

Providing Support for Adaptive Scripting in an On-Line Collaborative Learning Environment

Gahgehe Gweon, Carolyn P. Rosé, Regan Carey, Zachary S. Zaiss
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh PA 15213
{ggweon,cp3a,zzaiss}@andrew.cmu.edu, reganmcarey@hotmail.com

ABSTRACT

This paper describes results from a series of experimental studies to explore issues related to structuring productive group dynamics for collaborative learning using an adaptive support mechanism. The first study provides evidence in favor of the feasibility of the endeavor by demonstrating with a tightly controlled study that even without adaptive support, problem solving in pairs is significantly more effective for learning than problem solving alone. The results from a second study offer guidelines for strategic matching of students with learning partners. Furthermore, the results reveal specific areas for needed support. Based on the results from the second study, we present the design of an adaptive support mechanism, which we evaluate in a third study. The results from the third study provide evidence that certain aspects of our design for adaptive support in the form of strategic prompts are effective for manipulating student behavior in productive ways and for supporting learning. These results also motivate specific modifications to the original design.

Author Keywords

Collaborative Learning, Adaptive Support, Computer Mediated Communication.

ACM Classification Keywords

Education, E-Learning.

INTRODUCTION

Typical on-line education programs with full faculty support and other benefits are not always significantly cheaper than on-campus post-secondary learning [17]. In contrast, free on-line programs such as the Open Learning Initiative at Carnegie Mellon University or the OpenCourseWare program at Massachusetts Institute of

Technology potentially offer educational materials to a wider audience, including the economically disadvantaged. Clearly, if these materials could be used effectively, these web-based resources could have a tremendous impact where the need is greatest - for example on the education of the low income population or those in developing nations.

Unfortunately, evaluations of these programs suggest important problems that must be solved. For example, students cannot benefit from materials unless they spend time working through them. An evaluation of MIT's open courseware program, offering on-line materials from 900 courses, revealed that although responders to a survey reported high satisfaction, only 11% of people who access the materials are return visitors¹. Thus, it has not succeeded in maintaining the involvement of the majority of initially interested students. Evidence from relative levels of participation in on-line discussion groups associated with MIT's OpenCourseWare program suggest that a key factor in eliciting the involvement of students in on-line learning communities such as these is providing sufficient structuring in the environment. Anecdotal evidence from student posts suggests that providing an infrastructure that supports effective, synchronous collaborative learning discussions would be highly attractive to students. Yet, such support is not provided in existing OpenCourseWare environments.

Beyond attracting and maintaining the involvement of students, such an environment could be used to enhance the instructional effectiveness of learning "in the wild", that is where faculty support is not available. While it is desirable for students to take initiative in their learning [2], and on-line learning environments provide the opportunity for autonomous learning, there is evidence that students require structuring of their learning experiences, both in terms of navigating the instructional materials [7, 22] and in terms of relating to one another in productive ways [5, 8]. The focus of our work is on the design of an adaptive support mechanism to enhance the instructional effectiveness of collaborative peer problem solving in the wild.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2006, April 22-28, 2006, Montréal, Québec, Canada.
Copyright 2006 ACM 1-59593-178-3/06/0004...\$5.00.

¹ http://ocw.mit.edu/NR/rdonlyres/90C9BC91-7819-48A0-9E9A-D6B2701C1CE5/0/MIT_OCW_2004_Program_Eval.pdf

In this paper we present a series of three studies in which we experimentally investigate foundational issues related to the design of adaptive support for on-line collaborative learning. Our empirical investigations focus on support for calculus learning since success in calculus is a key determining factor in college success in math and science based majors, and thus an area of great potential impact. The first study provides evidence in favor of the feasibility of the endeavor by demonstrating with a tightly controlled study that even without adaptive support, problem solving in pairs is significantly more effective for learning than problem solving alone. Nevertheless, there are many individual differences between students in terms of their level of engagement with the problem solving and their relative ability level. It has been conjectured that students learn from one another despite the erroneous information they communicate to one another in the process. Beyond this, it is argued that in fact the effectiveness of collaborative learning may be in part *because of* the exposure to erroneous information [3]. The idea of the instructional benefit of errors has its roots in Piaget's notion of perturbation or cognitive conflict [11]. Cognitive conflict plays an important role in stimulating cognitive restructuring by making students aware of a deficiency in their understanding for explaining the world around them.

Reflections from the first study lead us to ask what strategies for matching students with learning partners would produce the optimal conditions for learning. Previous work in this area provides mainly correlational evidence in favor of some combinations of ability levels and style over others [19, 23]. Building upon the results from the first study, we experimentally explored these group dynamics issues in greater depth in a second study in which we carefully manipulated engagement and ability level in confederate peer learners in order to precisely measure the causal impact of these variables on the behavior and learning of participant students working with them in pairs. We predicted an interaction between these variables since intuitively if a high ability and a low ability student are working together, for example, it makes sense for the high ability student to take the lead. The results from the second study offer limited evidence of the instructional value of exposure to errors. Beyond that, the results suggest that certain combinations of these two variables make students dangerous learning partners. Furthermore, we identified a disturbing lack of teaching behavior in the conversational logs from the first two studies, thus demonstrating a need for support in this area.

The results from the second study motivate the design of adaptive prompts both to encourage deep explanation and teaching behavior and to manipulate student behavior in an attempt to keep it out of the danger zone identified in the second study. We evaluated this design in a third study. The results from the third study provide evidence that adaptive support in the form of strategic prompts is

effective for manipulating student behavior and for supporting learning.

In the remainder of the paper we first describe the motivation for our approach as well as a review of the computer supported collaborative learning literature. We then present the three studies mentioned above. We conclude with some discussion and current directions.

MOTIVATION

The motivation for our work comes from observations of the lack of success to date with building active discussion groups in support of OpenCourseWare (OCW) environments. Our long term goal is to design and build a thriving on-line community to enhance the effectiveness of these free educational resources. At the time of writing this paper, out of 180 on-line discussion groups in connection with MIT's OCW community², only 7 had more than 10 posts, and only 2 had more than 50.

The two most active of MIT's OCW groups were for the two most highly instrumented courses, which provided resources such as suggested readings, video lectures, assignments, and tests. Discussion about what made these courses attractive was one of the topics of discussion found in the posts. For example, "There are other course in OCW that has few resource and hard to digest. The resources are more complete and easier to use." An informal analysis of the complete set of 142 postings to the Linear Algebra group posted over an 18 month period of time revealed frequent expressions of a desire for collaborative learning. For example there were posts simply asking "Is anyone out there still doing this?" and exclamations such as "Let's do it together!" or "I hope to drop in on future discussions!".

Despite expressions of a desire for collaborative learning in this on-line setting, there were occasional indications of discomfort with the newsgroup style interaction and desire for more synchronous communication. For example, while students talked about how great it was to "have students to study with", there were no instances of organized study sessions conducted in the newsgroup environment. Instead we saw students posting about isolated issues when they reached an occasional impasse. Out of 50 participants who ever posted to the discussion group, only 6 of them posted more than 5 times altogether, and only 2 posted more than 10 times. On one occasion we observed a pair express a desire to study together and then decide to take the interaction into a different environment, such as synchronous chat. This was the last post contributed from one of these students.

These informal findings suggest that the scant infrastructure that is commonly provided for OpenCourseWare courses is too impoverished to support a thriving, on-line learning community. What is needed is an environment that offers

² <http://mit.ols.usu.edu/courses/>

more structure and support for extended discussions. The focus of the research presented in this paper is on the second of these two concerns, namely supporting productive collaborative learning discussions in a computer-mediated environment. Not all instructional conversation between learners is equally effective [14], and sometimes is not better than non-interactive text for some populations of learners [15, 18]. Webb and colleagues present a series of studies in different educational settings that demonstrate the importance of the depth of instructional explanations, both for the speaker as well as the recipient [19]. Much research shows the value of drawing out student reasoning in the form of elaborated explanations. In particular, one of the best substantiated educational findings in cognitive science research related to education is the educational benefit of explanation, and in particular, the self-explanation effect [13]. Nevertheless, previous discourse analyses of collaborative conversations reveal that the majority of conversational interactions between students do not display the “higher order thinking” that collaborative learning is meant to elicit [19].

Meloth and Deering [8] present evidence of the importance of the teacher’s role in supporting effective collaborative learning. The teacher’s input is essential for keeping group discussions moving in a productive direction. The teacher provides key insights and models a productive learning process. Nevertheless, a human teacher is not required to provide this structuring. In order to encourage productive patterns of collaborative discourse, researchers both in the Computer Supported Collaborative Learning (CSCL) tradition [12, 16] and the Educational Psychology tradition [6, 10] have separately developed approaches for scaffolding the interactions between students, to help them coordinate their communication, and to encourage deep thinking and elaborated explanations. These simple forms of support are implemented as “scripts” or sets of prompts that students are provided with and expected to respond to. There is much evidence that argues for the effectiveness of these simple forms of support for boosting productive conversational behaviors [5].

Applying supportive scripting in a distance education context where we are concerned about student interactions in the environment over an extended period of time raises new questions not previously explored in the literature on scripting. While previous approaches to scripting vary along numerous dimensions, previous approaches to scripting were all static, one-size-fits-all approaches that were not sensitive to what was actually happening in the interactions. This can lead to over scripting [9] or interference between different types of scripts [21]. We hypothesize that over long periods of time, students will begin to ignore the prompts that scripts are composed of if they see them as not adapted to what is actually happening in the conversation since they will be seen as irrelevant.

We hypothesize that prompts offered only when deemed necessary will have more of an influence on the

conversation over time. For example, several studies have evaluated the impact of providing a social script that encourages productive consensus building behavior such as transactivity, which is a measure of the extent to which student contributions directly address the contributions of the other students in the group [20, 21]. Such conversational behavior is accomplished by assigning students to roles (i.e., case analyst or constructive critic) and providing prompts that target particular ways in which contributions may relate to each other, for example “We have not reached consensus concerning the following points:”. Rather than providing this prompt each time students formulate a contribution, as is the current, non-adaptive approach, a more adaptive approach would be to offer this prompt only in cases where non-productive forms of consensus building are detected, for example, where students fall into a pattern of quick consensus building rather than discussing the reasons for their differing points of view. Recent work demonstrates that patterns such as this can be detected with a high degree of reliability in collaborative discourse [4]. Nevertheless, the potential disadvantage of this adaptive scripting approach is that students receive much less scaffolding overall. Thus, it is necessary to experimentally verify whether this dramatically reduced level of scaffolding will be sufficient to yield a noticeable effect on behavior and learning.

METHODS AND RESULTS

Experimental Paradigm

One innovative aspect of the work presented in this paper is the experimental paradigm, which provides a highly controlled way to examine mechanisms by which one peer learner’s behavior influences a partner learner’s behavior and learning. This was accomplished by pairing real students with confederate peer learners who were staff members on the research team behaving in a highly prescribed manner. By holding the behavior of one member of a dyad constant within conditions but varied systematically across conditions, we can measure the causal effect of the variables we manipulate. Furthermore, this approach allows us to observe the interaction between both typical and unusual combinations of the variables we manipulate. While this approach lacks the high degree of external validity found in more naturalistic observations of collaborative learning interactions, it provides complementary insights not possible within that framework. Confederate peer learners were used in the two latter studies reported in this paper. In the first study, where we contrast solitary problem solving (SOL) and naturalistic peer problem solving (P2P), no confederate peer learners were necessary. An identical experimental procedure and infrastructure were used across all three studies.

Experimental Procedure

The experimental procedure consisted of 5 phases, composed of three test phases alternating with two

instructional phases. The experimental manipulation took place during phase 4, which was an instructional phase.

We strictly controlled for time in all phases. During the pre-instructional testing phase (phase 1), students filled out a consent form, took a pretest to assess their domain specific knowledge (for 15 minutes), and read the instructions for the first instructional phase. During the first instructional phase (phase 2), which was a human tutoring phase lasting 45 minutes, students received tutoring on the general concept of differentiation as well as 7 specific rules of differentiation from a human tutor. Although requiring students to learn independently in the first instructional phase would have been closer to what students face in real on-line environments, we needed to provide students with some common ground quickly for the purpose of the short term lab study. The tutor was blind to the student's condition and adhered to a rigid schedule for covering all of the content in a consistent way across students. During the mid-instructional testing phase (phase 3), students took a short middle test to assess their learning during phase 2 (for 10 minutes). They also read the instructions for the second instructional phase. The second instructional phase (phase 4), was a problem solving phase where students worked through as many of 12 multi-step derivation problems as possible during the allotted 35 minutes. Finally, in the post-instructional phase (phase 5), students took the post-test (for 15 minutes) and filled out a questionnaire.

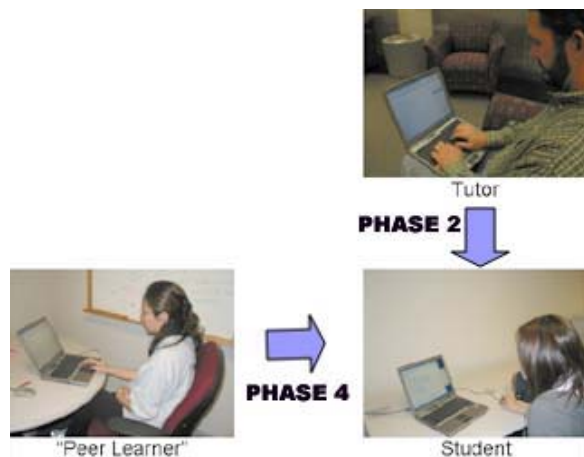


Figure 1. The student participant and tutor work together during Phase 2. The student participant and confederate peer learner work together during Phase 4.

The experimental setting is displayed in Figure 1. The student participant, tutor, and confederate peer learners were all located in separate rooms. The tutor and the peer learner roles were each played by 2 of the members on the research team each time. All students were told that their participation was part of a contest up front in all 3 studies. Also, the role of the student participants in all collaborative conditions across all studies was the same. Pairs working together interacted with a shared web-based problem solving interface using RealVNC software. The interface is

described in greater detail below under the “Problem Solving Interface” section. They communicated with one another using MSN Messenger. During the tutoring phase (Phase 2), time stamped logs of chat behavior were recorded. During the problem solving phase (Phase 4), submitted solutions, points assigned for each problem, and all chat behavior were collected in time-stamped logs. In addition, all activity with the problem solving interface was recorded using Camtasia Studio software made by TechSmith Co. In the SOL condition, students worked alone during Phase 4 using the same web based interface. Their interactions with the interface were also recorded using Camtasia Studio. They inserted think aloud comments in the MSN Messenger interface as well.

Materials

The experimental materials consisted of the following:

- An 8 page web based lesson designed in collaboration with a calculus instructor from the Math department at Carnegie Mellon University. This lesson that focused on derivatives provided material for the tutor and student to work through during Phase 2. It consisted of an overview and individual units on each of 7 specific rules of derivation. Each unit consisted of some explanation of the rule and an example problem for the student to work through using a structured problem solving interface.
- 12 on-line problem solving exercises for Phase 4, each requiring the applications of multiple derivation rules.
- 2 extensive tests (Test A and Test B) were used for the pre-test (in Phase 1) and the post-test (in Phase 5). These tests each consisted of 7 algebraic manipulation problems, 7 simple calculus problems to test knowledge of each specific differentiation rule, and 6 complex calculus problems requiring both multiple rule applications and algebra. In order to maintain consistency of content coverage and difficulty across tests, each problem on test A had an isomorphic problem on Test B, which required the use of the same skills. To further control for test difficulty and coverage, we counterbalanced the order of the tests. In Phase 3, students took a middle test with 8 simple calculus problems, isomorphic to the second section of tests A and B, and three complex calculus problems requiring multiple rule applications.
- Instructions for Phase 2 were provided on paper. The instructions before the first instructional phase were identical for all conditions except that students in the solitary learning condition in the first study were told that they were preparing to solve problems independently, whereas students in other conditions were told they were preparing to solve problems with a peer.
- Instructions for Phase 4 were again identical for all but the SOL condition. The instructions for all but the SOL condition began with the following:

“During this portion of the experiment, you and another student will be working together to complete Calculus problems based on the rules you learned in Part 1. You will have 35 minutes to complete as many problems as you can (up to 12). Both students will be manipulating the webpage that you see on the screen. Each problem is worth up to 2 points. In order to get 2 points on a problem, both students must contribute equally to the problem solving and the solution must be correct. A correct solution where one student does the most problem solving will only be worth 1 point.”

The instructions for SOL were identical except that all mention of a peer problem solver and division of labor were deleted since these are not relevant for solving problems alone. Thus, in the SOL condition, students were told they would receive 2 points for a correct solution and 0 otherwise.

Problem Solving Interface

All on-line problem solving was done using a structured problem solving interface designed for solving differentiation problems. See Figure 2. Students first select a rule from a menu. Based on their selection, some explanation about the rule and slots to fill in were presented to the student. In some cases, additional menus were presented, allowing for embedded rule applications.

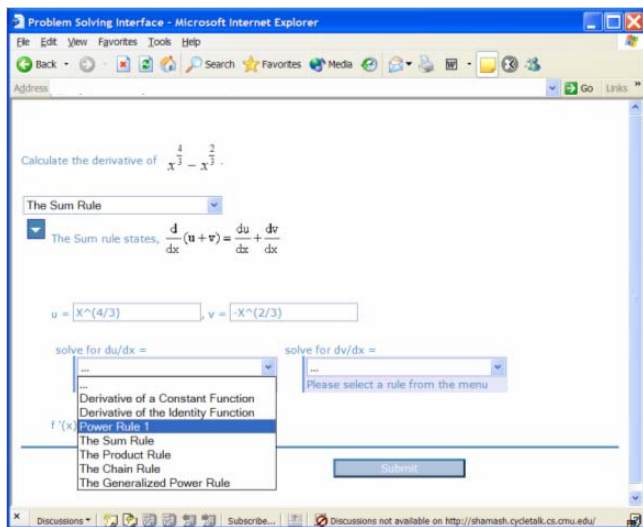


Figure 2 Structured problem solving interface.

No feedback was provided by the system based on the students’ selections from the menu or entries in the text input boxes during the problem solving process. When the student or pairs of students were satisfied with their solution, they submitted it. If it was incorrect, they were then shown their incorrect derivation next to the correct one as a worked example including both the derivation and some explanation. The purpose was for them to compare and see how the problem should have been worked out and where their mistake occurred. See Figure 3 for an example. This approach of offering a correctly worked out version of

the problem as feedback was motivated by a long and well established tradition of learning from worked out examples. Presenting students with correctly worked out examples has been demonstrated to be highly effective for learning, even more effective than problem solving at early stages of skill acquisition [1, 13].

In the case of a correct submission, the students moved on to the next problem. There was no need to compare their solution with an ideal solution if their solution was correct. At all times their current score was displayed next to an unchanging Highest Score.

PeerLearner: okay.. i think that's the answer, what do you think?

RealStudent: I have no idea

PeerLearner: so let's try submitting then

...

PeerLearner: damn it...we got it wrong again... i think we almost had it, right? we got product rule

RealStudent: shall we move on?

PeerLearner: but do you understand this?

RealStudent: hell no

PeerLearner: cause we're going to keep getting things wrong if we don't understand... so should we study this a little maybe?

Figure 3 Dialogue between a confederate peer learner and a student participant when comparing their submitted faulty solution with the correct solution displayed next to it

Study 1: Exploring the Benefits of Collaboration Under Non-ideal Circumstances

Collaborative learning in on-line learning communities without faculty support can be seen as risky since presumably all of the participating students are still in the process of learning the material. So the support they can offer each other is necessarily imperfect. In classroom settings where collaborative learning has been used successfully, the teacher plays an important supportive role in facilitating productive student interactions [8]. Where this support is absent, it is not clear whether collaborative learning will be beneficial. Nevertheless, we hypothesized that even in the absence of a teacher-facilitator students would benefit from solving problems with a peer if they were rewarded for cooperating with one another. Although students in our study did receive some faculty support in the form of tutoring during Phase 2, the collaborative problem solving phase in which the experimental manipulation took place was unsupported. The students in our studies believed they were participating in a contest in which they would be rewarded both for their correct problem solving behavior as well as for the extent to which they kept their distribution of labor equal. This competition scenario (where dyads cooperated with one another, but competed with other

dyads) is a typical collaborative learning configuration that provides a light form of non-invasive support for collaboration. The purpose of Study 1 was to measure the effect of collaborating with a peer on calculus problem solving in the context of this light support.

Experimental manipulation

The experimental manipulation, which occurs during Phase 4 of the experimental procedure discussed above, consisted of Solitary problem solving (SOL) and Naturalistic Peer-to-Peer problem solving (P2P). In the SOL condition, students solved problems alone during Phase 4, whereas in P2P students solved the same problems, but in pairs. Note that in contrast to the collaborative problem solving conditions in the second two studies, in the P2P condition, both students are student participants, not experimenters.

Subjects

21 participants for whom we measured learning were undergraduates or administrative personnel at the Carnegie Mellon University. They were randomly assigned to the two conditions. 12 students were assigned to the P2P condition in 6 pairs, 4 of which were same gender pairs. 9 students were assigned to the SOL condition.

Results

Using an ANCOVA with Post-test score as the dependent variable, Condition (SOL versus P2P) as the independent variable, and Pre-test score as a covariate, we verified that students in the P2P condition learned more than their peers in the SOL condition $F(1,18) = 6.0$, $p < .05$, $MSE = 5.64$, effect size = 1.1 standard deviations. We did not use the mid-test score as a covariate along with pretest score because it was not reliably correlated with post-test score with this population of students after we first factored out the effect of pretest score. Note that this is not a methodological problem because the experimental procedure up until the mid-test was identical across conditions. Although both high and low pretest students benefited from collaboration, there was a trend for high pretest students to benefit more than low pretest students. The gap between gain in the solitary condition and in the collaborative condition widens as pretest score increases. Specifically, in the 2 pairs where high pretest students were paired with very low pretest students, the high pretest student gained substantially more than predicted based on their pretest score.

Discussion

These results are important because they demonstrate in a highly controlled setting the value of collaboration in problem solving despite the fact that students are fallible. Students contribute both correct and incorrect problem solving actions, advice, and feedback. Nevertheless, the interaction is beneficial although the degree of benefit may differ. In the second study, we systematically explore the impact of the erroneous contributions made by peer problem solvers. And in the subsequent study, we explore a

mechanism by which we can improve the effectiveness of collaboration using an adaptive computer based support mechanism.

Study 2: Exploring the Interaction Between Engagement and Competence

While the first study demonstrated that randomly assigned pairs of students collaborated with one another in a way that lead to significantly more learning than a solitary problem solving control condition, this initial success lead us to ask what strategies for matching students with learning partners would produce the optimal conditions for learning. We chose engagement and ability level as variables to manipulate since they are directly related to the standard by which we are measuring student performance, specifically correctness of solutions and evenness of distribution of labor. There are many reasons to believe these variables might interact with one another. For example, while we observed a benefit for collaboration in the first study even in the face of errors contributed by students working together, there is reason to believe that as errors are contributed with much higher frequency, they would become a hindrance and a distraction.

Experimental manipulation

The experimental design for the second study was a 2X2 factorial design in which we varied two factors describing characteristics of a scripted confederate peer problem solver, namely Lazy(LA)/Engaged(EN) referring to the frequency of the confederate problem solver's contributions, and High(HI)/Low(LO) referring to the accuracy of the confederate peer learner's contributions.

During this phase of the experiment, one member of our team acted as a confederate student and another experimenter kept track of score, timing, and distribution of labor in order to ensure that all students within the same condition were treated in a consistent way. The confederate student behaved according to the following rules:

- LA/EN: In the Lazy condition (LA), the confederate student contributed to solving the problem either by offering part of the solution in the chat window or by performing an action in the problem solving interface every 45 seconds. In the Engaged condition (EN), the confederate peer learner contributed every 8 seconds.
- HI/LO: In the High performing condition (HI), the confederate student provided only correct contributions. In the Low performing condition (LO), the confederate student provided incorrect contributions 2/3 of the time. 2/3 was chosen after some pilot testing since it seemed unrealistic for even a low performing student to get incorrect answers 100% of the time.

Subjects

36 university students and staff participated in the study, randomly assigned to the 4 conditions.

Results

As predicted, we found a significant interaction effect using an ANCOVA with Post-test scores as the dependent variable, LA/EN and HI/LO as factors, and Pre-test and Middle-test scores as covariates $F(2,30) = 7.47, p < .05, MSE = 7.41$. In a post-hoc analysis using a Bonferroni test, the students in the Engaged High performing condition achieved significantly higher post-test scores than the students in the Engaged Low performing condition, $p < .05$. There was a marginal trend in favor of Lazy Low in comparison with Lazy High $p < .1$. Thus, within the Lazy condition, Low performing partners were marginally more effective while in the Engaged condition, Low performing partners were significantly worse.

We found that the strongest predictor of student learning was the number of correct problems the pairs managed to submit during the problem solving phase (CorrectProb). We computed this with a linear regression between CorrectProb and Post-test score with effect of Pre-test score factored out. $R\text{-squared} = .70, p < .001, N = 36$. There was a main effect of the HI/LO factor on the number of correct solutions contributed, with the effect of Pre-test and Mid-test scores used as covariates, $F(1,33) = 49.1, p < .001, MSE = .93$, effect size = 2.4 standard deviations. This makes sense since errors contributed as part of the problem solving process must be corrected in order to submit a correct solution. Thus, errors cause more work to be required for a correct solution, and problem solving with additional work takes more time. On the other hand, errors may be left uncorrected, in which case the problem solving may not take more time, but the solution that is submitted will not be correct, and thus will not increment the number of correct solutions.

Based on the above reasoning, a reduction in number of problems submitted is predicted. Thus, we would predict that Low performing confederate peer learners would be less effective as learning partners since their errors slow down the rate at which correct problems are submitted. With this in mind, it is surprising that students in the Lazy Low condition performed marginally better (rather than significantly worse) than the students in the Lazy High condition. Furthermore, an ANCOVA with post-test as the outcome measure, LA/EN and HI/LO as factors, and pre-test and CorrectProb as covariates, we found a significant crossover interaction effect explaining an additional 4% of the variance that provided some weak evidence that the errors contributed by the fake peer learners sometimes had a positive effect on student learning. $F(2,30) = 4.96, p < .05, MSE = 10.68$. Student participants paired with Lazy Low performing confederate peer learners learned more than would be predicted based on their pretest score and how many correct problems they managed to submit.

Discussion

The results from the second study offer limited evidence of the instructional value of exposure to errors. More importantly, the results suggest that students who are high

engagement but low in ability level are dangerous learning partners. Working with high engagement, low ability level confederate peer learners was less effective for learning than any of the other conditions. It was significantly worse than working with high engagement, high ability level confederate peer learners. Thus, one goal for the design of an adaptive support for effective collaborative problem solving would be to slow down high engagement, low performing students so that they won't produce a harmful level of erroneous problem solving behavior that might confuse, distract, or hinder their peer. Furthermore, we identified a disturbing lack of teaching behavior in the conversational logs from the first two studies, thus demonstrating a need for support in this area. In study 3, we evaluate the effectiveness of an adaptive support mechanism whereby prompts are strategically offered to students when either of these two needs are evidenced in the collaboration.

Study 3: Evaluating the Impact of Adaptive Collaboration Support

While we hypothesize that prompts offered only when deemed necessary will be more beneficial to students over an extended period of time than non-adapted collaboration scripts, the potential disadvantage of this adaptive scripting approach is that students receive much less scaffolding overall. Thus, it is necessary to experimentally verify whether this dramatically reduced level of scaffolding will be sufficient to yield a noticeable effect on student behavior and learning.

Experimental manipulation

The experimental design for the third study was a 2X2 factorial design in which we varied one factor relating to characteristics of a confederate peer problem solver and one characteristic relating to adaptive collaboration support. Specifically, High(HI) versus Low(LO) was a replication from the previous study. Prompt (PR)/No Prompt(NP) referred to the presence or absence of adaptive collaboration support in the form of prompts. In all conditions, the confederate peer learner in this study followed the rules for Lazy (LA) peer learners from the previous study. Prompting was offered in one of four cases outlined below. The prompts given in each case were canned text worked out in advance so that they are presented the same way each time. Each one relates either to curbing frequency of contribution of high engagement student participants or eliciting reflection and explanation from the student participant. The prompts were not meant to change the role of the student participant, but to encourage behavior for instructionally beneficial collaboration [19]. In the prompt condition, students were told that automated prompts would appear on their screen to support their collaboration, but not on the other student's screen. The list of circumstances under which students received prompts are enumerated below. The exact prompts associated with these circumstances are listed in Table 1.

- #1: Vague Help Offered - Confederate makes incorrect problem-solving action and participant offers vague / incomplete help (e.g., “That’s wrong.”)
- #2: Answer Corrected but No Help Offered - Confederate makes incorrect problem-solving action and participant changes to correct problem-solving action
- #3: Participant Working Independently - Participant makes 5 correct problem-solving actions without confederate contribution
- #4: Insufficient Review of Incorrect Answer - Participant wants to move on from wrong answer page before 2 minutes have elapsed

Case	Prompts
1	The other student would benefit from more explanation. Please elaborate on your correction.
2	Help the student understand your correction. The other student seems to be struggling with this section of the problem. Please offer your assistance.
3	Please be sure you are working with the other student to solve the problem. It seems like the other student has not contributed lately. Why don’t you see if they need help?
4	It seems like you are moving on before understanding your errors. Please spend more time reviewing this page. Does the other student understand the errors made on this problem? Please share your understanding of this page with the other student.

Table 1. Prompts.

There were minor differences in the instructions in study 3 compared to study 2. The student participants in all conditions were told as part of the instructions prior to Phase 2 (with the tutor) that the other student would not receive tutoring. They were told that they should prepare to work with and teach the other students if they need help during the problem solving that occurs in Phase 4. This was reiterated in the second instruction sheet before the peer learning session in Phase 4. Students were also told that they would receive a bonus if the other student’s score improves in the post test.

Subjects

40 university students and staff participated in the study, randomly assigned to the 4 conditions.

Results

Overall, the results from our experiment offer evidence in favor of the effectiveness of adaptive support for improving student behavior and learning. They also point towards specific ways in which our design for adaptive support should be modified in order to be more effective. Here we will first examine in depth the effect of the prompts on student behavior. We will then examine the effect on learning.

We first evaluated whether the prompts offered to students had a significant effect on their behavior. Remember that the prompts were primarily for two purposes: namely, to regulate the frequency of contribution of students, and to increase the amount of teaching behavior students offered. To evaluate whether the prompts were effective for regulating the frequency of contribution of students, we first analyzed trends in change of distribution of labor over time in the problem solving logs for each student participant. We looked at the number of contributions made by student participants and confederate peer learners for each problem solution submitted. From this, we computed for each problem submitted a LaborDistribution score between 0 and 1 indicating how different from equal the distribution of labor was, with 0 being the best and 1 being the worst. This was computed by the following formula, where PLC indicates number confederate peer learner contributions and SPC indicates number of student participant contributions:

$$2 \times \left| 0.5 - \frac{PLC}{PLC + SPC} \right|$$

We then computed for each student an improvement score indicating the extent to which the distribution of labor became more equal during problem solving. We did this by subtracting the LaborDistribution score of the final problem submitted with that of the first problem submitted. Positive values indicate an improvement in distribution of labor, whereas negative values indicate the opposite.

On average the LaborDistribution scores in the no prompt conditions remained stable over time, whereas in the prompt conditions where students received prompts there was improvement over time and a significant correlation between amount of improvement and number of prompts received (R-squared=.27, P< .05, N=20). Students remained out of the danger zone in the experimental condition, with an average LaborDistribution score of .32. On average only one prompt related to distribution of labor was required over the entire collaborative problem solving session, although some students received as many as 3. Because not many prompts related to distribution of labor were required, there was no significant effect of the prompting manipulation on average distribution of labor between conditions. Nevertheless, based on the significant correlation between number of prompts and improvement in LaborDistribution score, we have some evidence from this that prompts are effective for manipulating behavior.

The effect of the explanation oriented prompts was more obvious upon inspection. We observed that they had a local effect on explanation behavior in that we saw students attempt in all cases to follow the instruction offered in the prompt.

Thus, prompts had a positive effect on student behavior in the intended direction, offering evidence in favor of our design for adaptive support. However we also observed some negative effects of prompts on student performance that also negatively interacted with student learning within the LO condition. This finding led us to revise our design for adaptive support. In particular, there was a non-significant trend for distribution of labor prompts to reduce the number of correct problems solved within the LO conditions. Although the effect was not significant, the added noise in terms of number of correct problems submitted obscured the difference in learning between the PR and NP conditions. Remember that there was a large and statistically significant correlation between number of correct problems submitted and student learning. When we factor out this effect by including correct problems submitted as a covariate in an ANCOVA comparing pre to post test gains of students in the PR condition to students in the NP condition, we see a significant benefit for prompting on student learning. $F(1,37) = 4.12, p < .05$, effect size .66.

In future iterations we plan to modify our distribution of labor prompts so that the ideal distribution of labor is dependent upon the relative ability levels of the two students. Upon reflection, it does not make sense to encourage students who contribute errors 2/3 of the time to take an equal role in the problem solving. However, we do not want to completely discourage their involvement. So we need to explore further how to balance the concern over maximizing the number of correct problems submitted and optimizing balance of engagement between partners.

Further analysis of the learning gains reveal further insights for appropriate matching of students with optimal learning partners. We found a significant aptitude-treatment interaction, showing that High performing peer learners become less effective as learning partners as student pretest scores increase while Low performing peer learners become more effective as partners as student pretest scores increase. $F(2,35) = 5.97, p < .05, N=40$. The difference in effectiveness between High performing peer learners as partners versus Low performing peer learners as partners for high pretest students is in favor of Low performing peer learners, but as in the previous study, the difference is only marginal. However, the lack of significance could simply be due to a Type II error, thus we are continuing to collect data. We did not find this aptitude-treatment interaction in our previous study because the range of pretest scores was much higher in the first study.

Discussion

The results from the third study are particularly interesting from the standpoint of supporting effective collaborative

learning. First, these results demonstrate that an approach to automatic strategic prompting based on patterns in collaborative discourse have a significant impact on learning. In our current work, we are working on automating this automatic prompting. Prior work has demonstrated excellent results automating the application of a sophisticated multi-dimensional coding scheme for characterizing the collaborative learning process to naturally occurring collaborative learning data [4]. Thus, we believe the goal of adapting this technology for use in creating an environment that automatically offers students this form of strategic collaboration support is within our reach. Furthermore, the results from our investigations yield insights that can be used in matching students for effective learning together. If more data renders significant the difference in effectiveness of High Performing versus Low Performing confederate peer learners as learning partners, we can also use these results to motivate the design of more effective pedagogical agents that are tailored to the competence of the students who will use them as virtual learning partners.

CONCLUSIONS AND CURRENT DIRECTIONS

In this paper we have argued that the scant infrastructure that is commonly provided for OpenCourseWare (OCW) courses is too impoverished to support a thriving, on-line learning community. Based on observational data from existing discussion groups connected with MIT's OCW environment, we have argued that what is needed is an environment that offers more structure and support for extended discussions. We have presented 3 controlled lab studies where we have explored foundational issues related to the design of an adaptive support mechanism to facilitate productive collaborative learning interactions on-line.

We are currently working on a detailed analysis of the corpus. We also plan to continue refining the design of our adaptive support mechanism through further lab studies and more naturalistic classroom studies. Furthermore, we are building a working prototype of our adaptive support mechanism by applying techniques published by Donmez [4] for automating the application of coding schemes for characterizing patterns found in collaborative discourse. Ultimately, our design must be validated through a long-term study in an on-line community.

Beyond these concerns most directly related to the issues we have focused on in the series of studies presented in this paper, there are other more basic issues that must be addressed before we will have met our goal of making OCW resources a viable option for serious education. For example, only 1 out of every 15 users who have an account in the OCW discussion groups we reported on in the Motivation section of this paper ever posted to any group. Thus, much is not known about the experiences of those users and why they chose not to post. We plan to explore ways of eliciting the active participation of these users. Furthermore, what is currently lacking is a needs

assessment of the target learners, namely low income and minority adults who are not currently on track for education in an on-campus setting. Based on data from student profiles in the groups we examined, students with accounts, whether or not they ever posted, were almost exclusively students enrolled in an on-campus university program or college graduates. It is not known whether the course materials currently available on these open educational resource websites are appropriately adapted for the target user population. This important issue must be addressed before this work can have the intended impact.

ACKNOWLEDGMENTS

This work was supported by NSF Grants SGER REC-0411483 and NSF/IERI REC-043779.

REFERENCES

1. Atkinson, R. Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps, *Journal of Educational Psychology*, 95, 4 (2003).
2. De Jong, T., Van Joolingen, W.R., Scientific Discovery Learning With Computer Simulations of Conceptual Domains. *Review of Educational Research*, 68, 2 (1998), 179-201.
3. De Lisi, R., Golbeck, S. Implications of Piagetian Theory for Peer Learning, in A. O'Donnell & Alison King (Eds.) *Cognitive Perspectives on Peer Learning*, Lawrence Erlbaum Associates, New Jersey, 1999.
4. Domnez, P., Rosé, C. P., Stegmann, K., Weinberger, A., Fischer, F. Supporting CSCL with Automatic Corpus Analysis Technology, *In Proc. Computer Supported Collaborative Learning*. (2005).
5. Fischer, F., Bruhn, J., Gruesel, C & Mandl, H. Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, 12 (2002), 213–232.
6. King, A. Transactive peer tutoring: Distributing cognition and metacognition. *Computer-Supported Cooperation Scripts 52 Educational Psychology Rev.* 10 (1998), 57-74.
7. Kinshuk, Lin, T. User exploration based adaptation in adaptive learning systems. *International Journal of Information Systems in Education*, 1, 1 (2003), 22-31
8. Meloth, M. S., Deering, P. D. The Role of the Teacher in Promoting Cognitive Processing During Collaborative Learning, in O'Donnell & King (Eds.) *Cognitive Perspectives on Peer Learning*, Lawrence Erlbaum Associates: New Jersey, 1999.
9. O'Donnell, A. M. Structuring Dyadic Interaction Through Scripted Cooperation, in O'Donnell & King (Eds.) *Cognitive Perspectives on Peer Learning*, Lawrence Erlbaum Associates: new Jersey, 1999.
10. Palinscar, A.S., Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction 1* (1984), 117-175.
11. Piaget, J. The equilibrium of cognitive structures: the central problem of intellectual development, Chicago University Press, 1985.
12. Pfister, H-R., Mühlpfordt, M. Supporting discourse in a synchronous learning environment: The learning protocol approach. In Stahl, G. (Ed.), *In Proc. CSCL* (2002), 581-589.
13. Renkl, A. Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning & Instruction*, 12 (2002), 529-556.
14. Rosé, C.P., Aleven, V., Carey, R., Robinson, A., Wu, C. A First Evaluation of the Instructional Value of Negotiatble Problem Solving Goals on the Exploratory Learning Continuum, *In Proc. AI in Education.*, (2005)
15. Rose, C.P., VanLehn, K. & NLT Group. Is Human Tutoring Always More Effective than Reading, *In Proc. AIED Workshop on Tutorial Dialogue Systems: With a View Towards the Classroom*. (2003)
16. Rummel, N., Spada, H., Caspar, F., Ophoff, J. G., Schornstein, K. Instructional support for computer-mediated collaboration – results from process analyses. *In Proc. CSCL* (2003), 199-208.
17. Steeples, C., Jones, C. Networked Learning: Perspectives and Issues, Sringer, London, 2002.
18. VanLehn, K., Graesser, A., Tanner, J., Jordan, P., Olney, A., Rosé, C. P. When is reading just as effective as one-on-one interactive tutoring? *Proceedings of the Annual Meeting of the Cognitive Science Society*
19. Webb, N., Nemer, K., Zuniga, S. Short Circuits or Superconductors? Effects of Group Composition on High-Achieving Students' Science Assessment Performance, *American Educational Research Journal*, 39, 4 (2002), 943-989.
20. Weinberger, A., Ertl, B., Fischer, F., Mandl, H. Cooperation scripts for learning via web-based discussion boards and videoconferencing. *In Proc. The first joint meeting of the EARLI SIGs "Instructional Design" and "Learning and Instruction with Computers"* Knowledge Media Research Center (2004), 22-28.
21. Weinberger, A., Ertl, B., Fischer, F., Mandl, H. Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, 33, 1 (2005), 1-30.
22. White, B., Frederiksen, J. Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, 42 (1990), 99-157.
23. Wood, D., Bruner, J., Ross, G. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17 (1976), 89-10