Help Providers and Help Receivers in a Computer Supported Collaborative Learning Environment

Gahgene Gweon, Carolyn P. Rosé, Emil Albright, Yue Cui
Language Technologies Institute/Human-Computer Interaction Institute
Carnegie Mellon University, Pittsburgh, PA 15213
ggweon@cs.cmu.edu, cprose@cs.cmu.edu

ABSTRACT
In this paper, we explore the effect of the form of feedback offered by a computer supported collaborative learning (CSCL) environment on the roles that students see themselves as taking and that their behavior reflects. We do this by experimentally contrasting collaboration in two feedback configurations, one which is identical to the state-of-the-art in intelligent tutoring technology (Immediate Feedback), and one which is based on a long line of investigation of the use of worked out examples for instruction (Delayed Feedback). While our conclusions remain tentative due to the small sample size, the data reveal a consistent gender by condition interaction pattern across questionnaire, test, and discourse data in which male students prefer and benefit more from collaboration in the Immediate Feedback condition where they are more likely to take on the role of a help provider rather than a help receiver while the patterns is the opposite for females.

Categories and Subject Descriptors
H.5.3 [**Information Interfaces and Presentation(e.g., HCI)**]: Group and Organizational Interfaces – *Computer-supported cooperative work.*

General Terms
Measurement, Design, Experimentation.

Keywords
Gender, Collaborative Learning, Communication.

1. INTRODUCTION
In this paper, we explore the effect of the form of feedback offered by a computer supported collaborative learning (CSCL) environment on the roles that students see themselves as taking and that their behavior reflects. For decades a wide range of social and cognitive benefits have been extensively documented in connection with collaborative learning. Based on Piaget's foundational work [3], 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. Furthermore, based on Vygotsky's seminal work [5], 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. Connectied with this Vygotskian model of collaborative learning, we predict a conflict with the design of feedback, or scaffolding, that is the hallmark of the state-of-the-art in intelligent tutoring technology and is based on the same principles, and designed to meet the same needs. Our hypothesis predicts that the presence of typical intelligent tutoring style feedback in a collaborative problem solving environment will reduce the extent to which students assume the roles of help providers and help receivers. Furthermore, a reduction in helping behavior may then lead to a reduction in the exchange of alternative perspectives on problem solving, thus also interfering with the benefits of collaboration from the Piagetian perspective.

The social benefits of collaborative learning may be even more valuable for fostering a productive classroom environment. By encouraging a sense of positive interdependence between students, where students see themselves both as offering help and as receiving needed help from others, collaborative learning has been used as a form of social engineering for addressing conflict in multi-ethnic, inner-city classrooms. Some examples of documented social benefits of successful collaborative learning interactions include increases in acceptance and liking of others from different backgrounds, identification with and commitment to participation in a learning community, improvements in motivation, and aptitude towards long term learning [4]. These social benefits of collaborative learning are closely connected with the Vygotskian foundations of collaborative learning because the positive interdependence that is fostered is related to the exchange of support, or scaffolding, that we hypothesize will be replaced with the scaffolding offered by the environment where typical intelligent tutoring technology is used.

While our data do not support the strong form of this hypothesis, we see a consistent gender by condition interaction pattern across questionnaire, test, and discourse data in which male students prefer and benefit more from collaboration in the Immediate Feedback condition where they are more likely to take on the role

CSCW '06, November 4-8, 2006, Banff, Alberta, Canada

of a help provider rather than a help receiver while the patterns for female students is the opposite.

2. INFRASTRUCTURE FOR SUPPORTING COLLABORATIVE PROBLEM SOLVING

In this section we discuss the experimental infrastructure used to conduct our investigation, both in terms of the technology we used and in how we set up the lab where the students worked. Because of its tremendous effectiveness for individual learning with technology, we are planning to build our eventual collaborative learning infrastructure on the foundation of Cognitive Tutors [1]. In our study we are contrasting two designs for feedback from the environment, which we refer to as Immediate feedback and Delayed feedback. These alternative feedback paradigms have been experimentally contrasted in individual learning settings in the past [e.g., 2]. Typically, immediate feedback consists of what is called flag feedback, which signals to students after each problem solving action whether it was correct or not, and hints on demand, which are typically arranged in hint sequences, beginning with less directive hints and ending with more directive hints. In a delayed feedback setting, flag feedback is typically withheld so that students must use their own self-monitoring skills to detect their errors. Furthermore, hints may be withheld altogether or changed in nature so as not to be as focused narrowly on the correct solution path so that students have a greater responsibility for keeping themselves on track. In our study, both flag feedback and hints were withheld from students in the Delayed feedback condition. Instead, when students decided that their solution was complete, they submitted the solution and then were presented with a fully worked out version of the problem, with some explanation about how the solution was constructed. In order to control for information access between conditions, the instructional content in the explanation was constructed from the content encoded in the hints that students had access to in the Immediate feedback condition.

As the students worked in the lab session, their computer's display was composed of two panels that were next to one another. Using VNC, a panel on the lefthand side of the screen was shared between the screens of the respective computers of a collaborating pair so that they were both free to contribute to the evolving joint solution. Students communicated with one another using MSN messenger in a panel on the righthand side of their screen.

3. METHOD

3.1 Experimental Procedure

We designed an experiment to test the hypothesis that if students are working together in an environment in which they can obtain immediate feedback and help from the environment that is always correct, they would be less likely to turn to each other for help and feedback. This hypothesis predicts that in an environment with this form of feedback students would give and receive less help, would perceive less help given and received, and would benefit less from the collaboration. In order to test our hypothesis, we conducted a two condition within subjects design study where we manipulate the characteristics of the feedback and help offered by the environment. In both conditions students work in pairs to solve fraction addition, multiplication, subtraction, and division problems. In the control condition, students get immediate feedback from the cognitive tutor (Immediate Feedback condition), whereas the experimental condition students get delayed feedback (Delayed Feedback condition).

The experimental procedure extended over 4 school days, with the experimental manipulation taking place during days two and three, which we refer to as the first and second lab day since the students worked together in pairs in a computer lab at their school. The fourth day of the experiment was separated from the third day of the experiment by a weekend. Because our study is a within subject manipulation, we used two different units of material, each of which was experienced by each pair in only one condition or another so that we could distinguish learning resulting from work in one condition from learning resulting from work in the other condition. The two units were fraction addition and subtraction (AddSub) and fraction multiplication and division (MultDiv). We counter-balanced the order of the units and conditions in order to control for ordering effects as displayed in Table 1.

Table 1. The experimental setup

	Pairs	Day 1	Day 2
Class 1	1~4	AddSub, Imm	MultDiv, Delay
	5~8	MultDiv, Delay	AddSub, Imm
Class 2	9~11	AddSub, Delay	MultDiv, Imm
	12~15	MultDiv, Imm	AddSub, Delay

3.2 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. Out of 15 pairs who participated in the study, 12 of them were mixed gender pairs, 2 of them were all female pairs, and one of them was an all male pair. Because only a small number of pairs were homogeneous gender pairs, we cannot distinguish between gender effects that are specific to mixed gender pairs, versus gender effects that are independent of group composition.

The materials for the experiment consisted of the following:

- A mathematics tutoring program. The two mathematics chapters were fraction addition & subtraction and fraction division & multiplication.
- 2 extensive isomorphic tests (Test A and Test B) were used designed for use as the pre-test and the post-test. These tests each consisted of 16 near transfer and 8 far transfer problems, balanced between the two units of material. 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, administered after each lab day. And a segment of material was held constant across all 4 tests. Thus, we were able to use 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 day).
- Questionnaire. As a subjective assessment of socially oriented variables, we used a questionnaire with 6 questions related to

perceived interdependence as well as information on gender and average hours of daily computer usage.

4. RESULTS

4.1 Quantitative Analysis

We began our analysis by investigating the socially oriented variables measured by means of the questionnaire. The goal with respect to the socially oriented variables measured by the questionnaire is to maximize perceived benefit, perceived help received, and perceived help provided. Neither of our conditions consistently maximized all three of these variables for both genders. Instead we see a consistent pattern across all analyses whereby we observe a male preference for the Immediate feedback condition and a female preference for the Delayed feedback condition, although the contrast is statistically significant ($p < .05$) only in some cases and marginal ($p < .1$) in others.

Consistent with prior work investigating the well known gender gap in math achievement for middle school children, we found a gender effect whereby boys saw themselves as more competent with the material, although there was no significant difference in ability according to grade so far in the class as reported by their teacher. Boys rated themselves on the questionnaire as knowing more math $F(1,29) = 5.01$, $p < .05$, effect size .7 s.d., although there was no significant difference in grade so far in the class $F(1,29) = 0.46$, p = n.s. There was, however, a significant difference in pretest score whereby boys scored higher than girls $F(1, 29) = 6.13$, $p < .05$, effect size 1.2 s.d., thus demonstrating that they came into the experiment with more prior knowledge about the specific material covered. Boys also rated themselves as benefiting significantly less from the collaboration than girls did $F(1,29) = 6.65$, $p < .05$, although there was a significant interaction with condition such that the difference is only significant in the Delay condition $F(1,29) = 4.63$, $p < .05$, effect size 2.5 s.d. This effect did not seem to be related to the relatively higher pretest scores of boys since there was no significant correlation between perceived benefit and either the pretest score of the student or that of the partner.

As noted, overall boys rated the collaboration as less beneficial than girls did within the Delay condition. Furthermore, we found a significant gender by condition interaction in connection with perceived help provided $F(1,29) = 4.84$, $p < .05$ whereby girls' ratings of the extent to which they offered help was significantly lower than that of boys, but only in the Delay condition. There was a corresponding marginal gender by condition interaction $F(1,29) = 2.62$, $p = .1$ whereby girls ratings of the extent to which they received help were higher in the Delay condition, whereas the opposite was the case for boys. Thus, in one condition boys rate themselves as offering more help and receiving less as well as benefiting less, whereas the pattern of preference is the opposite for girls, although the effect is not as strong.

The learning gains analysis further corroborates the interaction between gender and pattern of preference observed on the questionnaire. There was no measurable gain on far transfer items either within conditions or over the whole population, thus we suspect that the material may have been too difficult for these students, and we consider only learning on near transfer items for the remainder of our analyses. On the immediate learning test we see a marginal crossover interaction between gender and learning gains on near transfer items such that there was a trend for girls to learn more on average than boys in the Delay condition, and for boys to learn more on average than girls in the Immediate condition $F(1,32) = 3.43$, $p = .07$.

4.2 Qualitative Analysis

The log files with student chat information contain rich data on how the collaborative problem solving transpired. We conducted a qualitative analysis of the conversational data recorded from MSN messenger in order to illuminate the findings from the tests and questionnaire data discussed above. In order to quantify giving and receiving help, we designed a simple coding scheme, which we applied to our data. We first segmented the 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. All entries in the log files recorded by the tutoring software refer to the step the action is associated with.

We approached the design of our coding scheme with some focal questions in mind. For example, we wondered how many times each student requested help in each condition. Furthermore, we wondered how their partners responded to their help requests. A preliminary cursory analysis of the MSN messenger logs revealed that frequently students requested help but did not receive any verbal response from their partner. We also observed signs of frustration between students and some cases where students explicitly refused to help one another. Our objective was to investigate where these uncooperative episodes are occurring, what their impact is on the patterns of preference and learning that we observe, and how we can discourage students from falling into this uncooperative behavior pattern. Because our focal questions all pertain to issues related to help seeking and help provision, we designed a coarse grained coding scheme to identify the regions of the integrated logfiles where we can find the answers to our questions. In the future we may code additional types of behaviors or make finer grained distinctions. Our current coding scheme has 5 categories, namely (R) Requests received, (P) Help Provision, (N) No Response, (C) Can't Help, and (D) Deny Help.

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 log file was coded separately by 2 coders who then met and resolved all conflicts. Using the consensus coding, we then tabulated the number of occurrences of each code in each condition associated with each gender. While there was a significantly larger number of conversational segments in the logs from the Immediate Feedback condition, the proportion of segments that contained a help request was not stable across conditions. Thus, there was no significant main effect of condition on raw numbers of either help requests received or offered. There was, however, a significant gender by condition interaction on raw number of requests received $F(1,42) = 4.79$, $p < .05$, and a marginal gender by condition interaction on both help requests given and help requests received when the raw counts are normalized by number of segments: $F(1,42) = 3.62$, $p = .06$ and $F(1,42) = 3.10$, $p = .09$ respectively. In all cases there was no significant or marginal gender effect except in the Immediate feedback condition, where males received more requests than females as well as participating in a higher proportion of discourse segments in which they received a request than females did. In contrast, females participated in a higher proportion of segments in which they made requests than males did.

We find an asymmetric collaboration pattern in which males appear as the help providers and females appear as the help receivers. To further investigate this finding, we compared counts of response types across conditions, normalized by number of requests. Data from transcripts where no requests were received were dropped from this analysis. There was a significant main effect of condition on number of Can’t Help responses such that a larger proportion of requests were met with a Can’t Help response in the Immediate Feedback condition than in the Delayed Feedback condition, with no interaction with gender $F(1,42) = 4.86$, $p < .05$, effect size 1.5 standard deviations. This suggests that the nature of help requests may have been different in the two conditions. Our coarse grained coding of the collaborative behavior does not allow us to further address the question of what caused this difference at this time.

For the other three response types, we see a significant gender by condition interaction but no main effect of condition: Help Provision $F(1,40) = 4.84$, $p < .05$; Deny Help $F(1,40) = 3.96$, $p < .05$; No Response $F(1,40) = 4.91$, $p < .05$. For girls, the proportion of Help Provision and Deny Help responses is lower in the Immediate Feedback condition than in the Delayed Feedback condition, but higher for No Response responses. The pattern is almost the opposite for boys, where proportion of Deny Help responses remains stable between conditions, but the proportion of No Response responses is lower in the Immediate Feedback condition than the Delayed Feedback condition, and the proportion of Help Provision responses is higher in the Immediate Feedback condition than the Delayed Feedback condition.

In terms of perceived help offered, there was no difference between how girls and boys rated themselves in the Immediate Feedback condition, but girls rated themselves as offering significantly less help in the Delayed Feedback condition than boys did. As mentioned, what we observed based on our corpus analysis is that girls responded to a higher proportion of help requests with a substantive answer in the Delayed Feedback condition, whereas boys responded to a higher proportion of help requests with a substantive answer in the Immediate Feedback condition. Thus, the surprising finding is that it appears that girls perceive themselves as benefiting more and receiving more help in the condition in which they are actually offering more help, and conversely, boys see themselves as receiving more help and benefit in the condition in which they are offering more help.

One possible explanation for why students perceived more help where they offered more help is that the act of offering help is an instructionally beneficial activity, and then when students engage in this activity, they perceive themselves as receiving help and benefit because they are learning. Recall that in the learning gains analysis reported above with the quantitative analysis, we observed that girls learned more in the Delayed Feedback condition where we see them offering more help, whereas boys learned more in the Immediate Feedback condition where we see them offering more help. As further evidence of this connection we see a significant correlation between total number of Help Provision responses and learning when we compute a multiple regression with pretest score and number of Help Provision responses as independent variables and posttest score as the dependent variable ($R^2=.84$, $p = .001$, $N=30$) and a significant gender by condition interaction on total number of Help Provision occurrences that mirrors the earlier analysis with respect to proportion of Help Provision responses $F(1,26) = 7.79$, $p=.01$. A Bonferroni posthoc analysis reveals a marginal difference between number of Help provision statements made by girls in the Delayed Feedback condition and that in the Immediate Feedback condition (effect size .89 standard deviations) and a marginal difference between number of Help provision statements made by boys in the Immediate Feedback condition and by girls in the Immediate Feedback condition (effect size .86 standard deviations).

5. CONCLUSION

In this paper we investigated the hypothesis that the presence of typical intelligent tutoring system style feedback in a collaborative problem solving interface would interfere with collaboration and dampen its positive effects. While our data do not support the strong version of this hypothesis, we are left with the challenge of reconciling the dichotomous needs and preferences of girls and boys. The observed learning effects support the value of encouraging students to take the role of a help provider.

6. REFERENCES

[1] Koedinger, K. J., Anderson, R. J., Hadley, W.H., Mark, M.A. Intelligent Tutoring Goes to School in the Big City. *International Journal of Artificial Intelligence in Education.* 8, (1997), 30-43

[2] Nathan, M. J. Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments*, 1998, 161-180

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

[4] Sharan, S. Cooperative Learning in Small Groups: Recent methods and Effects on Achievement, Attitudes, and Ethnic Relations. Review of Educational Research, Vol 50, No. 2 (Summer 1980), 241-271

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