Language, identity, and appropriating science discourse." *Journal of Research in Science Teaching* 43.1 (2006): 96-126.

Chi, M. T. H., & Menekse, M. (2015). Dialogue patterns that promote learning. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), Socializing Intelligence Through Academic Talk and Dialogue (Chi. 21, pp. 263-274). Washington, DC: AERA.

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. Cognitive science, 32(2), 301-341.

Chi, M. T. & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219-243.

Decker, P. J., & Nathan, B. R. (1985). *Behaviour Modeling Training: Principles and Applications*. Praeger Publishers.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three worlds of CSCL. Can we support CSCL?* 61-91.

Fischer, F., Kollar, I., Stegmann, K., Wecker, C., Zottman, J., & Weinberger, A. (2013). Collaboration Scripts in Computer Supported Collaborative Learning, in Hmelo-Silver, C., Chinn, C., Chan. C., & O'Donnell, A. (Eds) The International Handbook of Collaborative Learning, Routledge.

Gee, J. P. (2000). Identity as an analytic lens for research in education. *Review of research in education*, *25*, 99-125.

Joshi, M., & Rosé, C. P. (2007). Using transactivity in conversation for summarization of educational dialogue. In *SLaTE* (pp. 53-56).

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review*, *18*(2), 159-185.

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher*, 0013189X13480524.

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.

Madaio, M. A., Ogan, A., & Cassell, J. (2016, June). The Effect of Friendship and Tutoring Roles on Reciprocal Peer Tutoring Strategies. In International Conference on Intelligent Tutoring Systems (pp. 423-429). Springer International Publishing.

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in philosophy and education*, *27*(4), 283-297.

Murphy, N. A. (2005). Using thin slices for behavioral coding. *Journal of Nonverbal Behavior*, *29*(4), 235-246.

Ogan, A., Finkelstein, S., Walker, E., Carlson, R., & Cassell, J. (2012, June). Rudeness and rapport: Insults and learning gains in peer tutoring. In *International Conference on Intelligent Tutoring Systems* (pp. 11-21). Springer Berlin Heidelberg.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69-92.

Sinha, T., & Cassell, J. (2015). We click, we align, we learn: Impact of influence and convergence processes on student learning and rapport building. In *Proceedings of the 1st Workshop on Modeling INTERPERsonal SynchrONy And infLuence* (pp. 13-20). ACM.

Stenning, K., McKendree, J., Lee, J., Cox, R., Dineen, F., & Mayes, T. (1999, December). Vicarious learning from educational dialogue. In *Proceedings of the 1999 conference on Computer support for collaborative learning* (p. 43). International Society of the Learning Sciences.

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological inquiry*, *1*(4), 285-293.

Webb, N. (2013). Information processing approaches to collaborative learning. In Hmelo-Silver, C., Chinn, C., Chan. C., & O'Donnell, A. (Eds) The International Handbook of Collaborative Learning, Routledge (pp 19-40).

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer–supported collaborative learning. *Instructional Science*, *33*(1), 1-30.

Acknowledgments

We thank Nikol Rummel and Vincent Aleven for their guidance. This work was supported by Graduate Training Grant #R305B090023 from the US Department of Education (IES).

