Evaluating the Effect of Social Conversation on Learning, Interaction, and Perceived Interdependence in a Collaborative Math Problem Solving Environment

Rohit Kumar, Gahgene Gweon, Mahesh Joshi, Yue Cui, Adaeze Nwaigwe, Carolyn P. Rosé
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213,
{rohitk,gkg,maheshj,ycui,anwaigwe,cp3a}@andrew.cmu.edu

Abstract: In this paper, we explore the effect of social prompts offered by a computer supported collaborative learning (CSCL) environment on student attitudes and behavior towards each other. We do this by experimentally contrasting collaboration in two configurations, one that presents students with social questions in between math problems and then uses the answers to those problems to tailor the cover story and associated hints for the next problem, and a control condition that uses the same problems, but does not use social questions for tailoring. Our finding is that the social prompts changed the attitude students displayed towards one another. Whereas in the control condition students took on a more competitive attitude, in the experimental condition students were more playful. On the questionnaire students reported exchanging more help in the experimental condition. An analysis of the corpus reveals no significant differences in amount of help exchanged, but there was a significant difference in terms of non-help-related conversation such that the proportion of episodes where help was provided was higher in the Experimental condition. In terms of learning gains, the trend was consistently in favor of the experimental condition in terms of learning between the pretest and the quizzes after each lab session as well as the posttest, however these trends were not statistically significant overall. The strongest learning result is that students in the experimental condition learned marginally more on one unit of the material on the second lab day than students in the control condition (p=.06, effect size .55 standard deviations).

Introduction

We are in the beginning stages of an exploratory project, the goal of which is to enhance participation and learning in the Virtual Math Teams (VMT) online math service by designing, developing, implementing, testing, refining and deploying computer-supported tools to support facilitation in this lightly-staffed service. The key research goal in the long term is to optimize a design and implementation for dynamic feedback in support of collaborative problem solving that will maximize the pedagogical effectiveness of the collaboration by eliciting behavior that is productive for student learning in collaborative contexts, such as the Virtual Math Teams environment (Stahl, 2006). The study we report in this paper is one in a series of investigations into the design, implementation, and evaluation of conversational agents that play a supportive role in collaborative learning interactions (Gweon et al., 2006; Kumar et al., 2007; Wang et al., 2007). The ultimate goal of this long term endeavor is to support collaboration in a way that is responsive to what is happening in the collaboration rather than behaving in a "one size fits all" fashion, which is the case with state-of-the-art static forms of support such as including assignment of students to roles (Strijbos, 2004), provision of prompts during collaboration (Weinberger, 2003), design of structured interfaces including such things as buttons associated with typical "conversation openings" (Baker & Lund, 1997), instructions to guide learners to structure their collaboration (Webb & Frivar, 1999), or even various forms of collaboration training (Rummel et al., 2006). Our investigations thus far have been in lab and classroom studies. The far less controlled VMT environment provides a more challenging environment in which to test the generality and robustness of our prior findings, while at the same time providing a context where successful technology for supporting collaborative learning interactions can reach a broad spectrum of kids in need of support in their mathematics training.

In the VMT environment, collaboration is supported with a combination of scripting and human moderation. The script based structuring is stage-based. Students work in small groups on the same problem over 3 sessions. In the first session, they work out solutions to the problem. In between the first and second sessions, students receive feedback on their solutions from human moderators. In the second session, students discuss the feedback they received on their respective solutions and step carefully through alternative correct solutions. In that session and the subsequent session, they also discuss additional possible ways of looking at the problem including variations on that problem in order to take a step back and learn larger mathematics principles that apply to classes of problems rather than individual problems. Although the problem provides the opportunity to investigate multiple possible solutions and to engage in deep mathematical reasoning, VMT researchers have found from analysis of chat logs where students have worked together is that students tend to jump to finding one solution that works rather than taking the opportunity to search for alternative solutions. The moderator plays an important role in stimulating conversation between students, encouraging knowledge sharing and probing beyond a single acceptable solution.

Our long term goal is to use technology to support collaboration in this environment in two main ways, both of which leverage our prior research on automatic collaborative process analysis (Donmez et al., 2005; Wang et al., 2007b). The first approach is to deploy conversational agents to offer support by participating in the conversation as in (Gweon et al., 2006; Kumar et al., 2007; Wang et al., 2007), where in all cases we observed a positive learning effect from involvement of conversational agents in collaborative learning interactions in a chat environment. In this case, automatic collaborative learning process analysis is used to detect when a conversational agent should intervene in a conversation. The other way is to use the automatic analysis of the conversation to construct reports that inform human facilitators of which groups are most in need of support (Rosé et al., 2007; Joshi & Rosé, to appear). In this paper, we focus primarily on the first approach. Specifically, we report on an exploration of how the participation of chat agents in the collaborative learning interaction changes the nature of the interaction, and how that change in conversational processes affects how much students learn.

While there has been much work evaluating a wide range of conversational agents for supporting individual learning with technology (Kumar et al., 2006; VanLehn et al., 2007; etc.), a similar effort in collaborative contexts is just beginning. We have observed in our recent research that working collaboratively may change the way students conceptualize a learning task and similarly how they respond to feedback (Wang et al., 2007; Wang & Rosé, 2007). For example, Wang & Rosé found that students who worked in pairs approached an idea generation task more broadly when they worked in pairs rather than individuals, in particular behaving in a way that indicated more of a fluid boundary between tasks, whereas students who worked individually focused more narrowly on one task at a time. Correspondingly, students who worked in pairs with feedback showed even more evidence of a connection between tasks, where individuals with feedback during idea generation simply intensified their success within their original narrow focus. This difference in how students responded to feedback when they worked individually and in pairs tells us that before we will be able to effectively support collaborative learning with tutorial dialogue technology in particular as well as intelligent tutoring technology more generally requires re-evaluating established approaches

For decades a wide range of social and cognitive benefits have been extensively documented in connection with collaborative learning, which are mediated by conversational processes. Based on Piaget's foundational work (Piaget 1985), one can argue that a major cognitive benefit of collaborative learning is that when students bring differing perspectives to a problem solving situation, the interaction causes the participants to consider questions that might not have occurred to them otherwise. This stimulus could cause them to identify gaps in their understanding, which they would then be in a position to address. This type of cognitive conflict has the potential to lead to productive shifts in student understanding. Related to this notion, other cognitive benefits of collaborative learning focus on the benefits of engaging in teaching behaviors, especially deep explanation (Webb, Nemer, & Zunita 2002). Other work in the computer supported collaborative learning community demonstrates that interventions that enhance argumentative knowledge construction, in which students are encouraged to make their differences in opinion explicit in collaborative discussion, enhances the acquisition of multi-perspective knowledge (Fischer, et. al 2002). Furthermore, based on Vygotsky's seminal work (Vygotsky 1978), we know that when students who have different strengths and weaknesses work together, they can provide support for each other that allows them to solve problems that would be just beyond their reach if they were working alone. This makes it possible for them to participate in a wider range of hands-on learning experiences. Because of the importance of these conversational processes, in our evaluation of the design of conversational agents for supporting collaborative learning, we must consider both the learning that occurs when individuals interact with these agents in the midst of the collaboration (i.e., learning from interaction with the agents) with learning that is mediated by the effects of the agents on the interaction between the students. While in our previous recent studies we have focused on the first source of learning, in the study reported in this paper, we focus on learning from changes in conversational processes.

In the remainder of the paper, we begin by describing our current collaborative problem solving environment, which we eventually plan to replace with the VMT environment augmented with our dynamic collaborative learning support technology. We then discuss our hypotheses and experimental design. Finally, we discuss our findings and current directions.

Infrastructure and Materials for Supporting Collaborative Problem Solving

The study we report in this paper was a classroom study where students worked in their school computer lab in pairs using the collaborative problem solving environment displayed in Figure 1. In this section we discuss this experimental infrastructure, which was used to conduct our investigation. We will discuss this infrastructure both in terms of the technology we used and in how we set up the lab where the students worked.

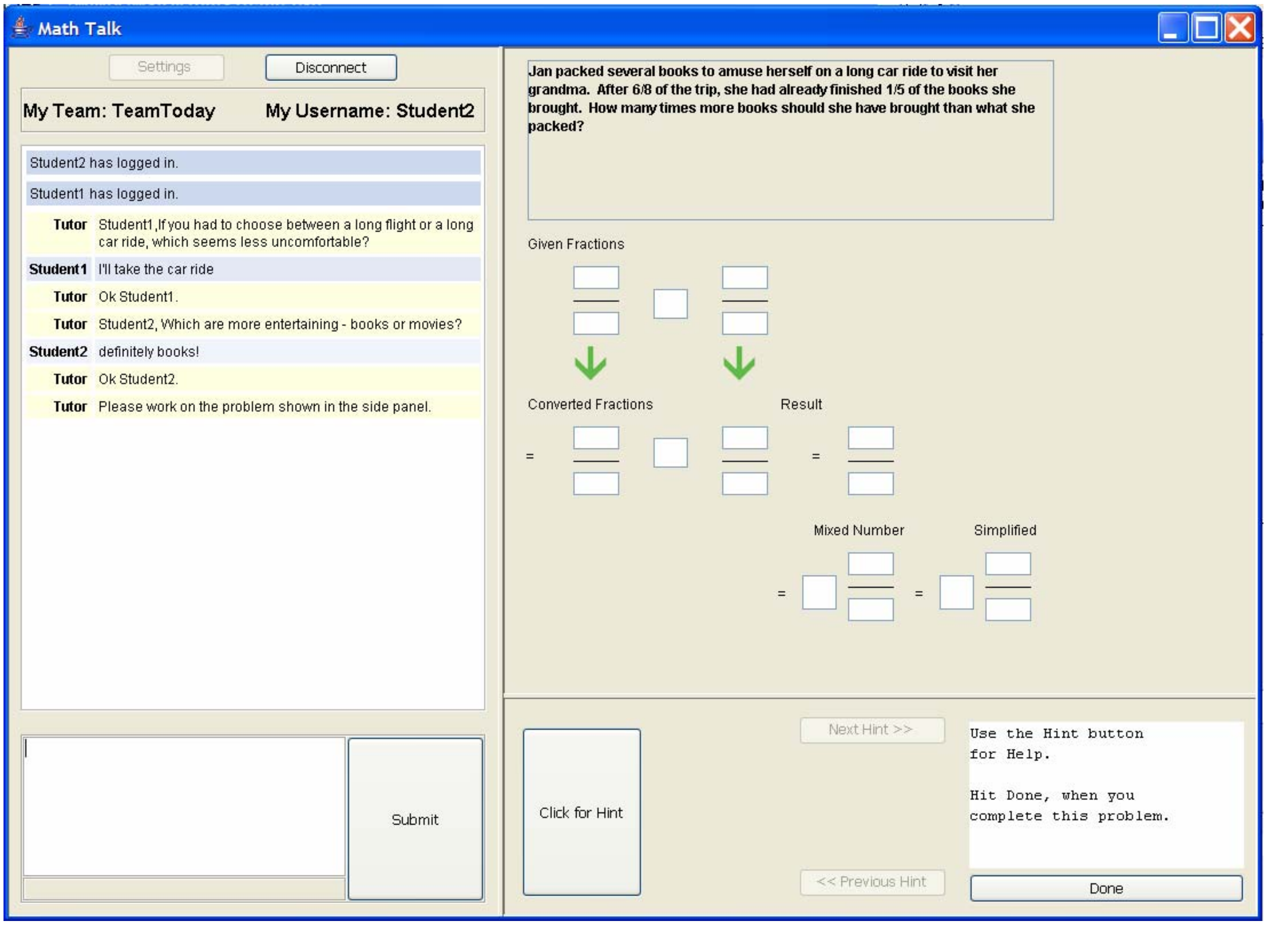

Figure 1. Collaborative Problem Solving Environment for Experimental Condition.

The interface in Figure 1 has two panels. On the left is a chat interface, which allows students to interact with each other as well as with the conversational agents that are triggered at different occasions during the problem solving session. The panel on the right is the problem solving interface that allows students to collaboratively work on a given problem. In this case the interface in the right panel was built using the Cognitive Tutor Authoring Tools (CTAT) [14]. The CTAT panel has a problem layout and a hint button. The hint button triggers support built into the CTAT environment. The hint messages are displayed in the Chat buffer. Both panels of the interface maintain a common state across both the participants at all times creating a shared experience for the students. All actions performed by a student in either of the panels are immediately communicated and reflected on the interface of the other student. This integrated shared experience of problem solving is unique to this interface in contrast to systems used in our earlier experiments where VNC was used to manage the shared problem solving interaction (Gweon et al., 2006; Gweon et al., 2007).

Figure 2 shows an overview of the architecture used to develop the infrastructure for this study. This architecture is principally similar to that used in our earlier work (Kumar et al., 2007; Wang et al., 2007). However the present implementation of this architecture allows for a richer set of communications that enable creation of the integrated shared problem solving experience. The filters module is responsible for managing the interaction. All interface events resulting from student contributions to the chat interface and to the structured problem solving interface are sent to the Filters module. Its purpose is to identify significant events in this stream that it then reflects back to the interfaces of both students. It also uses these identified events to update its internal state. Other triggers such as timers that keep track of time elapsed since the beginning of the session or since the last significant contribution of each student are also used to manipulate the Filter module's internal state. The internal state then is used to select strategies for selecting dialogue agents to participate in the chat session. In our prior experiments we have used different kinds of triggers including topic based filters, time-outs, interface actions, and conversational actions that are indicative of the degree of engagement of the students in the discussion. Some of these event identifiers rely on functionality provided by the TagHelper tools verbal protocol analysis toolkit[1] (Donmez et al., 2005; Wang et al., 2007b). Our generic architecture is meant to be easily extended to work with other types of triggers such as cues from other modalities like speech, eye-gaze, etc. We continue to improve the architecture to provide richer communication and modularization.

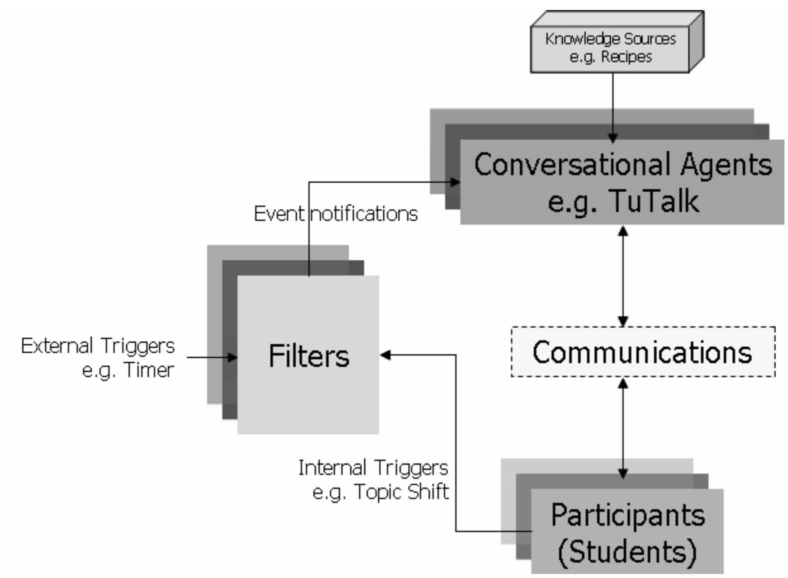

Figure 2. Architecture underlying the Collaborative problem solving interface with Conversational Agents

We employ two types of conversational agents for this experiment: simple social dialogue agents and cognitive support gents implemented with the TuTalk authoring environment (Gweon et al., 2005; Jordan et al., 2007). The social dialogue agents were designed to show a personal interest in the students by asking them to reveal their personal preferences about things like food and extra-curricular activities. These agents simply prompted students with a question such as, "Would you prefer pizza or hamburgers for dinner?" Strict turn taking is enforced in this social dialog and a robust understanding module is used to map the student responses to one of the expected answers.

The cognitive support conversational agents were authored with the TuTalk dialogue agent authoring system (Gweon et al., 2005; Jordan et al., 2007). The TuTalk agent can lead a student through a knowledge construction dialogue, which is an interactive directed line of reasoning meant to lead a student to understand a concept, such as fraction division. As displayed in Figure 1, when the Filters module sends a notification to the Conversational Agents module to trigger a particular cognitive support agent, the scheduled TuTalk agent is appended to a queue of TuTalk Agents. The TuTalk agents in the queue act upon completion of the problem in the problem solving interface. This strategy is motivated by the intention of not interrupting an ongoing conversation among the students. Problem completion signals from the structured problem solving panel are assumed to correspond to major topic shifts in the chat among the participants.

Here's an example knowledge construction dialogue. In this case, all of the student contributions were from the same student, but both students have the opportunity to participate. The goal of the conceptual dialogues was to encourage students to think about the ideas behind the fraction arithmetic problems they were working on. Some concepts related to fraction division are particularly counter-intuitive for middle school aged students.

[1] TagHelper tools can be downloaded from http://www.cs.cmu.edu/~cprose/TagHelper.html.

Tutor: Another way to think of dividing by fractions is this: If I have 1 divided by 1/2, can I turn that into something divided by 1?. If I want to turn the denominator, which is 1/2, into 1, what do I need to do to it?
Student 1: yep
Tutor: Actually, you need to multiply by two. And If I multiply the denominator by 2, what do I have to do to the numerator?
Student 1: 2
Tutor: Right. So 1 divided by 1/2 is the same as 2 divided by 1. So that is one reason why the flip and multiply strategy for dividing by fractions works. If you multiply the denominator by something to turn it into 1, you have to multiply the top by the same value. So let's try another one. If I want to divide something by 2/3, what do I need to multiply the numerator and denominator by?
Student 1: 3
Tutor: Actually, it is 3/2. You should think about why it's 3/2. We'll talk about it more later.

The arrangement of the lab in which our study was conducted was such that each student was sitting at his own computer in such a way that collaborating pairs could not easily talk face-to-face since in all cases there was a row of desks with computers in between that student's row and the row where the partner student was sitting. The students were not told who their partner was or where they were seated, and they were asked not to reveal their identities, although in some cases they did. The purpose of this arrangement was to encourage communication through the chat interface so that it could easily be recorded and eventually processed on line during collaboration.

Methodology and Results

Experimental Design

The purpose of our study was to test the effect of social prompts on student interactions with each other and with the cognitive support agents. Our hypothesis was that the social prompts would increase student engagement in the conversation. Our experiment was a simple two condition between subjects design in which students in the experimental condition experienced interaction with social agents during two collaborative problem solving sessions, and students in the control condition did not.

In the experimental condition, a social dialogue agent was notified when the student interface was ready to begin a new problem. The social dialogue agents took the students through a directed system initiative dialogue to elicit their preference on certain items. Based on the student's preferences, the next math problem offered to the pair was formulated to include the given responses to the social prompts. For example, the agent might ask student 1 "Student 1, if you had to choose between a long flight or a long car ride, which seems more uncomfortable?" The student might indicate that a car ride would be preferable. Then the tutor agent might ask, "Student 2, which are more entertaining–books or movies?", and the student might respond that books are more amusing. These two pieces of information were then used to fill in slots in a template that was then used to generate the math problem that would then be displayed in the structured problem solving panel. In this case, the resulting story problem might say, "Jan packed several books to amuse herself on a long car ride to visit her grandma. After 1/5 of the trip, she had already finished 6/8 of the books she brought. How many times more books should she have brought than what she packed?" The goal of the social dialogs was to give students the impression that the support agents were taking a personal interest in them and that they had the opportunity to work together to create the math problems they were solving.

In order to control for content and presentation of the math content, we used the same problem templates in the control condition, but rather than presenting the social prompts to the students, we randomly selected answers to the social questions "behind the scenes" from the same set of choices offered to the students in the experimental condition. Thus, students in both conditions worked through the same distribution of problems.

The cover stories for the math problems used in this study were designed to be highly familiar to students. The design of the pool of cover stories as well as the accompanying social questions were based on data collected from an earlier pilot study in which urban middle school kids participated in a two week math camp focused on fraction arithmetic. During those two weeks, students participated in 5 hour intensive instructional sessions 3 times a week in which they were exposed to a progression of topics related to fraction arithmetic such as basic fraction concepts, addition and subtraction of fractions with like denominators, addition and subtraction of fractions with unlike denominators, fraction multiplication, and fraction division. During the two week period of time of the summer camp, the students kept math journals about how they saw themselves using their math in their every day lives outside of the in-class sessions. Each class session began with students sharing from their math journals about how they used their math at home. Although many of the examples the students included in their journals were examples of using numbers, but not necessarily the math they were learning at camp per se, the journal entries provided a stimulus for a type of "math story game" in which the group worked together to weave the math stories from the journals into fraction arithmetic problems, which the group then solved together. Often the group problem solving discussion was conducted with the student who's journal inspired the problem at the board leading the discussion, with the support of the instructor. We observed that students became very excited about solving math problems that they saw as being about themselves. Furthermore, we noticed certain consistent themes about the stories students came up with. For example, students talked about buying things, cooking, measuring familiar objects such as their room, or taking trips.

Experimental Procedure

The experimental procedure extended over 4 school days, with the experimental manipulation taking place during days two (i.e., Lab Day 1) and three (i.e., Lab Day 2). The fourth day of the experiment was separated from the third day of the experiment by a weekend. Teams remained stable throughout the experiment. The students were instructed that the teams would compete for a small prize at the end of the study based on how much they learned and how many problems they were able to solve together correctly. The second and third days were lab days in which the students worked with their partner. Each lab session lasted for 45 minutes. At the end of each lab period, the students took a short quiz, which lasted about 10 minutes. At the end of the second lab day only, students additionally filled out a short questionnaire to assess their perceived help received, perceived help offered, and perceived benefit of the collaboration. On the fourth experiment day, which was two days after the last lab day, they took a post test, which was used for the purpose of assessing retention of the material.

Subjects and Materials

Thirty sixth grade students from a suburban elementary school participated in the study. The students were from 2 different classes taught by the same teacher, with 16 students in the first class and 14 students in the second class. Students were arranged into pairs by the experimenter in such a way as to maintain a roughly consistent average grade so far in the course between pairs, and a balanced average grade so far in the course per condition.

The materials for the experiment consisted of the following:

- A mathematics tutoring program covering problems on fraction addition, subtraction, multiplication, and division.
- 2 extensive isomorphic tests (Test A and Test B) were designed for use as the pre-test and the post-test. Likewise, we had Quiz A and Quiz B, which were designed to be isomorphic to a subset of the pre/post tests. Thus, quizzes are shorter versions of the tests. Thus, we were able to use gains on quizzes to measure learning within sessions and pre to post test gains as a measure of retention (since there was a two day lag between the last lab day and the post-test).
- Questionnaire. As a subjective assessment of socially oriented variables, we used a questionnaire with 8 questions related to perceived problem solving competence of self and partner, perceived benefit, perceived help received, and perceived help provided. Each question consisted of a statement such as "The other student depended on me for information or help to solve problems." and a 6 point scale ranging from 0, labeled "strongly disagree", to 5, labeled "strongly agree". This questionnaire was developed and used in our previous work (Gweon et al., 2007).

Results

We began our analysis by investigating the socially oriented variables measured by means of the questionnaire, specifically perceived problem solving competence of self and partner, perceived benefit, perceived help received, and perceived help provided. Recall that students responded to each question using a 6 point likert scale, ranging from 0, which signified strong disagreement, to 5, signifying strong agreement. The only significant differences were in terms of perceived help received and perceived help provided. Students in the experimental condition rated themselves and their partner significantly higher on offering help than in the control condition.

Table 1 Questionnaire Results

	Control	Experimental
Perceived Self Competence	4.2 (.56)	4.1 (.23)
Perceived Partner Competence	4.3 (.62)	3.9 (.49)
Perceived Benefit of Collaboration	4.5 (.74)	4.4 (.70)
Perceived Help Received	1.8 (1.3)	3.3 (.69)
Perceived Help Provided	1.8 (1.1)	3.1 (1.1)

In order to investigate whether students in the experimental condition actually offered each other more help, we coded the chat logs from each lab day with a coding scheme developed in our previous work (Gweon et al., 2007). In order to make the sometimes cryptic statements of students clearer during our analysis, and also to provide an objective reference point for segmenting the dialogue into meaningful units, we merged the logfile data recorded by the problem solving interface with the chat logs recorded from the chat window using time stamps for alignment. We then segmented the conversational data into episodes using the log files from the tutoring software as an objective guide. Each episode was meant to include conversation pertaining to a single problem solving step as reified by the structured problem solving interface. Conversation related to a single social prompt also counted as 1 episode. And conversation related to one cognitive support agent also counted as 1 episode. All entries in the log files recorded by the tutoring software refer to the step the action is associated with as well as any hints or other feedback provided by the tutoring software. Note that steps where no conversation occurred did not have any episode associated with them.

Our coding scheme has 5 mutually exclusive categories, namely (R) Requests received, (P) Help Provision, (N) No Response, (C) Can't Help, and (D) Deny Help. Along with the "other" category, which indicates that a contribution does not contain either help seeking or help providing behavior, these codes can be taken to be exhaustive. The first type of conversational action we coded were Help Requests (R). Help Requests are conversational contributions such as asking for help on problem solving, asking an explicit question about the domain content, and expressing confusion or frustration. Not all questions were coded as Requests. For example, there were frequent episodes where students discussed coordination issues such as whether the other student wanted to go next, or if it was their turn, and these questions were not coded as help requests for the purpose of addressing our research questions. Adjacent to each coded help request, in the column associated with the partner student, we coded four types of responses. Help provisions (P) are actions that attempt to provide support or substantive information related to the other student's request, regardless of the quality of this information. These actions are attempts to move toward resolving the problem. Can't help statements (C) are responses where the other student indicates that he or she cannot provide help because he or she doesn't know what to do either. Deny help (D) statements are where the other student responds in such a way that it is clear that he or she knows the answer but refuses to stop to help the other student. For example, "Ask [the teacher], I understand it" or "Hold on [and the other student proceeds to solve the problem and never comes back to answer the original question]" are type D statements. And finally, no response (N) are statements where the other student ignores help requests completely. Each chat log was coded separately by 2 coders who then met and resolved all conflicts. Note that often where help requests are not met with a verbal provision of help, the students are still able to collaboratively or independently work out an answer to their questions, at least at the level of moving forward with the problem solving. In some cases, however, the students seem to move forward through guessing.

Here are two example episodes where a help request is met with the provision on help.

> **Student 1**: What operation do we do?
> *<student 2 tries multiplication and gets negative feedback from the problem solving environment>*
> *<student 2 tries divide and gets positive feedback from the problem solving environment>*
> **Student 2**: We divide. Now look at the problem, what is the other fraction we must divide by?

Student 1: What do we put on top of the fraction?
Student 2: Did you find a common denominator?
<student 1 correctly finds the common denominator>

Here are two example episodes where a help request is met with a Can't help response. In the second example, the student who requested help eventually figured out what to do on his own.

Student 1: Why 16?
Student 2: I don't know.

Student 1: I need help.
Student 2: Same
Student 1: 23/2
Student 2: What's 23/2?
Student 1: 11.5

Here are two example episodes where a help request is met with a Deny help response. In the first case, the student who asked for help was able to figure out the answer by guessing.

Student 1: I don't get it
Student 2: hold on
<then student 1 tried something and got negative feedback from the problem solving environment.>
<finally student 1 tried something else, which was correct, and got positive feedback from the problem solving environment>

Student 1: I don't know what to do
Student 2: click on the help button

Here are two example episodes where a help request is met with no response. In both cases the students seem to find the answer by guessing.

Student 1: I don't get it
<student 2 tries something and gets negative feedback from the problem solving environment>
<student 2 tries something else and gets negative feedback from the problem solving environment>
<student 2 clicks on the help button>
<student 1 tries something that is correct and gets positive feedback from the problem solving environment>

Student 1: ?
<student 2 tries something and gets negative feedback from the environment>
<student 1 tries something, which is correct, and gets positive feedback from the environment>

Table 2 displays the results from our coding of the corpus. First, we see that there is a significantly larger total number of episodes on the transcripts from the Experimental condition. Recall that all episodes contain some conversation. Steps where no conversation occurred do not count in our total number of episodes. The larger number of episodes in the Experimental condition is primarily due to the fact that episodes in which social prompts were given to students only occurred in the Experimental condition, and two of these occurred between every problem solved during the Experimental condition.

Table 2 Results from Corpus Analysis

	Experimental (Day 1)	Experimental (Day 2)	Control (Day 1)	Control (Day 2)
Total Episodes	47.1 (8.2)	61.3 (12.3)	33.8 (17.9)	49.1 (26.9)
Social Prompt Episodes	24.1 (9.9)	33.7 (16.2)	0 (0)	0 (0)
Help Episodes (P)	.79 (1.6)	.36 (1.1)	1 (1.3)	1.4 (2.9)
Unanswered Help Requests (C+R+N)	2.4 (2.7)	1.4 (1.9)	2.2 (1.9)	1.4 (1.4)
Non-Help Episodes	19.9 (5.6)	35.8(9.3)	30.6 (16.3)	46.3 (25.1)

Looking at the totals in Table 2, our finding regarding the average number of help provisions was that contrary to what we might suspect based on the questionnaire data, there was no significant difference between conditions, although there was a non-significant trend for fewer verbal help provisions to be given in the Experimental condition. The number help requests met with no verbal form of help was not different between conditions. However, there were significantly more non-help related conversational episodes in the control condition transcripts. Thus, the students in the control condition may have perceived less help behavior because there was a lower proportion of helping behavior. Overall, we observed that students displayed more negative affect in the control condition. Insults like "looser", "you stink", "stupid" only occurred in the control condition, never in the experimental condition, based on a keyword search analysis. Here is an example:

Student 1: finally
Student 2: Shut up
Student 1: oooooooooo burn
Student 2: I don't like you
Student 1: fine be that way
Student 2: how did you get that
Student 1: Guessing
Student 2: good, do you got it?
Student 1: no
Student 2: well too bad

The learning gains analysis offers some weak evidence in favor of the experimental condition on learning. The trend was consistently in favor of the experimental condition when comparing pretest to Quiz 1, Quiz 1 to Quiz 2, and pretest to posttest, although none of these comparisons were statistically significant. The strongest effect we see is on lab day 2 where students in the experimental condition gained marginally more on the segment of the test related to interpretation problems ($p=.06$, effect size .55 st. dev.).

Conclusions and Current Directions

In this paper we report on a study investigating the effect of social agents on conversational processes in a collaborative problem solving environment. Our finding was that while the social prompts encode no instructional material, they were not just extraneous entertainment either. The social prompts affected student attitudes and behavior towards each other. Furthermore, our study provides some weak evidence in favor of social prompts in connection with student learning.

While the effect of the social agents' prompts on student attitudes and behavior towards each other is promising, we are not satisfied with the interaction between the cognitive support agents and the student pairs. In particular, although we consistently see positive learning effects in our studies in conditions with the cognitive support agents in contrast to conditions without (Wang et al., 2007; Kumar et al., 2007), the agents often seem to appear as an interruption in the conversation. And students are frequently observed to "talk around" the agents rather than interacting with the agents. Based on results from an earlier study in which student interaction with a cognitive support agent in an individual learning scenario was intensified with the inclusion of social conversation (Rosé & Torrey, 2005), we hypothesized that the social agents in this study would positively effect the interaction between students and the cognitive support agents. However, we did not observe any difference between conditions on behavior with the cognitive support agents. Thus, in our current work an important focus is on improving the social graces of our cognitive support agents so that they enter the conversation with more sensitivity, and engage students' attention before proceeding with their attempted interactive instruction.

This work was supported by National Science Foundation grant number IERI REC-043779.

References

[1] Aleven, V., Sewall, J., McLaren, B. M., & Koedinger, K. R. (2006). Rapid authoring of intelligent tutors for real-world and experimental use. In Kinshuk, R. Koper, P. Kommers, P. Kirschner, D. G. Sampson, & W. Didderen (Eds.), *Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies* (ICALT 2006), (pp. 847-851).

[2] Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning*, 13, 175-193.

[3] Donmez, P., Rose, C. P., Stegmann, K., Weinberger, A., and Fischer, F. (2005). Supporting CSCL with Automatic Corpus Analysis Technology, *Proceedings of Computer Supported Collaborative Learning*.

[4] Fischer, F., Bruhn, J., Gruesel, C & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, 12:213–232.

[5] Gweon, G., Arguello, J., Pai, C., Carey, R., Zaiss, Z., Rosé, C. P. (2005). Towards a Prototyping Tool for Behavior Oriented Authoring of Conversational Interfaces, *Proceedings of the ACL Workshop on Educational Applications of NLP*.

[6] Gweon, G., Rosé, C. P., Zaiss, Z., & Carey, R. (2006). Providing Support for Adaptive Scripting in an On-Line Collaborative Learning Environment, *Proceedings of CHI 06: ACM conference on human factors in computer systems*. New York: ACM Press.

[7] Gweon, G., Rosé,C. P., Albright, E., Cui, Y. (2007). Evaluating the Effect of Feedback from a CSCL Problem Solving Environment on Learning, Interaction, and Perceived Interdependence, *Proceedings of CSCL 2007*.

[8] Jordan, P., Hall, B., Ringenberg, M., Cui, Y., Rosé, C. P. (2007). Tools for Authoring a Dialogue Agent that Participates in Learning Studies , *Proceedings of AIED 2007*.

[9] Joshi, M. & Rosé, C. P. (to appear). Using Transactivity in Conversation Summarization of Educational Dialogue, to appear in the *Proceedings of the SLaTE Workshp on Speech and Language Technology in Education.*

[10] Kumar, R., Rosé, C. P., Aleven, V., Iglesias, A., Robinson, A. (2006). Evaluating the Effectiveness of Tutorial Dialogue Instruction in an Exploratory Learning Context, *Proceedings of the Intelligent Tutoring Systems Conference.*

[11] Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., Robinson, A. (2007). Tutorial Dialogue as Adaptive Collaborative Learning Support, *Proceedings of AIED 2007*

[12] Piaget, J. (1985). *The equilibrium of cognitive structures: the central problem of intellectual development*, Chicago University Press.

[13] Rosé C. P., & VanLehn, K. (2005). An Evaluation of a Hybrid Language Understanding Approach for Robust Selection of Tutoring Goals, *International Journal of AI in Education* 15(4).

[14] Rosé, C. P., & Torrey, C. (2005). Interactivity versus Expectation: Eliciting Learning Oriented Behavior with Tutorial Dialogue Systems, *Proceedings of Interact '05*

[15] Rosé, C. P., Gweon, G., Arguello, J., Finger, S., Smailagic, A., Siewiorek, D. (2007). Towards an Interactive Assessment Framework for Engineering Design Learning, *Proceedings of ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*

[16] Rummel, N., Spada, H. & Hauser, S. (2006). Learning to collaborate in a computer-mediated setting: Observing a model beats learning from being scripted. In S. Barab, D. Hickey & K. Hay (Eds.) *Proceedings of the Seventh International Conference of the Learning Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

[17] Stahl, G. (2006). Group cognition in an online chat community: Analyzing collaborative use of a cognitive tool. *Journal of Educational Computing Research (JECR) special issue on Cognitive tools for collaborative communities*.

[18] Strijbos, J. W. (2004). *The effect of roles on computer supported collaborative learning*, PhD Dissertation, Open Universiteit Nederland, Heerlen, The Netherlands

[19] VanLehn, K., Graesser, A., Jackson, G. T., Jordan, P., Olney, A., Rosé, C. P., (2007). Natural Language Tutoring: A comparison of human tutors, computer tutors, and text. *Cognitive Science* 31(1), pp 3-52

[20] Vygotsky, L.S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press

[21] Wang, H. C., Rosé, C.P., Cui, Y., Chang, C. Y, Huang, C. C., Li, T. Y. (2007). Thinking Hard Together: The Long and Short of Collaborative Idea Generation for Scientific Inquiry, *Proceedings of CSCL 2007*.

[22] Wang, H. C. & Rosé, C. P. (2007). Supporting Collaborative Idea Generation: A Closer Look Using Statistical Process Analysis Techniques, *Proceedings of AIED 2007*

[23] Wang, Y. C., Joshi, M., & Rosé, C. P. (2007b). A Feature Based Approach for Leveraging Context for Classifying Newsgroup Style Discussion Segments, *Proceedings of the Association for Computational Linguistics*.
[24] Webb, N. & Farivar, S. (1999). Developing Productive Group Interaction, in O'Donnell & King (Eds.) *Cognitive Perspectives on Peer Learning*, Lawrence Erlbaum Associates: New Jersey.
[25] Webb, N., Nemer, K., & Zuniga, S. (2002). Short Circuits or Superconductors? Effects of Group Composition on High-Achieving Students' Science Assessment Performance, *American Educational Research Journal*, 39, 4, pp 943-989.
[26] Weinberger, A. (2003). Scripts for Computer-Supported Collaborative Learning Effects of social and epistemic cooperation scripts on collaborative knowledge construction, PhD dissertation, University of Munich