On socio-cognitive processes that promote learning from peer collaboration and how immediate transfer tests cannot always detect their effects

Christa S. C. Asterhan, Baruch B. Schwarz, School of Education,
The Hebrew University of Jerusalem, Jerusalem 91905, Israel,
Email: christa.asterhan@mail.huji.ac.il, msschwar@mscc.huji.ac.il

Abstract: Research into the benefits of collaborative work on learning have shifted from questions regarding *whether* it promotes learning, to research into the *conditions* that promote learning and the identification of processes that make collaborative group settings particularly effective. In this symposium we will present findings from recent research into a number of socio-cognitive processes that have been found to foster conceptual gains following group learning. The papers that will be presented as part of this symposium will focus on three different phenomena: argumentation, production feedback and the incubation effect of collaboration. In addition, these studies also show that the effects of collaborative learning may not be apparent immediately following interaction, but need some time to materialize. This finding emphasizes the need for multiple and delayed assessment, as well as alternative assessment tools, such as prospective (instead of retrospective) measures of learning. .

Introduction

The potential benefits of peer interaction in learning activities have been and still are a popular topic of psycho-educational research, theory and debate. Overall, however, results from collaborative arrangements have been inconsistent. This may in part be attributed to the immense variability in research paradigms, operational variables, subject material and learning contexts that have been employed in the peer interaction literature. Most importantly, benefits of interaction are very likely to be dependent on the type of learning goals and material. According to Damon (1984; Damon & Phelps, 1989), collaborative peer interactions may create good opportunities to explore new ideas and may therefore be especially beneficial for concept learning. Indeed an increasingly large body of research works seems to corroborate this conjecture (e.g., Kruger, 1993; Schwarz, Neuman & Biezuner, 2000).

However, as the literature on peer collaboration has extensively and repeatedly shown, simply putting two (or more) students together is not sufficient (e.g., King & Rosenshine, 1993; Teasley, 1995). To fully exploit the potential benefits of peer interaction, the tasks and settings have to be adequately designed and structured to actively support those social and cognitive processes considered to be responsible for collaborative learning effects. Moreover, the benefits of such processes may take some time to materialize and may remain hidden when standard formats of immediate assessments are administered.

The present symposium will contribute in both these directions: Each of the three contributions will present recent research that provides new insights into the (socio-)cognitive processes that may be responsible for peer collaboration benefits in concept learning. They isolate these processes by creating conditions in experimental and/or instructional designs. In addition, they reveal the importance of multiple assessments to allow for the effect of these processes to be detected.

The paper by Christa Asterhan and Baruch Schwarz focuses on the role of argumentation within peer collaboration settings that are designed within a socio-cognitive conflict paradigm. The engagement in collective argumentation has been identified by many as beneficial to individual thinking skills, as well as to individual and collective knowledge construction. The researchers present two experimental studies that treat argumentation as an experimental condition and test its effect on conceptual gains in evolutionary theory. The results of these studies show that whereas peer collaboration by itself led to immediate gains in conceptual understanding, only those students that were explicitly instructed or scripted to engage in collective argumentation preserved these gains to later test occasions. This finding has important implications for the design of successful peer collaboration settings, in particular for learning tasks that are designed within the cognitive conflict paradigm for instruction.

Christine Howe has extensively researched the effects of peer collaboration vs. individual settings on conceptual understanding of scientific concepts. Several of her studies have shown that peer collaboration often results in conceptual growth that is detectable several weeks after collaborative group work. Moreover, these gains were not found to be predicted by the views expressed *during* the interaction. Borrowing from the developmental psychology literature she proposes three possible mechanisms for these *incubation* effects. Her paper presents findings from a set of three controlled experiments designed to uncover the mechanism behind these effects. The results strongly suggest that peer collaboration can 'prime' children to make good use of subsequent output.

Daniel Schwartz and Taylor Martin have undertaken a different approach. They relied on previous findings from research that identified the beneficial role of *production feedback* in peer collaboration. Production feedback is the feedback one receives when observing how somebody else uses one's ideas. They then created a computer-mediated environment that isolated and simulated this particular aspect of peer collaboration (a Teachable Agent). First results from implementations of this environment in classroom settings show that the inclusion of production feedback indeed improved learning and that these effects increased from a modest advantage at the first learning unit to a large advantage at the third and last learning unit. Moreover, prospective measures of learning reveal that it also improved students' preparation for learning new topics later on, when they were no longer using the technology.

Sometimes collaborating is not enough: the role of argumentation in learning notoriously difficult scientific concepts

Christa S. C. Asterhan and Baruch B. Schwarz
School of Education, The Hebrew University of Jerusalem, Jerusalem 91905, Israel,
Email: christa.asterhan@mail.huji.ac.il, msschwar@mscc.huji.ac.il

Many practitioners and researchers have identified argumentative dialogue in the classroom as a means of students' understanding of subject matter (e.g., De Vries, Lund & Baker, 2002; Duschl & Osborne, 2002; Nussbaum & Sinatra, 2003; Schwarz & Asterhan, in press; Schwarz, et al, 2000). As a practice, argumentation is best defined as an activity in which interlocutors cooperate in solving a particular problem to which a number of different solutions are proposed. Each of these solutions is perceived to be of a different epistemic status and interlocutors feel obliged to choose between them (Baker, 2003). In contrast to other forms of conflict resolution, they attempt to reach this goal by engaging in reasoning.

However, these potentially beneficial effects have thus far not been successfully subjected to rigorous experimental testing. Recently we conducted two experimental studies on the effects of argumentation on conceptual understanding in evolutionary theory (Asterhan & Schwarz, 2007). The learning tasks in these two studies were designed within the socio-cognitive conflict paradigm, according to which collaborating peers are either confronted with anomalous data or contradicting views and/or are paired with peers who have different views (Limon, 2001; Mugny & Doise, 1978). In the interventional phase of both studies students watched an educational movie on evolution which presented the scientifically accepted Darwinian account and they solved an evolutionary phenomenon that had been found to cause cognitive conflict in many students with misconceptions. In study 1, they were paired with a peer student in the intervention phase and asked to collaboratively solve two evolutionary phenomena. In the second study, individual students were exposed to a misconception other than theirs, which was read aloud by a confederate.

In the first study, undergraduates were randomly assigned to dyads and collaboratively tried to explain an evolutionary phenomenon (i.e., the evolution of webbed feet of ducks). Half of the dyads were instructed to engage in argumentative dialogue on their respective explanations and received some written examples of argumentative moves; the other half was merely instructed to collaborate. Individual evolutionary understanding was assessed as the quality of the explanatory schemas they used to explain different evolutionary phenomena on three separate test occasions: Prior to, immediately after and a week following the dyadic intervention. In addition to this measure of conceptual understanding, we also assessed the number of correct Darwinian principles that students explicitly mentioned in their written responses.

When controlled for pretest performance and nested effects of the individual within the dyad, delayed posttest explanations of individuals in the argumentative condition were found to testify of better conceptual understanding than those of control students. Furthermore, the pattern through which this advantage was attained revealed that students in both conditions improved their conceptual understanding immediately following the intervention. However, students who were merely instructed to collaborate lost this temporary gain, whereas students in the argumentative condition retained the same level of performance at the delayed posttest. The improvement in conceptual understanding as seen in the explanatory schemas students used could not be attributed to an increase in the number of correct Darwinian principles they produced: Students in both conditions showed immediate gains which disappeared on the delayed post-test a week later. Potential intervening variables, such as whether students arrived at the Darwinian solution *during* the interaction and the length of their discussions, were not found to be dependent on condition.

Taken together, these findings seem to suggest that the differences in conceptual understanding may be the result of different levels of processing during or after the intervention phase: The engagement in argumentation may have led to superior processing which resulted in consolidation of their immediate cognitive gains. On the post-test occasion, these restructured cognitive schemas were then retrieved and applied to new phenomena. In contrast, the immediate gains of subjects that did not engage in argumentation may have been due to the social affordances of the collaborative situation (such as exposure to superior solutions from more knowledgeable or dominant partners) or recall of information from the movie.

Our conjecture that the difference between the two conditions may be attributed to superior processing as a result of argumentation, would be further supported if such patterns of change could be replicated in an additional study. Therefore, a follow-up experiment was designed to address these and other issues in a more rigorously controlled design that further isolated the engagement in dialectical argumentation. To determine whether argumentation leads to superior processing *during* or *after* the intervention, students were also asked to report post-intervention reflections and deliberations on evolution.

The design and procedure of study two were almost identical, except for the following changes in the interventional phase: A confederate played the role of one of the participants in both conditions. Participants in the experimental condition were prompted to engage in dialectical argumentation on their own and the confederate's solution, by answering structured questions read aloud by the confederate, who chose a piece of paper from an urn and invariantly picked up the role of the "reader". In the control condition, the subject and the confederate only read aloud their solutions to each other, without discussing them further and performed a filler task to control for time-on-task. Thus, students in both conditions were prevented from conducting a natural dialogue and were exposed to the same naïve conception in evolution.

So as to ensure uniformity of exposure to another explanatory schema (i.e., the confederate's), while preserving a minimum difference between that and the student's explanatory schema, two different answer sheets were prepared for the confederate. Each contained a solution according to an explanatory schema that was qualitatively different, but belonged to the same schema category (see Asterhan & Schwarz, 2007 for more details). The solution that was read by the confederate as her own was thus contingent on the participant's explanatory schema.

The task scenario in the experimental condition was designed to ensure that participants engaged in dialectical argumentation in a controlled design, while preserving the perceived equal-status, peer-collaborative nature of the first study. First, participants were requested to read aloud their answer to the 'duck item' (for which students were asked to explain the evolution of webbed feet). They were then asked to discuss the strengths of that solution, to criticize it, and to discuss whether it explained the change that occurred to the ducks' feet. Then the confederate was requested to read "her" solution aloud, after which the participants were asked to discuss that solution according to the previous steps. In both conditions, students interacted with the confederate and the additional solution was always presented as being the confederate's, who personally read it to a student. In sum, the conditions were identical on factors such as social facilitation, actual exposure to an alternative view, the nature of this alternative solution and the personification of viewpoints. They differed only in engagement of dialectical argumentation.

The results showed that students who were instructed to engage in scripted dialectical argumentation on their own and another person's solution showed greater conceptual gains than control students. Thus, the advantage of argumentation observed in collaborative dyadic situations was replicated in a situation of scripted argumentation directed at and prompted by a peer, a design that isolated the argumentation from the interaction features of the collaboration.

Similar to the findings from study one (the dyadic study), students in the argumentative condition preserved the conceptual gains obtained during the intervention. However, students in the control condition did not show improvement on any of the tests. This suggests that control subjects' temporary gains in the dyadic study derived from the peer interaction *per se*, and not from the movie they saw. When students were not allowed to discuss each others' solutions and were exposed to the same misconception, such temporary gains disappeared. Instances of post-intervention reflection and deliberation were equally distributed among experimental and control students. It seems, therefore, that the preserved effects of argumentative conditions are not the result of processes that occur following, but during the interaction.

The combined results of these two studies first and foremost provide first experimental support for the assertion that argumentation promotes conceptual understanding in the Sciences in a cognitive conflict learning paradigm. Even a meticulously designed task meant to cause socio-cognitive conflict did not lead to lasting cognitive gains, unless the students were specifically instructed (dyadic study) or scripted (confederate study) to engage in dialectical argumentation. Secondly, our findings particularly emphasize the importance of delayed assessment, especially in the case of dialogical argumentation, since its potential benefits may not become apparent at immediate test occasions. Whereas peer collaboration by itself was found to have a positive effect on conceptual understanding, these gains proved to be merely temporary and disappeared at delayed post-tests.

Dialectical argumentation requires explaining oneself and justifying one's standpoints, as well as dialectically discussing and evaluating different solutions. We suggested that the advantage of argumentation for conceptual understanding is achieved through superior cognitive processing. This conjecture is indirectly supported by two findings: (1) neither mentioning the Darwinian account during the discussion nor reaching an agreed upon solution was related to learning gains; and (2) the particular pattern in which the advantage of argumentative conditions was achieved in both experiments (preserved gains versus loss of temporary gains or no gains). The implications for theory development and further research on the mechanisms behind these effects will be discussed in the presentation.

Incubation and the delayed benefits from group work in science

Christine Howe, Faculty of Education, University of Cambridge, Cambridge CB2 8PQ, UK. cjh82@cam.ac.uk

Over the past 25 years, evidence has mounted that conceptual understanding in science can be promoted by collaborative group work between children (Azmitia & Montgomery, 1993; Howe, Tolmie & Mackenzie, 1995; Howe & Tolmie, 1998). There is, however, continuing uncertainty about the mechanisms by which children learn from their collaborative experiences. One popular account sees children as constructing collective insights that are superior to their starting points. These insights are then individually internalised (Brown, Ash, Rutherford, Nakagawa, Gordon & Campione, 1993; Tomasello, Kruger & Ratner, 1993; Vygotksy, 1978). There can be little doubt that such internalisation operates on occasion, but it cannot be the whole story. In particular, it cannot account for conceptual growth that is detectable several weeks after collaborative group work and is demonstrably stimulated by the group work experience, but is not apparent during group work itself. Here views expressed (and agreed) during group interaction are often: a) inferior to individual starting points; b) unrelated to patterns of change. Such delayed effects have been reported repeatedly over the past 20 years (e.g. Howe, Tolmie & Rodgers, l992; Tolmie, Howe, Mackenzie & Greer, 1993), but they have never been satisfactorily explained.

One clue to a possible interpretation comes from the fact that delayed benefits can be regarded as a specific manifestation of the 'incubation' phenomenon that has intrigued cognitive psychologists for some considerable period of time (e.g. Dorfman, Shames & Kihlstrom, 1996). Therefore, the author has recently completed three studies, which examine the relevance of explanations proposed for incubation within cognitive psychology to the delayed effects of group work in science. These studies will provide the focus for this presentation (see Howe, McWilliam & Cross, 2005 for details). All three studies were conducted with primary school children aged 9 to 12 years and addressed mastery of the factors relevant to floating and sinking. The group task used in the studies required children to predict which, from sets of small objects, would float and which would sink, test their predictions by immersion in a tank of water, and interpret outcomes. The children worked through the task in triads or foursomes. Mastery of factors was established by paper-and-pencil tests that were administered to whole classes as pre-tests prior to the group tasks and as post-tests up to eight weeks afterwards.

Study One explored the possibility that collaborative group work can result in unhelpful mental sets, which can only be broken after an interval, whereupon productive (but essentially individual) work resumes. Thirty children (Condition 1A) worked in groups using a version of the group task that was believed from previous work to have relatively weak set-inducing powers. A further 31 worked on a version that was believed to have relatively strong set-inducing powers (Condition 1B). Analysis of group interaction confirmed these beliefs: by the end of the task, the 1B children had fixated on a much smaller set of factors than the 1A children. However, analysis of conceptual growth from the pre-tests to 'immediate' post-tests administered within one hour of the group task's completion, and from the immediate post-tests to delayed (eight-week) post-tests offered no support for the mental set hypothesis. There were no statistically significant differences between the conditions, with further analyses indicating that the mental sets had, in fact, already been broken by the time of the immediate post-test.

Study Two addressed the possibility that collaborative group work stimulates children to engage in private accretion and appraisal of factors once the task is completed, and this is what results in delayed learning gains. Thirty-six children (Condition 2A) worked in groups on the 1A group task, and then completed ‘reflective exercises'. These exercises took place two, four and six weeks after the group tasks and required the children, working on their own, to: a) write down 'all the ideas you have ever come across' about why pictured objects floated or sank; b) evaluate the ideas as 'good', 'bad' and 'in-between'. Similar numbers of children engaged in the reflective exercises without experiencing the group task (Condition 2B), experienced the group task without engaging in the reflective exercises (Condition 2C), and experienced neither the group task nor the reflective exercises (Condition 2D). The results suggest that the post-group accretion and appraisal of factors plays little role in the positive effects of group work. The performance of the 2A children during the reflective exercises was indistinguishable from the performance of the 2B children. Although the 2A children progressed more from pre-test to delayed (eight-week) post-test than both the 2B and 2D children, their pre- to post-test growth was equivalent to that of the 2C children. This suggests that the reflective exercises added nothing to what followed from participation in the group task itself.

Study Three examined the possibility that group tasks provide children with frameworks for making productive use of subsequent experiences, and unlike Studies One and Two, this study obtained positive results.

Thirty-six children (Condition 3A) worked in groups on the 1A task, and then viewed demonstrations two, four and six weeks afterwards that had the potential to consolidate their group experiences. Similar numbers of children viewed the demonstrations without participating in the group task (Condition 3B), participated in the group task without viewing the demonstrations (Condition 3C), and neither participated in the group task nor viewed the demonstrations (Condition 3D). Consistent with the possibility that group work primes effective use of subsequent experiences, the children in Condition 3A responded more productively to the demonstrations than the children in Condition 3B, and their subsequent learning was more strongly influenced by those responses. The 3A children also progressed more from pre-test to eight-week post-test than the children in Conditions 3B, 3C and 3D.

Analyses of the dialogues held by the children in Condition 3A suggests that unresolved contradiction during group interaction was a key factor in priming attention to subsequent events. Data obtained during three other studies (including the present Study One) have also been found to support this suggestion (Howe, 2006). Therefore, the main conclusion to be drawn in the presentation will be that when collaborative group work triggers delayed conceptual growth, the likely mechanism is productive use of events experienced afterwards. Such usage is stimulated by unresolved contradiction during the group task itself. This conclusion has important implications for theory and practice, which will be outlined during the presentation.

Simulating one aspect of collaboration and its effects on preparation for future learning

Daniel L. Schwartz[1] and Taylor Martin[2]
[1]School of Education, Stanford University, Stanford, CA 94305-3096, daniel.schwartz@stanford.edu, [2]School of Education, University of Texas at Austin, Austin, TX 78712, taylormartin@mail.utexas.edu

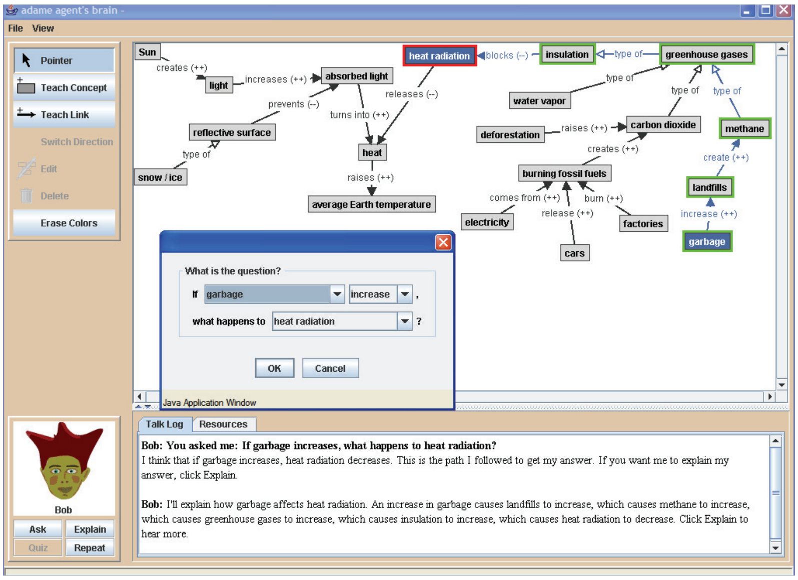
In this paper, we will discuss a technology that we have designed to simulate and capitalize on one of the benefits of collaboration. The beneficial feature we have identified is the "production feedback" that develops when observing one's collaborator use ideas developed during the collaboration. We describe an initial instructional experiment that maximized production feedback through a technology called a Teachable Agent. The results suggest that the use of this technology improved students' readiness to learn related science content when students were no longer working with the technology.

An important educational goal is to help students learn to work collaboratively, because collaboration is a fundamental norm of work in life after school. For example, Barron (2004) found that young adolescents working in groups often let social preferences interfere, so that they missed good ideas that were introduced by less preferred group members. Ideally, appropriate experiences will help students learn to collaborate more effectively in the future. Another important education goal is to create situations that enhance or even simulate the benefits of collaborative work for learning, while avoiding some of the pitfalls. This is the focus of the current work.

The literature on learning through collaboration has begun to identify specific situations where collaboration is particularly effective. For example, Chi, Roy and Haussman (in press) found that collaborative learning works well to the extent that an activity generates "deep questions" for discussion. Schwartz (1995) found that groups provide a special benefit on tasks that involve producing external representation. In an important study, Sears (2006) compared tasks and outcomes for students working alone or together. He found that groups and individuals exhibited no differences in abilities to apply learning after tasks that emphasized practice and checking answers. However, on tasks that emphasized representational production, groups showed a strong benefit over individuals in their readiness to learn new concepts when working alone. Thus, collaborative work seems effective in situations where groups naturally generate new questions, ideas, and representations, and the benefit may occur later when students subsequently have a chance to learn new ideas.

Our focus here is on a second potentially valuable learning characteristic of group work called "production feedback." Production feedback refers to the feedback that arises when one has an opportunity to see how somebody else uses one's ideas. Okita and Schwartz (2006), for example, found that people can learn more by observing their pupil answer questions than by answering those same questions themselves.

We have developed a technology called a Teachable Agent (TA) that can deliver frequent production feedback (Biswas et al, 2005). TA's are a type of pedagogical agent (Baylor, 2007) where students teach the computer, rather than the computer teaching the student. For the TA used in this study – Betty's Brain – students teach by creating a concept map that connects entities (fossil fuels, carbon dioxide) with causal links (make/increase). Betty includes a simple query feature. Using simple artificial intelligence techniques, Betty can visually answer questions posed by the student. For example, in the figure, Betty has answered the query, "What happens to 'heat radiation' if 'garbage' decreases?" based on the map the student taught the her. Betty is intended to help students develop well-structured knowledge that supports complex chains of inference. Betty does not replace instruction, but rather, the technology helps students organize the content of instruction.

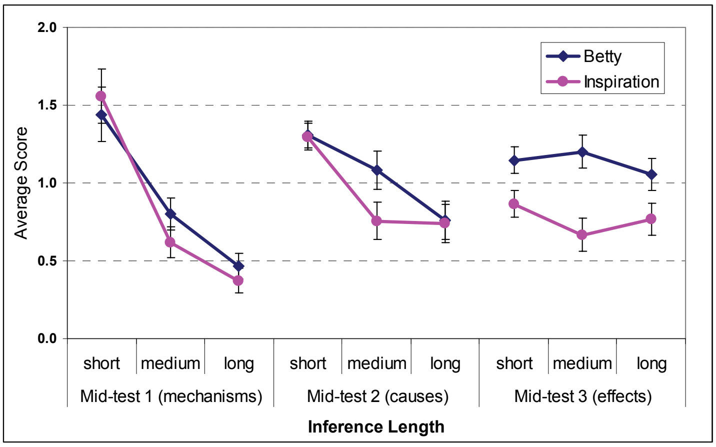


Betty can provide production feedback by including a hidden expert map that can evaluate the quality of Betty's map. For example, it is possible for students to submit Betty take a quiz, or Betty can play a question-answering game and students can wager on whether their agent will give the right answer. We have created a number of affiliated technologies that try to maximize the benefits of production feedback; for example, the teacher can project all the agents at the front of the class and ask a common question. The agents demonstrate their reasoning, and the teacher can lead a discussion over the different answers and agent representations.

We conducted an initial study to determine whether the inclusion of production feedback improved learning, and whether it improved students' preparation for learning new topics later on when no longer using the technology. Our thought was that having students observe their agents answer questions would help them develop a more integrated representation of cause relations, because they would anticipate and follow their agent's reasoning, instead of just given a "factual answer" themselves. The study involved two conditions. In the Betty condition, students created concept maps by teaching Betty, and Betty answered questions in a variety of settings to generate production feedback. In the Inspiration condition, students created concept maps with the commercial software Inspiration (which they regularly used in class). The key difference was that the students in the Inspiration condition answered the questions in a variety of settings themselves. Thus, they received direct feedback to their own answers, compared to the Betty condition where the feedback was directed at the agent.

Two classes of 6th-grade students (n=58) participated. The intact classes were randomly assigned to either the Betty or Inspiration condition. A researcher taught both classes. Eleven days of instruction and testing were spread over three weeks. The main topic of instruction, global warming and climate change, was divided into three units: mechanisms, causes, and effects. Both groups received an identical mix of hands-on experimentation, reading, watching video clips, concept map construction, lecture, and discussion. Participants received one homework assignment per unit. At the end of each unit, participants received an eight-question test on the content from that unit (and its integration with prior units if relevant). The questions were all short answers that varied in complexity, with questions requiring short, medium, or long chains of inference to answer. Questions were not in a form that the TAs could answer, and often times, they required the integration of prior knowledge with current instruction (e.g., how will a floor mirror effect the temperature in a greenhouse?). Answers were scored from a low of 0 points to a high of 2 points.

The figure shows the test results broken out by condition, unit test, and the length of inferential chain needed to answer the question. At the first unit-test, the Betty students show a very modest advantage for the longer inferences. At the second unit-test, the Betty students show a strong advantage for the medium-length inferences. By the final unit-test, the Betty students show an advantage for short, medium, and long inferences, $p < .01$. One interpretation of this pattern is that the Betty students were getting progressively better at reasoning about longer and longer chains of inference. This may be due to the increase in production feedback technology at each unit, or it may simply be a cumulative effect of working with Betty for longer, regardless of feedback. A separate study, one that uses a different design and more than two intact classes, is planned to address this question. Nevertheless, the steady improvement in length of inference is exactly what one would expect the TA software to yield, because this is what Betty models and enforces.



After the intervention was completed, we also gave students from both conditions a new assignment to learn about what they could do to stop global warming. They received a relevant reading passage, and their task was to make a paper-and-pencil concept map. They received four starter nodes. This latter assignment was actually a test of students' preparation for future learning when no longer using their respective technologies. Students in the Betty condition integrated roughly 2.5 concepts from the passage versus 1.0 for the Inspiration condition, $p < .01$. Concept maps are often used to assess student understanding (Ruiz-Primo, et al., 2001), and these maps indicated that the production feedback from the earlier lessons prepared students to develop a more complete and integrated understanding of the passage.

The intent of the current work has been to embody a positive feature of productive collaborative interactions into technology. The goal has not been a severe experimental test of the hypothesis that production feedback is an important ingredient. This has been done elsewhere (Okita & Schwartz, 2006). Instead, the goal was to see if we can design sociable technologies that capture some of the hypothesized positive aspects of collaboration, for those situations when students cannot collaborate (e.g., homework in the U.S. is typically done alone). We wanted to see if there were benefits of the technology that extended beyond the time when interacting with the technology. In this case, we found that if students learn the initial casual content in an integrated way, they will be more prepared to learn in the future (Bransford & Schwartz, 1999). There are alternative interpretations for the effects, so additional research will be necessary. In the meantime, the goal of putting collaborative learning theories to test by embodying them in technologies seems like a worthwhile direction for the study and design of instruction.

References

Asterhan, C. S. C., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology, 99,* 626-639.

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues and development of scientific reasoning. *SocialDevelopment, 2,202—221.*

Baker, M. (2003). Computer-mediated interactions for the co-elaboration of scientific notions. In: J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning_environments.* Utrecht: Kluwer Academic Publishers.

Barron, B. J. (2004). When smart groups fail. *Journal of the Learning Sciences, 12,* 307–359

Baylor, A. (2007). Introduction to the special issue on pedagogical agents. *Educational Technology, 47(1)*, 3-4.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N. & TAG-V (2005). Learning By Teaching: A New Agent Paradigm for Educational Software. *Applied Artificial Intelligence, 19*, 363-392.

Bransford, J.D. & Schwartz, D.L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Educational Research, 24,* 61-100.

Brown, A.L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J.C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 188—228). Cambridge: Cambridge University Press.

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (in press). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science.*

Damon, W. (1984). Peer education: The untapped potential. *Journal of Applied Developmental Psychology, 5,* 331-343.

Damon, W., and Phelps, E. (1989). Critical distinctions among three approaches. In N. M. Webb (Ed), *Peer interaction, problem-solving, and cognition: Multidisciplinary perspectives*. New York: Pergamon Press.

De Vries, E.,, Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue : Explanation and argumentation as vehicles for understanding scientific notions. *Journal of the Learning Sciences, 11*, 63-103.

Dorfman, J., Shames, V.A., & Kihlstrom, J.F. (1996). Intuition, incubation and insight: implicit cognition in problem solving. In G. Underwood (ed.), *Implicit cognition* (pp. 257-296). Oxford: Oxford University Press.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38,* 39-72.

Howe, C. (2006). Group interaction: and conceptual understanding in science: co-construction, contradiction and the mechanisms of growth. Paper presented at BPS Developmental Psychology Section, Royal Holloway College, London.

Howe, C., McWilliam, D. & Cross, G. (2005). Chance favours only the prepared mind: incubation and the delayed effects of peer collaboration. *British Journal of Psychology,* 96, 67-93

Howe, C., & Tolmie, A. (1998). Productive interaction in the context of computer-supported collaborative learning in science. In K. Littleton & P. Light (Eds.), *Learning with computers: Analysing productive interaction* (pp. 24—45). London: Routledge.

Howe, C. J., Tolmie, A., & Rodgers, C. (l992). The acquisition of conceptual knowledge in science by primary school children: Group interaction and the understanding of motion down an inclined plane. *British Journal of Developmental Psychology, 10,* 113—130.

King, A. & Rosenshine, B. (1993). Effects of guided cooperative questioning on children's knowledge construction. *Journal of Experimental Education, 61*(2), 127-148.

Kruger, A.C. (1993). Peer collaboration: conflict, cooperation or both*? Social Development, 2,* 165-182.

Limon, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning & Instruction, 11,* 357-380.

Mugny, G. & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8,* 181.

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28*, 384-395.

Okita, S. Y., Schwartz, D. L., (2006) When observation beats doing: Learning by Teaching. Proceedings of the *7th International Conference of the Learning Sciences* (ICLS) Bloomington, IN.

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278

Schwarz, B. B., & Asterhan, C. S. C. (in press). Argumentation and Reasoning. To appear in: K. Littleton, C. Wood, & J. Kleine Staarman (Eds). *Elsevier Handbook of Educational Psychology*: *New Perspectives on Learning and Teaching*. Elsevier.

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right...if they argue together! *Cognition & Instruction, 18*, 461-494.

Schwartz, D. L. (1995) The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, *4*, 321-354.

Sears, D. A. (2006). Effect of innovation activities on dyads. Unpublished dissertation. Stanford University, Stanford, CA.

Teasley, S.D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 31*, 207-220.

Tolmie, A., Howe, C.J., Mackenzie, M., & Greer, K. (1993). Task design as an influence on dialogue and learning: Primary school group work with object flotation. *Social Development, 2,* 183-201.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495—552.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Acknowledgements

The research presented by Asterhan and Schwarz was supported by Israel Foundations Trustee grant 27.35 to Christa Asterhan. The research outlined in Howe's paper was supported by ESRC award R000239092. The material in the paper by Schwartz and Martin is based upon work supported by the U.S. National Science Foundation and Department of Education under Grants No. 0354453, 0633856, and R305H060089. Any opinions, findings, and conclusions or recommendations expressed in the material are those of the authors and do not necessarily reflect the views of the funding agencies.