ELABORATIVE AND CRITICAL DIALOG: TWO POTENTIALLY EFFECTIVE PROBLEM-SOLVING AND LEARNING INTERACTIONS

by

Robert G.M. Hausmann

B.A., Bowling Green State University, 1998

M.S., University of Pittsburgh, 2001

Submitted to the Graduate Faculty of

Arts and Science in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Robert G.M. Hausmann

It was defended on

July 5, 2005

and approved by

Dr. Jonathan W. Schooler

Dr. Christian D. Schunn

Dr. Kevin Crowley

Dr. Michelene T.H. Chi
Dissertation Director

ELABORATIVE AND CRITICAL DIALOG: TWO POTENTIALLY EFFECTIVE PROBLEM-SOLVING AND LEARNING INTERACTIONS

Robert G.M. Hausmann, PhD

University of Pittsburgh, 2005

Recent research on learning individual monologs and collaborative problem solving suggests that students learn best when they are required to be active participants in interactive dialogs. However, some interactive dialogs are more conducive to learning than others. Two dialog patterns that seem to be effective in producing successful problem solving and deep learning are elaborative and critical interactions. The goal of the present study is to evaluate the relative impact of each dialog on learning and problem solving by experimentally manipulating the types of conversations in which dyads engage.

Undergraduate participants were randomly assigned to one of four conditions: a singleton control, a dyadic control, an elaborative dyad, or a critical dyad. The domain chosen for the experiment was a bridge optimization task in which individuals or dyads modified a simulated bridge, with the goal of making it as inexpensive as possible.

Both problem solving and learning from the simulation were assessed. Performance on the task included a combination of two factors: the quality of the design and the price. Overall learning was measured by the gain from pre- to posttest on isomorphic evaluations, and was further decomposed into text-explicit and inferential knowledge. The results suggest elaboration is easier to train and led to stronger problem solving and learning than the control condition, whereas the critical interactions were more difficult to instruct and led to problem solving and learning equal to the control condition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

My friend H. once said, “Science is hard.” This became our mantra during the ABD years. Science, however, need not be hard when you enlist the help of your friends and colleagues. I am indebted to a great many individuals for their assistance throughout the duration of this project.

First, I would like to recognize the financial assistance I received. The first was in the form of an award from the American Psychological Association. The second was in the form of a grant from the Pittsburgh Science of Learning Center, funded by the National Science Foundation (award number: SBE-0354420). I am extremely grateful to both agencies for selecting my project to support. I would also like to recognize COL Stephen J. Ressler and the West Point Bridge Designer software team for writing bulletproof software. Not once did their program crash during the entire duration of this experiment.

Second, I would like to thank all of the day-to-day support I received from the hard working folks at the Learning Research and Development Center, especially the computing services. Gary Wild deserves a special thanks for loaning me some equipment for longer than was anticipated.

Third, I greatly appreciate the hard work and insightful feedback from my committee members: Jonathan Schooler, Chris Schunn, Kevin Crowley, and Micki Chi. Their feedback provided the much-needed varnish for this project. Also, I am deeply indebted to Micki for taking me into her lab and constantly inspiring me to work beyond my comfort zone. When this happened, great things were accomplished.

Finally, I would like to thank: Ryan Tweney, M. Neil Browne, and Stuart Keeley for helping me to get into graduate school; all of my friends, Steve Jones, John Connelly, Mark McGregor, Tim Smith, and H. Chad Lane, for helping me stay in grad school; my parents and brother for their undying support; and most especially, my beloved wife. Thank you for reading several drafts, offering your sage advice, and providing the reliability for the assessments. Without you, *none* of this would have been even remotely possible.

1. INTRODUCTION

Certain verbalizations have been shown to lead to the modification of the underlying representation held by the speaker. One such verbalization, *self-explaining*, has been shown to be particularly effective in bringing about representational change during procedural and conceptual learning (Chi, Bassok, Lewis, Reimann, & Glaser, 1989a). Self-explaining might be usefully conceptualized as a personal monolog that results in an increased understanding of a domain through the modification of the underlying mental representation (Chi, 2000). The *self-explanation effect* seems to be a domain-general learning strategy, given it has been shown to be useful for learning physics (Chi & Bassok, 1989; Chi et al., 1989a; Chi & VanLehn, 1991), the human circulatory system (Chi, DeLeeuw, Chiu, & LaVancher, 1994; Hausmann & Chi, 2002), probability (Renkl, 1997), calculating interest rates (Renkl, Stark, Gruber, & Mandl, 1998), Lisp programming (Pirolli & Bielaczyc, 1989), algebra and geometry (Aleven & Koedinger, 2000), and even how to use spreadsheets (Reimann & Neubert, 2000).

Self-explaining is certainly not the only type of monolog that has been linked to learning. Other types of verbalizations have also been shown to facilitate the acquisition of new information, which include summarizing (Palincsar & Brown, 1984), self-questioning (Wong, 1985), generating precise elaborations (Stein & Bransford, 1979), and elaborative interrogation (Seifert, 1993).

On the other hand, there are types of monolog that are *not* associated with learning. For instance, in the same study that demonstrated support for the self-explanation effect, Chi et al. (1989a) found that paraphrasing was not correlated with learning (see also Hausmann & Chi, 2002). Similarly, merely repeating facts (O'Reilly, Symons, & MacLatchy-Gaudet, 1998) and underlining (Seifert, 1993) have also been shown to be ineffective in producing learning gains.

What empirical regularities arise from the studies linking individual speech with learning? It seems apparent that verbalization, by itself, does not lead directly to learning. Instead, the pattern of results suggests that only certain types of monologs are associated with learning. What are the commonalities among the types of monologs that correspond to learning? The shared feature among facilitative monologs is modification of the student's underlying representation, either through the generation of new knowledge or the reorganization of prior knowledge. For present purposes, these types of monologs will be referred to as *constructive monologs*. Monologs might be considered constructive under three conditions. A monolog might be considered constructive when: a) the student processes the information at a semantic level (Craik & Lockhart, 1972), b) it creates connections between the target material and prior knowledge, or c) it generates new information or modifies prior knowledge.

While the research on constructive monologs has resulted in strong learning gains, so too has research on dialogs. Several studies have shown that collaborative problem solving and learning is superior to individual problem solving and learning. Several of those studies are reviewed elsewhere (E. G. Cohen, 1994; Dillenbourg, 1999; Hill, 1982; Webb & Palincsar, 1996). Estimates of the effect sizes for the impact of collaborative learning range from $\sigma = 0.21$ (Slavin, 1990) to $\sigma = 0.88$ (Johnson & Johnson, 1992). A targeted review is presented below to highlight the effects of collaboration.

Phelps and Damon (1989) observed fourth graders' performance on mathematic problems across two years. Specifically, they were interested in the children's ability to reason about spatial problems (i.e., model-copying and spatial-perspective tasks) and mathematical problems (i.e., rote math and proportions). Phelps and Damon argued that the effects of peer collaboration were most helpful for tasks that involved problem solving. They found an advantage for collaboration for the spatial-reasoning tasks, but not the rote math problems or a model-copying task. This suggests that collaboration is especially helpful for difficult problems.

Okada and Simon (1997) found similar results for a scientific-discovery task. They contrasted dyadic and individual scientific discovery on a computer simulation of a molecular genetics laboratory (Dunbar, 1993). Participants were asked to discover gene regulation either independently or collaboratively. Using a nominal groups analysis (i.e., non-interacting groups matched for the number of members), Okada and Simon found dyads were more successful in finding the correct inhibition hypothesis than the nominal individuals.

These studies suggest that collaboration can be an important tool for learning and problem solving. To understand why there is an advantage for collaboration, research has turned to the interactions in which the groups engage. Several types of dialog have been shown to be effective, but the present study will focus exclusively on the interactions between relative novices of approximately symmetrical knowledge. In other words, interactions between non-peers such as tutorial dialogs, which have been considered elsewhere (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001), will not be reviewed here. In the next two sections, an argument is made for the inclusion of critical and elaborative interactions as constructive dialogs.

1.1. THE IMPACT OF CRITICAL INTERACTIONS

One type of dialog that might be considered constructive is interactions that contain an element of conflict. Conflict, defined here, is the discrepancy between an individual's expectation and what is actually observed. The source of conflict can arise from an individual observing an empirical result or through interacting with another person. Certain types of instruction rely on conflict as a pedagogical technique. Physics instructors sometimes use demonstrates to provoke conflict in their students' minds. For example, when students are confronted with the following scenario, "*A hunter with a blowgun is hunting for monkeys. He sees one hanging from a tree. He knows that a monkey will always drop from its branch the moment a hunter fires his dart. How does the hunter aim his gun to make sure he hits the monkey?*" they often mistakenly answer in the middle or the base of the tree. The instructor can use the differences in predictions to help the students identify their faulty assumptions (Hatano & Inagaki, 1991). Thus, interacting with the individual might also serve as a source of conflict.

1.1.1. Definitions of critical interactions

A subset of constructive dialogs might contain elements of conflict. At least three different classes of interactions have been introduced in the literature that contain conflict. The first class of interactions is *socio-cognitive conflict* (Druyan, 2001; Kruger, 1993; Tudge, 1989). Socio-cognitive conflict is derived from the Piagetian tradition in which children are observed interacting with other children of higher (or lower) levels of development. Development, according to Piaget, progresses in discrete steps from intuitive understandings based on perceptual cues to formal operations that rely on advanced reasoning. The implication is that children can be classified into discrete categories, based on their present developmental status. For example, children who are classified as *conservers* understand that physical manipulations do not change the quantity of items present. If two rows of checkers are presented to a child and an adult experimenter spreads one row of checkers out, a conserver will correctly answer that both have the same number of checkers. On the other hand, a *non-conserver* will answer that the longer row contains more checkers. Experiments concerning socio-cognitive conflict often test the effects of pairing conservers with non-conservers and observing their interactions and cognitive outcomes. Socio-cognitive conflicts sometimes lead to the conserving child convincing the non-conserver of the correct answer (Miller & Brownell, 1975), while other times the dialog leads to regressions (Tudge, 1989). The child's ability to localize or not localize the source of the conflict might explain the differences in outcome. Thus, children who ignore the conflict are less likely to learn from their peer.

At the opposite end of the spectrum of development is a second class of interactions, which has gone under the label *argumentation*. Argumentation is defined with respect to a formal structure in which the participants propose a claim ("C") that is backed by a piece of information, or datum ("D), and linked by other information, which is assumed to be true, called a warrant ("W"). This formal argument structure, proposed by Toulmin (1958), and adopted by others (Voss, Tyler, & Yengo, 1983), can be used to model a debate between two (or more) individuals (Keefer, Zeitz, & Resnick, 2000; Leitao, 2003; Pontecorvo & Girardet, 1993; Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993). Argumentation is best suited for modeling dialogs in which the knowledge structures of the individuals is fairly well developed.

For instance, one student could propose an argument, to which another student disagrees and proposes a counter-argument. To be able to articulate a counter-argument, one must have a well-developed understanding of the domain (Leitao, 2003).

An intermediate point between socio-cognitive conflicts and argumentation is a third class of interactions, which will be referred to as *critical interactions*. The term critical interactions has not been used in the argumentation literature but is a useful distinction because it delineates a type of interaction where the goal of the conversation is not to persuade, but to learn a new domain or solve a problem.

Critical interactions can unfold along several different paths. However, one way to organize the trajectories of critical interactions is to consider them in terms of the depth of interaction (see Table 1 for an ordered list of critical interactions). At the most shallow level, an individual can criticize an idea by offering a counter-suggestion without any justification. This is considered the most shallow style of interaction because it does not provide a method with which to improve an idea. A slightly less shallow response is to make an attempt to understand a speaker's idea. The listener may ask the speaker to be more precise by asking for clarification of ambiguous terms. Once the ambiguity has been resolved, the listener can request justification in support of the idea. Asking for a reason is a deeper question than asking for clarification because presumably the listener has understood the idea enough to ask for more information. Once the listener has understood the speaker's idea, as the speaker intended, the listener is then in a position to evaluate the idea. The overarching goal is that that the speaker and listener attempt to produce an idea that is of higher quality than the initial idea.

Table 1. *A linear ordering of the depth of critical interactions.*

Critical Statement	*Goal*	*Depth*
Counter-suggestion	Non-collaborative	Shallow
Clarification question	Understanding	Shallow
Request justification	Understanding	Deep
Evaluation	Diagnose worth	Deep

1.1.2. Empirical evaluations of critical interactions: Problem solving

Although the literatures on argumentation and socio-cognitive conflict are fairly extensive, to the best of the author's knowledge, the effects of critical interactions on problem solving have not been studied directly. However, the effects of the individual components within critical interactions (i.e., counter-suggestion, clarification question, request justification, and evaluation) have been assessed across different studies.

The first goal of critical interactions is to introduce, and make comprehensible, ideas into the conversational space. Studies by Teasley (1995) and Barron (2000a) provide evidence for the existence of this goal. Teasley (1995) compared talking and no-talking individuals and dyads on Klahr and Dunbar's (1988) BigTrak task. BigTrak is a programmable toy that moves forward, turns, and fires a canon. The challenge to the participants was to discover the function of the mystery key, labeled "RPT." Participants worked with BigTrak until they were confident enough to announce the function of the RPT key. To gain a deeper understanding of the processes used to discover the function of the mystery key, Teasley coded the dyads' utterances into several different categories. The category relevant to the current discussion is "checks with partner," in which partners asked clarification questions. According to the coding results, about 7% of the dialog was dedicated to clarifying ideas. Unfortunately, correlations between the number of clarification questions and performance on the dependent measures were not reported, probably because clarification questions were not part of the experimental goals. However, this study demonstrates that interacting dyads actively pursue understanding, and it provides a baseline frequency of utterances associated with the pursuit of understanding against which comparisons can be made.

Barron's (2000a) research illuminated the issue of whether clarification questions translated into better problem solving. In her study, Barron asked high-achieving math students to collaboratively solve a complex story problem. By contrasting successful and unsuccessful triads, Barron found that asking clarification questions was not an effective method of interaction. Specifically, she found that the unsuccessful triad produced more clarification questions than the successful triad. While this difference did not reach traditional levels of significance, it does provide suggestive evidence that asking for clarification may not be an effective method of interaction.

Barron also coded the number of rejections produced by successful and unsuccessful triads. A *rejection* was coded as a response to an alternative solution, which can either be accompanied by a justification or not. She found that the unsuccessful triad produced more rejections than the successful triad. This suggests that merely dismissing an idea can also be detrimental to the success of a collaborative group, possibly because it does not offer any insight as to what the group might change to improve the original suggestion.

Once the individuals in a collaborative problem-solving group have understood the speaker's message, the members can begin evaluating the ideas. Evidence for requesting justification can be found in a study by Okada and Simon (1997). As stated in the introduction, they contrasted nominal individuals with dyads and found that the dyads performed the task more successfully than the nominal individuals. In addition to analyzing the behavioral outcomes, they also coded the individuals' monologs and the dyads' dialogs. They found that the dyads produced more *requests for justification* than the individuals. Although the results are explained by a myriad of factors, request for justification is one component of the interaction that may have led to superior performance on the simulated discovery task.

Vroom, Grant, and Cotton (1969) found positive evidence for the impact of evaluation on problem-solving performance. They chose an idea generation task for their problem-solving domain. The normative model for brainstorming tasks is to divide the session into two parts: generation and evaluation. During the generation phase, members of the group are instructed to generate as many ideas as possible. Critique is supposed to be suspended during the generation phase. Once a threshold of ideas is reached, the evaluation phase begins. During the evaluation phase, members of the group critique each of the ideas.

Vroom et al. manipulated whether groups interacted or not during the evaluation phase. They found that the interacting groups produced ideas of better quality than the nominal groups. Their results suggest that interactive evaluations result in ideas of higher quality than individually produced evaluations.

1.1.3. Empirical evaluations of critical interactions: Learning

In addition to the few studies that have focused on elements of critical interactions and their effects on problem solving, there have been several studies that have examined the outcome of critical interactions for learning. One of the strongest demonstrations of the effects of critical interactions on learning can be found in a study by Schwartz, Neumann, and Biezuner (2000). They specifically paired students on the basis of their incorrect rules for comparing decimals and fractions with the hope that differences in background knowledge would lead to critical interactions.[1] Based on case-study data, Schwartz et al. found that students with different rules were more likely to spontaneously engage in critical interactions and discover the correct fraction rules than students who had the same incorrect rules. This study is noteworthy because it provides evidence that two people, both with incorrect knowledge, can combine their knowledge through the use of critical dialog to produce a correct understanding of the target domain.

Chinn, O'Donnell, and Jinks (2000) found a pattern of similar results.[2] They examined critical interactions and learning in several small groups by measuring the impact of critical interaction on students' scientific writing. They found a significant, positive partial correlation between the quality of collaboratively constructed critical interactions and the outcome measure ($r = 0.37$), while controlling for performance on a transfer problem. The outcome measure was the student's ability to write a conclusion about an experiment they had not previously encountered (i.e., a new transfer problem).

Evidence from a classroom study by Mercer, Wegerif, and Dawes (1999) experimentally manipulated students' interactions. They developed a training procedure for assisting children to talk in productive ways. Their training program, called TRAC (Talk, Reasoning, and Computers), included elements of critical interactions. For instance, children were encouraged to challenge one another, provide reasons for their statements, and discuss alternative ideas. The results showed that students trained with the TRAC program had higher gain scores on the Raven Progressive Matrices task than a group of control students who were exposed to the same curricula, minus the dialog training. Thus, in a classroom setting, they demonstrated that it is

[1] While Schwartz et al. originally used the term "argumentation" in the formal structure sense. Because their students' knowledge was under refinement, the term "critical interactions" will be used to be consistent with the classification of interactions outlined above.
[2] Chinn et al. also used the term "argumentation" in the formal structure sense. They based their analysis of argumentation on Toulmin's (1958) formalization.

possible to train students to converse in particular ways, and that inducing critical interactions can help in the development of general reasoning skills.

Similar effects have also been replicated in the laboratory. In a series of analyses, Chan and her colleagues investigated the effects of prior knowledge and grade level on knowledge-building activities, conflict, and conceptual change (C. Chan, Burtis, & Bereiter, 1997; C. Chan, Burtis, Scardamalia, & Bereiter, 1992; C. K. K. Chan, 2001). The domain they selected was evolution, which is a difficult domain because of the many misconceptions held by children and adults alike (Chi, 2005). Information was presented to students that either maximally conflicted with their prior beliefs, or was supposed to be easily assimilated into their prior knowledge. The researchers coded knowledge-building activities of interacting groups at five levels of depth and found that conflict alone did not produce conceptual change, but was mediated by the knowledge-building activities in which the students engaged. Through a path analysis, the researchers demonstrated that conflict is helpful only insofar as it leads groups of students to engage in deep knowledge-building activities.

1.1.4. A process model of critical interactions

Although the results from the learning studies reviewed above generally show a positive association between critical interactions and learning, there are cases in which being critical does not lead to positive learning gains (Barron, 2000a, 2000b, 2003). Indeed, there are good reasons to suspect that critical interactions do not always lead to learning. Consider the following process model for conflict (see Figure 1, patterned after Posner, Strike, Hewson, & Gertzog (1982); Lee, Kwon, Park, Kim, Kwon, & Park (2003)). There are four stages for the possibility of conflict to lead to learning. The solid arrows represent positive links, while the dashed lines represent negative links leading to the failure of learning from conflict.

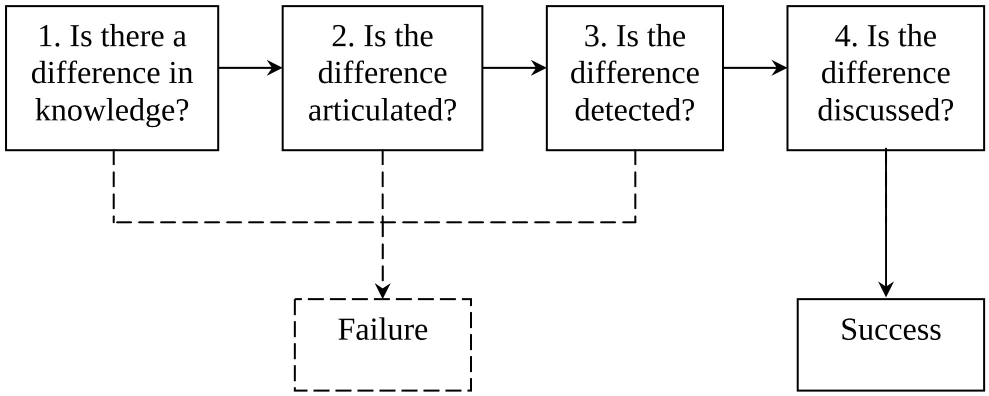


Figure 1. Process model of conflict during collaborative problem solving.

The first stage is a pre-requisite for conflict. Before two individuals can disagree, there must be something to disagree about. Schwartz, Neumann, and Biezuner's (2000) study nicely illustrates this by showing that different prior knowledge is important, even if it is incorrect.

The second stage states that the differences must be articulated or brought into the conversation. Sharing information that is unique to an individual, however, is not guaranteed. Work on group decision making suggests that groups, composed of members of varying levels of expertise, tend not to discuss unshared information (Stasser, Vaughan, & Stweart, 2000). The decision tasks in Stasser's research are constructed such that the unshared information is critical for optimal decisions, often resulting in groups making sub-optimal decisions.

Stage three is of particular interest. In the face of conflicting information, both adults and children often react to conflicting information in striking ways. In the extreme case, some may completely dismiss the conflicting data altogether (Chinn & Brewer, 1998). For example, Chinn and Brewer (1998) have developed a taxonomy of responses to anomalous data. At the lowest level, individuals completely ignore anomalous data. They ignore the data, not because they did not attend to it, but because they do not believe the data are relevant to the present issue. That is, they fail to see the relevance of the anomalous data with respect to their particular theory. If the data are considered relevant, the individual may reject the data because they have an explanatory reason for doing so.

The distinction between ignoring and rejecting anomalous data can be observed in children engaged in a balance-beam task (Karmiloff-Smith, 1988). In the task, the experimenter surreptitiously weighted one end of a block of wood with a lead weight, and the child's task was to balance the block of wood. Very young children (4-5 year olds) were able to complete the task, mostly through proprioceptive feedback from the block of wood. Their behavior was completely driven by a bottom-up, perceptual process; therefore, they were not disturbed by the uneven balance. Slightly older children (6-7 year olds) had a more difficult time with the task because their "theory" of balance states, "items balance on their geometric center." When they attempted to balance the oddly weighted blocks at their center, the older children found it difficult because the wood did not conform to their theory. Thus, when they gave up, claiming the task to be "impossible," they used their explanatory theory to reject the offending data.

Slightly more advanced than completely dismissing conflicting data altogether, some individuals are not aware of the conflict and assimilate the data directly into their background knowledge. For example, children can be categorized on the basis of their mental models of the human circulatory system (Chi, 2000; Chi et al., 1994; Hausmann & Chi, 2002). Some children believe that the heart pumps blood to all the major organs, which includes the liver, kidneys, and lungs (i.e., the single-loop model). They do not understand that the lungs are the sites for reoxygenation of the blood. Thus, when a child reads a statement, such as "*The heart pumps the blood to the lungs*," they understand the sentence to mean that the lungs are going to receive blood as any other organ might. They do not recognize the conflict between their model and the model described in the text. Instead, they assimilate that chuck of information directly into their mental model.

Alternatively, individuals might struggle to resolve the incompatible information by modifying their prior knowledge. In the developmental literature, for example, it has been shown that children attempt to integrate incoming information to be compatible with their current representations. For example, Vosniadou (1994) categorized children's mental models of the earth into six discrete categories. Some children believe the earth is flat like a rectangle. When asked what would happen if they walked in the same direction for days and days, the flat-earth children reliably stated that they would fall off the edge of the earth. As children grow up, they hear that the earth is round, which does not make sense according to their rectangular flat-earth model. To reconcile the conflict between "roundness" with "flatness," the children report that the earth is still flat, but round like a pancake.

Collaborators may react in an analogous way to ideas that conflict with their partner. Partners may fail to learn from each other during collaboration when one partner's statement conflicts with the other's underlying representation, and that conflict is ignored, rejected, or directly assimilated. Instead, a more productive way to interact with conflicting information is for the dyad to attempt to resolve the differences.

Finally, the fourth stage may be necessary for an interaction to lead to any sort of cognitive advancement. It is not enough for two people to "agree to disagree." Instead, the partners must be motivated to discuss their conflicting ideas. The type of dialog the individuals have may determine the quality of their problem solving or learning. Critical dialog can occur at two levels: *shallow* and *deep*.

1.1.5. Critical interactions: Deep versus shallow

Shallow critical interactions are characterized either by disagreeing without any sort of justification or by asking for clarification. Disagreeing without justification might be considered shallow. For example, Chinn, O'Donnell, and Jinks (2000) modeled student arguments as a network of nodes, where the conclusion was the top-level node with each reason represented as individual nodes connected to the conclusion. Conclusions that included several reasons were considered "deeper" arguments than conclusions that were either without justification or only a single reason was provided. They found a significant correlation between argument depth and posttest performance, while controlling for pretest performance. While deeper arguments seem to be related to learning, the converse seems to hold as well. That is, shallow arguments have been shown not to lead to learning (as will be demonstrated below).

Deep critical interactions are characterized either by requesting a reason or evaluating a partner's idea. Requesting a reason suggests that the partner has understood the contribution at a high level but is unable to evaluate the statement without more information. Thus, the listener and speaker work together to make explicit the reasoning behind the speaker's idea. Once a contribution is proposed at a sufficiently high level, the process of evaluation can occur. Evaluation represents the deepest level of reasoning because it is computationally the most complex. To evaluate a contribution, the individual must first comprehend the statement, and then compare it either to information stored in long-term memory or an inference generated from the available information. The comparison must then localize the discrepancy between the two and produce a new contribution that seeks to remove the difference. Of course, evaluative statements themselves can differ in depth. But speaking comparatively, evaluation is a deeper contribution than asking for clarification, disagreeing without justification, or requesting justification.

Interpreting why some students fail to learn from critical interactions becomes clearer when viewed from the perspective of the preconditions (outlined in Figure 1) and the depth of the interactions (presented in Table 1). One might expect critical interactions to lead to learning when students either do not recognize conflict, or treat the conflict by directly assimilating it, rejecting it, or completely ignore it all together.

The problem-solving literature has shown that, even when there are high amounts of conflict, groups can still fail to solve a problem. An example of high amounts of conflict and ineffective problem solving can be found most notably in Barron's study. Barron showed that even groups comprised of high-achieving students can fail to solve problems collaboratively (Barron, 2000a, 2000b, 2003). In a case-study analysis, a successful triad produced more *discuss* responses than an unsuccessful group. Moreover, the unsuccessful group displayed more *ignore-reject* responses than the successful group (Barron, 2003). This suggests that the successful groups, while not immune to conflict, reacted to it through discussion, while the unsuccessful dyads either ignored or rejected the conflicts.

Barron's results accord well with the three reactions to conflicting information outlined above (i.e., completely dismiss, assimilate, or resolve conflict). The unsuccessful triads completely dismissed the conflict, while the successful students struggled to resolve their conflict. Thus, one key feature of successful critical interactions is the development and justification of an alterative proposal. Simply rejecting a statement out-of-hand does not lead to successful problem solving. Thus, there seems to be a curvilinear relationship between critical dialog and learning. Too much disagreement can stifle productive dialog, whereas hasty agreement can be just as ineffective (Keefer et al., 2000).

In summary, there are at least two preconditions for learning from critical interactions. First, students must produce a fully specified idea for another person to understand and detect a discrepancy between the two ideas. Second, the critical dialog should take place at a level where the partners are neither too quick to form a hasty agreement nor antagonistic toward one another. Furthermore, the dialog should be sufficiently deep that learning is most likely to occur. When these preconditions are met, critical dialog should serve as a catalyst for provoking individuals to engage in knowledge-building activities, such as explaining or searching for additional information.

1.2. THE IMPACT OF ELABORATIVE DIALOGS

The problem solving and learning literature demonstrated that critical interactions seem to impart positive consequences for a group when they are produced at a sufficiently deep level (i.e., asking for justification or producing evaluations). Other types of interactions might also lead to learning. Several types of non-critical dialogs have been linked with learning, including exploratory talk (Mercer, 1996), interpretive talk (Teasley, 1995), and cooperative interaction (Forman & Cazden, 1985). Another constructive dialog type that does not contain conflict is elaboration (Brown & Palincsar, 1989). Within an interaction, elaboration can be defined as a conditionally relevant contribution that significantly develops another person's previously stated idea (Hogan, Nastasi, & Pressley, 1999). What are the cognitive implications of elaborative dialogs? Before answering this question, it may be useful to step back and consider how elaboration impacts individual problem solving.

1.2.1. Individual problem solving

The dominant paradigm for studying human problem solving comes from an information-processing approach to cognition (Newell & Simon, 1972). This approach, overly simplified, suggests that problem solving can be conceptualized as a two-step process. The first step, *understanding*, generates a representation of the problem that is useful to the problem solver. The representation includes an initial state, legal problem-solving operators, and finally a goal state. The second step, *search*, is the exploration of the problem space via the iterative application of the problem-solving operators. It is tempting to conceptualize the two-step process of problem solving as a serial process: The understanding process occurs first and produces a useful problem representation, upon which the search process takes over and runs until the problem is solved. However, during actual problem solving, the individual often pauses to check and potentially modify his or her representation (VanLehn, 1989).

If the problem is novel to the problem solver, the understanding process begins when the solver reads the problem statement (Hayes & Simon, 1979). For well-defined problems, all the information needed to solve the problem is included in the problem statement. The understanding process requires two sub-processes to operate, which serve to generate a fully functional problem representation. The first sub-process is the *language* process. The language process parses the linguistic input syntactically, then assigns semantic labels to the constituents, and finally integrates the representation into a coherent whole.

The language process passes the output (i.e., the deep structure of the text) to the second sub-process: the *construction* process. The construction process then augments the representation by supplying additional information. The process by which the additional information is added to the initial representation can be understood from a text-processing perspective. For instance, in the context of prose comprehension, elaborations serve to create connections between sentences, create expectations of the text, detect anomalies, and increase the retention of the text (Reder, 1980). The elaborations are generated from one's background knowledge. Thus, the same processes that are used during text comprehension (i.e., elaboration) might also operate when augmenting a problem representation.

An extreme example of augmenting one's problem representation can be found in a series of studies by James F. Voss and colleagues (Voss, Greene, Post, & Penner, 1983; Voss, Tyler, &

Yengo, 1983). They contrasted ill-defined problem solving of domain-relevant experts (i.e., political scientists) with experts of a domain that was not relevant to the current problem (i.e., chemists). Political scientists, when asked how they might increase crop production in the Soviet Union, dedicated 24% of their think-aloud protocol to developing the problem representation. In contrast, only 1% of the chemists' protocols were dedicated to the problem representation. The only difference between the two populations was the amount of their domain-relevant background knowledge, and this difference in prior knowledge translated into different problem solving processes. That is, experts used their vast background knowledge to augment the problem representation, which can be measured in their think-aloud protocols.

Once the problem solver has constructed a problem representation at a high enough level of specification, the search process can begin. During the search process, two findings come into play. First, the quality of the problem representation has a measurable impact on the probability of finding a successful solution (Hayes & Simon, 1977). Second, the solver will return to the understanding process when the individual detects a contradiction or runs out of things to try (Hayes & Simon, 1979).

1.2.2. Empirical evaluations of individual elaboration

What are the implications of elaborative monolog on learning? The cognitive effects of elaboration have been studied from three different perspectives: memory, text comprehension, and learning. Studies that have addressed the effects of elaboration are presented below.

1.2.2.1. Elaborative effects on memory. In the context of memory, elaboration serves to increase retention by adding partially redundant retrieval cues (Reder, 1980; Stein & Bransford, 1979). Stein and Bransford (1979) argue that not all elaborations are equally effective. They contrasted memory for sentences in two different conditions. In the first condition, sentence stems, such as "*The fat man read the sign*" was elaborated with an "imprecise" ending, "*that was two feet tall*." In the second condition, the same sentence stem was "precisely" elaborated with a slightly different ending, "*warning of thin ice*." The results indicated that precisely elaborate sentences were easier to recall than imprecise elaborations. The difference in the precision of elaborations suggests that quantity alone may not be important, but the quality has an effect on item strength.

Seifert (1993) provides evidence that the effects of elaboration generalize to entire textual passages. Seifert experimentally manipulated the types of elaborative activities in which his participants engaged. Students were instructed to learn from a text about animal characteristics and adaptation in three conditions. The first condition was instructed to underline the main ideas presented in the text. The second condition was also instructed to underline, but then the text contained an extra sentence that was an elaboration of the paragraph. In other words, the elaboration was *given* to the students. The third experimental condition read the same passage, but a simple prompt to elaborate the contents of the passage was in the adjacent column (i.e., "why"). This was the elaborative interrogation condition. Seifert found that producing elaborations helped individuals outperform the other conditions on memory for the main idea presented in the paragraph.

1.2.2.2. Elaborative effects on text comprehension. In the context of text comprehension, elaborations help increase comprehension and retention of the text (Reder, 1980). The elaborations that are formed during reading are most likely idiosyncratic because they are generated from an individual's prior knowledge (J. R. Anderson & Reder, 1979; Reder, 1980; Slamecka & Graf, 1978; Stein & Bransford, 1979). In addition to being idiosyncratic, some elaborations are easier to generate than others. For example, consider the following text: *John went into a restaurant. He ordered a salad and iced tea.* A simple elaboration would be to infer: `He = John` (i.e., anaphor resolution). A more complex elaboration might include: `John is on a diet`. Simple elaborations are generated automatically, while more complex elaborations require effortful processing.

Different types of elaborations may lead to different types of learning. For instance, Hamilton (1997) conducted a study in which he contrasted different types of elaborative statements while processing a text. He asked undergraduates to read a text on the topic of positive and negative reinforcement. While reading, the participants were instructed to generate different types of elaborations: create personal examples from the materials (integrative elaborations), contrast the ideas raised in the text (relational elaborations), or expand on the effects of the ideas (elaborative interrogation). Consistent with the memory research, relational elaborations produced the best performance on a recall of the definitions in the text. Furthermore, Hamilton found that the relational elaborations produced the best performance on solving application-type questions.

Implicit in the discussion of the aforementioned studies is some inferential mechanism that generates elaborations. How, then, are inferences generated while reading? At least five candidate inferential mechanisms might help explain how elaborations are generated. They include: inferences from the simulation of a mental model (Norman, 1983), inferences from category membership (Chi, Hutchinson, & Robin, 1989b), inferences from analogical reasoning (Markman, 1997), inferences from the integration of the situation and text model (Graesser, Singer, & Trabasso, 1994), and inferences from logical reasoning (Rips, 1990). A simple associative mechanism may also account for the production of elaborations.

1.2.2.3. Elaborative effects on conceptual learning. The studies relating elaboration to memory and text processing also seem to hold for learning conceptual domains as well. For example, van Boxtel, van der Linden, and Kanselaar (2000) investigated the effects of collaboratively producing a concept map in one experimental condition or a poster in a different condition. For the concept-mapping condition, they found a significant correlation between the definition posttest score and elaborative episodes (r = 0.83). The definition test represented shallow learning and the elaborative episodes did not correlate with a deeper measure of learning (i.e., essay questions). One way to explain their results is to look at the task demands. Concept maps lend themselves to discussions of the links between individual concepts, which suggests the individuals would have formed detailed representations of the concepts (i.e., definitions) through elaborative activities.

Stark, Mandl, Gruber, and Renkl (2002) extend these results by providing evidence of deep learning. They trained their participants to engage in elaborative behaviors. Individuals were trained to give elaborations while studying worked-out examples. After the training, participants then applied their study technique to the examples. According to their think-aloud protocols, the individuals were categorized according to different elaboration profiles. Stark et al. found that individuals who displayed an active, meta-cognitive orientation to the elaboration activity performed better on deep, far transfer problems than individuals whose profiles were characterized as more passive and superficial. Their finding is encouraging because it suggests that elaborative activities can lead to deep learning outcomes when the elaborative activities are trained.

1.2.3. Collaborative problem solving

Returning to the question of how elaborative dialog impacts problem solving, the same processes that operate during individual problem solving might also operate during collaborative problem solving. For example, members of a dyad may initially construct their own, independent representations of the task. Once they begin working on the problem, however, two conditions may arise that require them to modify their problem representations. First, the pair may detect a contradiction, either between each other or themselves and the task domain. Second, they may run out of things to try (Azmitia & Crowley, 2001). If either of these two conditions is satisfied, then the pair may be prompted to modify their problem representation. However, instead of independently modifying and augmenting their individual representations, the pair will work together to augment their shared understanding of the task by elaborating each other's statements (Teasley & Roschelle, 1993).

Another way in which elaborations may play a role in collaborative problem solving is to serve as a method for coordinating the dyad's actions by displaying evidence that the listener has understood the partner's contribution. During communication, "[t]he contributor and his or her partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes" (H. H. Clark & Brennen, 1991; p. 129). Displaying evidence of accepting a partner's contribution is a continuous function, with low levels of evidence on one side of the continuum (e.g., continued attention) and high levels on the other (e.g., collaborative completion). Finishing another person's statement demonstrates a high level of acceptance. Elaborative statements are sometimes stated as collaborative completions. Consider the following exchange from a banking simulation (see Table 2, from McGregor & Chi, 2002). Cass presents part of a contribution, which her partner finishes. Dana's completion signals to Cass that she has understood and accepted Cass's contribution.

Table 2. *Example of a socially distributed elaboration*

		Contribution	*Propositional Representation*
1	Cass:	Okay, the new system would give the- give the employees...	GIVE(SYSTEM, EMPLOYEES, …)
2	Dana:	more time to deal with the customers.	DEAL(TIME, CUSTOMERS)

Another function of an elaboration is to produce an utterance that satisfies Grice's maxims, two of which are relevant here. First, the maxim of *quantity* states that the individual should produce a contribution that is as informative as it needs to be, and no more than necessary. However, this is balanced by the second maxim, the maxim of *manner*, which states that the contribution must avoid ambiguity. Therefore, partners should optimally produce statements that are short, yet informative and unambiguous. Elaborative statements help satisfy these constraints by distributing the specification of the intended message across the speakers. To illustrate how this is accomplished, let us consider Cass and Dana's exchange further. If we represent each contribution propositionally, we see that Cass leaves one variable unassigned (i.e., a low level of specification). The verb *to give* requires the assignment of three variables: 1) a subject, 2) a predicate, and 3) a direct object. Cass has specified the first two variables (i.e., the subject and predicate), while Dana supplies the third (i.e., the direct object). Thus, elaborative dialogs in the context of problem solving may serve to produce efficient speech by distributing the process of variable assignments over speakers.

1.2.4. Empirical evaluations of collaborative elaborations

What are the implications of collaborative elaborations on learning? The cognitive effects of elaboration have been studied from various perspectives. Studies that have addressed the effects of elaboration on memory and learning are presented below.

1.2.4.1. Elaborative effects on memory. Very few studies have been conducted on the impact of collaboratively generated elaborations in the context of memory. O'Donnell et al. (1985) provides evidence that the quantity of elaborations has an impact on recall. They instructed their participants to elaborate either once or four times during the study phase. They found that groups that were instructed to produce multiple elaborations learned their lists better than those who elaborated only once. This suggests that the quantity of elaborative statements may be important for memory.

If we expand the definition of "collaboratively generated elaborations" to include outside influences, then there has been some research on contrasting self-generated elaborations with elaborations presented by the experimenter. One such study found that self- versus other-generated elaborations tends to depend on the prior knowledge of the individual (Kim & Van Dusen, 1998). That is, students with high prior knowledge tend to perform equally well when the elaborations are either self- or experimenter generated. Low prior knowledge students, on the other hand, tend to recall more information when the experimenter provides the elaborations for the student. This makes sense when we consider the source of the elaborations is generated from one's background knowledge (Chiesi, Spilich, & Voss, 1979). Thus, low prior knowledge students have a limited capacity to generate their own, high-quality elaborations.

1.2.4.2. Elaborative effects on conceptual learning. In the context of conceptual learning, Hogan, Nastasi, and Pressley (1999) reported positive effects of elaborative dialog. They classified small-group dialogs along three dimensions. On the lowest end, they coded a dialog as *consensual* when there was a low amount of mutual engagement (i.e., when only one member of the group contributed ideas). When two speakers were involved, they coded the dialog as *responsive*. When more than two speakers made contributions that clarified, built upon, or corrected each other, the dialog was coded as *elaborative*. They found that elaborative dialogs were more strongly related to higher levels of reasoning complexity than the other two patterns (i.e., consensual and responsive).

Similar results have been found for learning in a hypertext tutoring system. Pairs of students learning about glaciation were categorized as “high” and “low” learners according to their pre- to post-test change scores. High learners generated nearly twice as many elaborations as the low learners (A. Anderson, Mayes, & Kibby, 1995). This result is merely suggestive because other differences were also observed between the high and low pairs. For instance, high pairs also engaged in more cognitive conflict and more summarizing activities. Therefore, elaboration may not be the sole reason for the gain scores, but it may be a significant contributor.

One method that avoids the contaminating effects of other activities is to experimentally manipulate the types of dialogs in which students engage. For instance, Larson et al. (1985) trained dyads to interact using either meta-cognitive activities or elaborative activities. They found that the elaborative group more accurately recalled information from a science text than the meta-cognitively trained dyads. This result accords well with the memory and text comprehension literature.

In summary, elaborative dialogs are hypothesized to serve several functions during collaborative problem solving. First, elaboration may help dyads modify their problem representation as they work together on a task. During individual problem solving, singletons are often observed augmenting their problem representation through the application of background knowledge, especially when the problem is ill-defined. Second, elaboration may serve to coordinate actions by producing high-levels of evidence for accepting a contribution. Furthermore, they serve to increase the specification of individual contributions through the articulation of unfilled variable assignments, as was illustrated in the banking simulation dialog. Finally, elaborative statements may increase the memory strength for particular items when both a high number and quality of elaborations are generated. Producing elaborations has also been linked to increased reasoning complexity.

2. THE DIFFERENTIAL IMPACT OF CRITICAL AND ELABORATIVE DIALOGS

A few questions remain from the literature review. The results from existing research are typically correlational, thereby making it difficult to assess the causal impact of elaborative and critical interactions on problem-solving performance and learning. Moreover, the effect of each type of interaction on deep and shallow learning is typically not assessed. Finally, there are very few empirical studies that *directly* compare the effects of different types of dialog on both problem solving and learning.

To gain a more direct contrast between the impact of critical and elaborative dialog on problem-solving performance and learning, the following study experimentally manipulated the types of collaborative interactions in which dyads engaged. The study attempted to address the following research questions: (1a.) Is it possible to train collaborative partners to interact in a specific way? (1b.) Does interaction training have a measurable effect on problem-solving performance and/or learning? (2.) Do critical and elaborative interactions lead to quantitatively different levels of problem-solving performance and learning relative to a control condition that does not receive communication training?

A singleton condition was also added so that the collaborative results could be compared against individuals working in isolation. Although the main focus of the current research is to contrast the outcomes of critical and elaborative interactions, it is still important to establish whether interactions of any kind leads to improved problem-solving performance and/or learning. A review of the collaborative problem-solving literature produces mixed results regarding the efficacy of collaborative problem solving. Clearly, groups do not always outperform the same number of non-interacting individuals. Instead, the effects of working with a partner depend on the type of task. Even within the same problem, however, there is a large amount of variation in group problem solving. Several proposed hypotheses, both social and cognitive, attempt to explain the decrements in problem-solving performance. For example, the social psychology literature has proposed several hypotheses about why there is a decrement in performance, including groupthink (Janis, 1982); free riding, evaluation apprehension, and mutual production blocking (Diehl & Stroebe, 1987); hidden profiles (Stasser et al., 2000); social loafing, group polarization, and coordination problems (Oyster, 2000).

The literature investigating the cognitive components of collaboration, on the other hand, has attempted to explain why groups perform well compared to individuals. Some hypotheses that have been proposed to explain why groups outperform individuals include superior memory capacity in the form of recognition memory (S. E. Clark, Hori, Putnam, & Martin, 2000), transactive memories (Moreland, Argote, & Krishnan, 1996), and cross-cuing (Meudell, Hitch, & Boyle, 1995; Tulving & Pearlstone, 1966). Groups have internal error correction and monitoring by rejecting incorrect proposals (Laughlin, Bonner, & Miner, 2002; Laughlin, Zander, Knievel, & Tan, 2003) or by creating different or more abstract representations (D. L. Schwartz, 1995; Shirouzu, Miyake, & Masukawa, 2002). Groups may also have multiple representations, which may increase the probability of finding a solution (Hill, 1982; Moreland & Levine, 1992). Both the social and cognitive explanations intimate that sometimes collaborative problem solving is more effective than individual problem solving, while other times it is not. To shed further light on this pattern of results, the present research will focus on contrasting group and individual performance on a problem-solving task.

The first research question asks if group dynamics can be scripted through interaction training. That is, another way to look at group performance, which is not concerned with the effects of the group on the individual, is to take the group as an interacting unit. The interaction perspective, which combines both the social and cognitive processes, suggests that certain types of verbal interactions are more effective than others. Two types of interactions have already been reviewed thus far (i.e., conflict-driven and elaborative dialogs). The interaction perspective suggests that one way to increase group performance is to script or scaffold the partners' interactions.

Prior research suggests that, not only is it possible to manipulate collaborative dialogs, but doing so can have beneficial effects. Interaction training has been conducted for both individual monologs as well as collaborative dialogs. For individuals, Hamilton (1997) and Seifert (1993) both experimentally manipulated the types of elaborative activities in which their participants engaged through short training procedures. Both found positive effects of elaboration on problem solving and learning. Hamilton found a positive impact of relational elaborations on problem solving; and Seifert found that producing elaborations helped individuals outperform a control group on inferential questions.

For collaborative dialogs, Stark, Mandl, Gruber, and Renkl (2002) used a 20-minute training procedure that instructed participants to engage in different types of elaborative activity (i.e., cognitive and meta-cognitive). To train students to produce elaborative interactions, the experimenter first modeled the target behavior. Then participants were asked to engage in the modeled behavior. The experimenter was present to give the participants detailed feedback, as well as answer any questions they might have. The control group was instructed on the procedure for producing think-aloud protocols, thus controlling for any effect that the experimenter may have had on the dialog. Individuals who received elaboration training produced nearly twice the amount of elaborations as the control group. Based on the results of their training procedure, it seems plausible that an experimenter can shape an elaborative monolog.

O'Donnell et al. (1985) provides evidence that dyads can also be trained to produce elaborations in an interactive setting. As mentioned previously, their participants were trained in a particular learning strategy and instructed to elaborate either once or four times during the study phase. Based on the results of the study, it is clear that dyads can be trained to elaborate during a single-session study. No parallel research has been conducted on whether dyads can be trained to engage in critical interactions during a similar time frame. Mercer, Wegerif, and Dawes (1999) were able to train children aged 9-10 to challenge each other using the TRAC program in 9 1-hour sessions; however, their training procedure included more than just criticism. Therefore, the present study will also investigate the differential ease of training elaborative and critical dialogs.

The second research question asks if there is an advantage of one interaction style over another. That is, do elaborative or critical interactions lead to better problem solving and/or learning than unscripted interactions? While the two have not been compared directly, elaborative interactions seem to be favored mainly because there are many points of failure for learning from critical interactions. Furthermore, critical interactions are mostly likely to result in strong learning gains when they take place at a deep level. However, deep, critical interactions may be a more difficult style of interaction because of the need to detect the conflict and struggle to resolve the differences. Elaborative interactions, on the other hand, seem to exhibit fairly consistent results. The results from the memory, text comprehension, and learning literature all seem to suggest that elaborative comments help increase both memory strength and understanding.

2.1. METHOD

2.1.1. Participants

Participants for the current study were recruited from the University of Pittsburgh's undergraduate psychology subject pool. Participants were randomly assigned to condition, with twenty singletons assigned to the individual condition, twenty dyads assigned to the control condition, and twenty dyads assigned to each experimental condition, which yielded a grand total of $N = 140$ participants. However, two dyads were excluded from the study because they did not follow the experimental instructions. Therefore, the final sample size for each condition is as follows: individuals ($n = 20$), control dyads ($n = 20$), critical dyads ($n = 19$), and elaborative dyads ($n = 19$). 94% (128/136) of the sample received course credit for participating, while 6% (8/136) were monetarily compensated.

Dyads were tested in same-sex pairs to control for gender effects (Light, Littleton, Bale, Joiner, & Messer, 2000; Scanlon, 2000). Individuals within a pair were unfamiliar with each other prior to the experiment to protect against the effects of familiarity (Azmitia & Montgomery, 1993; Gould, Kurzman, & Dixon, 1994). Because the hypotheses assume that the participants were relative novices, students who have taken a course in either civil engineering or material science were excluded from participating in the experiment.

2.1.2. Materials

2.1.2.1. Simulation. The domain chosen for the experiment was a design task. The participants were asked to design a virtual bridge using a pre-existing software package (West Point Bridge Designer 2003). The West Point Bridge Designer software was originally created for a competition of high school students interested in civil engineering and was then made available to the public.

The domain of bridge design was selected for several reasons. First, bridge design is an open-ended task that does not have a strictly correct answer, yet it affords several dependent measures that help quantify designs in terms of their quality, which are described below. Second, virtual bridge construction was chosen because it allows for rapid design cycles by eliminating the need to cut and bond traditional materials (i.e., balsawood). Third, it was chosen to control for prior knowledge. Very few undergraduates have studied bridge construction directly. Fourth, pilot testing revealed that it is a highly engaging task. Most undergraduates seem to enjoy the task, as evidenced by their questions at the end of the experiment. They usually wanted to know how their performance compared with other pairs. An engaging task ensures that individuals will interact while solving the problem, which, in turn, allows the experimenter to manipulate and analyze the communication patterns more easily. Finally, working within a simulated environment allowed for the acquisition of both procedural and declarative knowledge by the participants. The knowledge can be easily captured by a content analysis of the software (see

APPENDIX K). Furthermore, the content analysis can be used to construct a formal problem space (Newell & Simon, 1972).

2.1.2.2. Content analysis. APPENDIX K contains a list of 24 concepts that can be derived from interacting with the simulation. The list of concepts was developed in two ways. First, the information contained in the companion text (Ressler, 2002) included both conceptual and procedural knowledge for redesigning a bridge. Key propositions from the text helped form a framework for the content analysis. The text recommends that the student first optimize member properties then optimize the shape of the truss. Within optimizing individual member properties, several factors come into play, such as the strength, price, and different types of stress (i.e., tension & compression). Thus, each of these topics (i.e., strength, price, stress, and configuration) formed the topic headings in the framework.

The details of the framework were made explicit through detailed interactions with the software. The author used the software for approximately 40 hours. During that time, careful notes were kept that documented the knowledge that was needed to complete the task. Each concept is represented under the four topic headings. For example, to establish the relationship between member properties (i.e., steel type, bar type, or cross-sectional diameter) and strength, the experimenter conducted several small experiments in the simulation. For instance, the first concept (i.e., *members with larger cross-section are stronger than members with a smaller cross-section*) was tested in the simulation by increasing the diameter of members under tension, then compression, and observing the change in the strength-to-force ratio, which is found in the Member List (see button #13 in the screenshot of toolbar found in APPENDIX G). The relationship between member properties and strength was conducted for each property and type of stress. Generalizations from the results of the experiments were then summarized and recorded in the Content Analysis. The same procedure was repeated for each of the topic headings in the framework.

2.1.2.3. Assessment. The assessment items were derived directly from the content analysis of the simulation. Nearly all of the concepts from the content analysis were transformed into questions for the pre- and post-test. The mappings between the concepts based on the simulation and their corresponding assessment items can be found in the right columns of APPENDIX K. There were six different types of assessment items. Definition questions merely asked the participants to provide a definition for key vocabulary terms. The multiple-choice (MC) items assessed participants' ability to validly discriminate between target and distracter items. Rank-order (RO) items asked the participants to order a list of properties both by cost, then by strength. The greater-than-less-than (GTLT) items assessed the participant's ability to select members that were more expensive. Finally, the short answer (SA) questions asked deeper questions about the effect of bridge configuration on strength and price.

To illustrate the process of constructing the questions, consider concept #18: *The center of the bridge is under more stress than the ends*. This piece of declarative knowledge can be learned from the simulation by clicking on the Member List (see button #13 in the screenshot of toolbar found in APPENDIX G), or it can be inferred by contrasting the color intensity in the color-coded feedback for different members. The magnitude of the tensile and compressive forces is encoded in the simulation as a continuous change in color intensity in the load test mode (contrast the intensity of blue [tension] for Member 1 with the intensity for Member 2 in APPENDIX H). Thus, the student can see that the impressed load is the same for both members, but the middle member is under more stress than the member on the end. To evaluate the student's understanding of this concept, concept eighteen was changed into a question and evaluated in the multiple-choice portion of the pre-test and post-test (i.e., MC9). This process was repeated to derive all of the items in the right-most columns of APPENDIX K (i.e. the pretest and posttest columns).

2.1.2.4. Text. The text was written by abstracting relevant propositions from a companion text to the software (Ressler, 2002). The text that was given to the participants is reproduced in APPENDIX C. Certain relationships from the content analysis were made explicit in the text. For example, the first concept (i.e., *members with larger cross-section are stronger than members with a smaller cross-section*) was directly stated in the text. There is nearly a one-to-one correspondence between the information in the text and the simulation. These items were given directly to the participants because the knowledge was easy to assimilate, yet would take time to learn from the simulation (i.e., the small experiments the experimenter conducted).

The concepts that were stated in the text will be referred to as *text-explicit* knowledge. Information that is not directly stated in the text, but is still encoded in the Content Analysis will be referred to as *inferential* knowledge because this type of information needs to be inferred either from reading the text or interacting with the simulation.

2.1.2.5. Training instructions. The training instructions were written to be as concise as possible to reduce the load on working memory. Acronyms were used so that the participants could easily remember to produce an *I*dea, *E*laborate upon the idea, or *R*espond to an idea (IER) for the elaboration condition. The critical condition was instructed to produce an *I*dea, *C*hallenge the idea, or *R*espond to an idea (ICR). The training instructions for the critical interaction condition are reproduced in APPENDIX D, and the training instructions for the elaborative interaction condition are reproduced in APPENDIX E.

The instructions for the elaboration condition were based upon the definition of elaboration in the literature (Hogan et al., 1999). There was an added emphasis on the idea that incomplete ideas should be made explicit, which comes from the idea that elaborations supply variable assignments (see Section 1.2.3).

The instructions for the critical condition were based upon the various moves in a critical interaction (see Table 1). That is, participants were encouraged to ask for clarification, request more information, or to disagree with an idea. The top-level goal given to the participants was to evaluate the ideas presented during the collaboration.

2.1.2.6. Argumentativeness scale. The argumentative scale was chosen because it has been shown to differential individuals on the likelihood of engaging in critical discussions (Infante & Rancer, 1987). This particular scale was included as a control variable because some individuals were instructed to engage their partner with critical comments (see APPENDIX B for the argumentation scale).

2.1.3. Design

The experimental design contained two control conditions and two experimental conditions. The first control condition was a simple baseline condition in which individuals completed the task without a partner or any communicative manipulation. The second control condition included dyads solving the problem together, without any communicative manipulation. For the experimental conditions, the first included dyads who were instructed to engage in elaborative interactions, while the other included dyads who were instructed to engage in critical interactions. Except for communication training, all other aspects of the experiment were identical for each condition.

2.1.4. Procedure

Before the task began, participants were asked to provide informed consent. Once consent was granted, the experimenter ensured that neither individual had taken a civil engineering or material science course. The experimenter also verified that members of the dyads did not know each other prior to the experiment. If either of these two conditions were not met, then both individuals were excused from the study. After a brief introduction to the experiment, an on-line pretest was administered (see APPENDIX A for the pretest). The pretest was designed to measure incoming knowledge of bridge design and material science.

As stated in the materials section, the knowledge that was assessed on the pre- and post-test was split into two types, text-explicit and inferential knowledge. The text-explicit information included declarative knowledge stated directly in the materials. An example of text-explicit knowledge includes: *Members with a larger cross-section are stronger than members with a smaller cross-section (#1)*. The inferential knowledge was not directly stated in the text, but was inferred by interacting with the software. An example of inferential knowledge includes: *To strengthen a member under compression, add a joint somewhere along the beam (#19)*. It was assumed that inferential knowledge was more difficult to acquire than the text-explicit knowledge.

After completing the pretest, participants then completed the argumentation survey. Once the pretest and Argumentativeness Scale were completed, the participants read a short text individually (see APPENDIX C for the text).

For the experimental conditions, instructions for engaging in a specific type of dialog were given (see APPENDIX D for critical interaction condition and APPENDIX E for elaborative interaction condition). After reading the instructions, the experimenter answered any questions the participants had. To make the instructions concrete, the participants were given a warm-up task, in which the experimenter listened to participant interactions and intervened when necessary (see APPENDIX F for warm-up task). The procedure for teaching the communication scripts was adapted from Renkl, Stark, Gruber, and Mandl (1998) with one exception; the second warm-up problem was excluded.

After the participants completed the warm-up task, they were then introduced to the software in two ways. First, they watched a short QuickTime movie demonstrating the tools and features of the interface (see APPENDIX G and APPENDIX H for screenshots of the user interface). The experimenter verbally introduced the task and answered any remaining questions. The participants were given a pre-existing bridge design (created by the experimenter) and told their goal was to optimize the design by attempting to make it as cheap as possible, while still being able to carry a load (see APPENDIX I for the problem statement). The participants were given 30 minutes to complete the task. Their dialog was videotaped for later transcription.

The product generated by the collaborative pairs was a single bridge design, which was assessed along two dimensions. The first performance measure was the total fabrication cost of the bridge. This information was continuously displayed for the participants as an indication of their progress (see the top of the screen shot in APPENDIX G). A second measure, which was more sensitive than the first, was the bridge's overall load efficiency. This was determined by inspecting the internal forces on each individual beam in the bridge. This measure was selected because one heuristic for optimizing a bridge in this particular simulation is to maximize the internal forces for each member. Thus, higher values indicate that the bridge is close to optimization.

After working on the bridge optimization task for 30 minutes, the participants were asked to finalize their design. The final design was saved to the computer's hard drive. A posttest, which was identical to the pretest, was administered individually to measure how much information was learned from the text, the simulation, as well as from their interactions (see APPENDIX J for a copy of the posttest). Upon completion, the participants were debriefed, thanked, and excused from the experiment.

2.1.5. Measures

Several different measures were used to capture the problem-solving performance, as well as the learning that resulted from reading the text, solving problems, and receiving feedback from a simulation.

2.1.5.1. Performance measures. During problem solving, dyadic activity was measured along several dimensions. First, the number of times the bridge design was tested was counted (i.e., each time the "test" button was clicked). This measure will be referred to as *iterations*. Second, performance was rated according to two measures: the overall *savings* and an *optimization score*. Savings was calculated by subtracting the final price of the bridge from the initial starting value, which was the same for all participants: \$256,678.63. Thus, higher savings scores indicated better performance on the task.

$$savings = \langle price_{initial} \rangle - \langle price_{final} \rangle \tag{1}$$

To calculate the optimization score, the stress-to-strength ratios for each member were summed, which was then divided by the total number of members:

$$optimization = \frac{\sum_{i=1}^{n} \left(stress_i / strength_i \right)}{n} \tag{2}$$

where n = the number of members per bridge. The optimization score represents the average load for each member. The optimization score was used in conjunction with the savings variable because the optimization score indicates how close a given configuration was to being completely optimized (independent of price). Higher optimization scores indicated better performance on the task

2.1.5.2. Learning measures. Two learning measures were used to capture the knowledge acquired during the entire experiment. Recall that the participants read a short text after completing the Argumentativeness Scale. The information that was explicitly stated in the text will be referred to as *text-explicit* knowledge, which is a shallow type of learning. Information that was not explicitly stated, but must be inferred either from reading the text or interacting with the simulation will be referred to as *inferential* knowledge. Inferential knowledge represents deep learning because the participants have to construct the information actively (inferential items are demarcated in bold font in APPENDIX K). Gain scores for both text-explicit and inferential knowledge were calculated by the following equation:

$$gain = \langle S_{post} \rangle - \langle S_{pre} \rangle \qquad (3)$$

The gain scores were calculated by summing across all the items within each category (i.e., text-explicit and inferential) and then the pretest was subtracted from the posttest. The gain scores represent the amount of information learned from reading the text and interacting with the simulation.

2.2. RESULTS

The results for the present experiment were analyzed using planned comparisons because not every pair-wise comparison was of theoretical interest. The analyses were limited to test the following research questions:

1. Is it possible to train collaborative partners to interact in a specific way?
2. Does receiving interaction training lead to more effective (*a*) problem solving and/or (*b*) learning, relative to dyads that did not receive explicit interaction training?

Before testing these research questions, an analysis of individuals and dyads will be presented. The same value for alpha will be used for all analyses ($\alpha = 0.05$). The problem-solving measures were evaluated using the dyad as the unit of analysis (n = 68), whereas the learning outcomes were measured individually (n = 136).

The results are presented in three major sections. First, the psychometric properties of the assessment items and a manipulation check will be presented. Second, a discussion of nominal groups analysis will be presented. Third, the effect of interaction training will be presented.

2.2.1. Preliminary results: Assessment validation

Seven questions (four definitions and three short answer) were open-ended and nineteen questions were multiple-choice. Human graders scored the open-ended questions. Twenty percent of the data was selected for reliability analysis; that is, two independent coders scored 28 pretests and 28 posttests. The correlations between the two raters, for each of the seven questions, ranged from $r(54) = 0.65$, $p < 0.001$ to $r(54) = 1.00$, $p < 0.001$. Differences between the two raters were resolved through discussion. Once the discrepant items were agreed upon, the author graded the remainder of the assessments.

In terms of the reliability of the assessment itself, split-half reliability was calculated for all of the items separately for the pretest and posttest. The split-half reliability for the pretest was low overall ($\alpha = 0.31$), whereas the split-half reliability for the posttest was slightly higher ($\alpha = 0.52$). The reason the split-half reliabilities were low can be explained by the nature of the items. There was very little variance in the pretest items. That is, most individuals were unable to correctly answer many of the pretest questions. One might not expect a high split-half reliability for the posttest items because the questions represented a wide range of knowledge. Some of the items asked about the relationship between strength and materials, while other asked about the effect of bridge configuration on price.

2.2.2. Preliminary results: Verbal coding validation

A manipulation check was conducted to see if the communication instructions had their intended effects. A stratified sample of the control (n = 8), critical (n = 8), and elaborative (n = 8) conditions was taken such that both good and poor performers were equally represented. Several types of statements were coded.

For the elaborative statements, three types were coded (see Table 3 for coding scheme definitions and examples). The first type, *elaborate suggestion*, modified the speaker's initial suggestion by providing a location, an additional change, or a specific value for the proposed modification. The second type, *provide reason*, gave a justification for a particular change. The last type, *provide implication*, gave a consequence for a particular change. The coding scheme for the elaborative statements was based on the task analysis.

Table 3. *Manipulation check: elaboration coding scheme*

Code	*Definition*	*Example*
Elaborate Suggestion	Modify speaker's initial suggestion by…	
Location	…proposing a specific location.	A: *We could make them thinner*. B: *Change that one right there*.
Additional Change	…adding an additional change.	A: *We could make them thinner*… B: …*and hollow, too*.
Specific Value	…providing a specific value for the proposed modification.	A: *We could make them thinner*. B: *Make it 150 millimeters*.
Provide Reason	Giving a justification for the proposed modification.	A: *We could make them thinner*. B: *Because thinner is cheaper*.
Provide Implication	Giving a consequence for the proposed modification.	A: *We could make them hollow*. B: *Which will decrease the stress here*.

For the critical statements, four types were coded (see Table 4 for coding scheme definitions and examples). The first type, *counter-suggestion*, offered an alternative idea to the speaker's most recent suggestion. The second critical comment, *clarification questions*, the listener asked for the speaker to be more precise about a particular suggestion. The third critical comment, *request reason*, was more direct because it asked for a justification for a particular suggestion. The fourth critical contribution, *evaluation*, assessed the viability or worth of a suggestion. The first three critical comments were derived directly from the instructions given to the participants (see the "Challenge" box from APPENDIX D).

Table 4. *Manipulation check: critical coding scheme*

Code	*Definition*	*Example*
Counter-suggestion	Offer an alternative to the proposed modification.	*It might be easier to* ___. or *Why don't we try ___ instead?*
Clarification questions	Request for speaker to be more precise about a proposed modification.	*Which members?* or *What do you mean?*
Request reason	Request for a justification for a proposed modification.	*How come?* or *Why do you want to try ___?*
Evaluation	Assess the viability or worth of a proposed modification.	*I don't think that will work at all.*

Both the total number of elaborative and critical statements was counted for all three conditions. The mean number of elaborative and critical statements is presented in Table 5. To determine if the communication manipulation had its intended effects, the following statistical method was adopted. The average number of total elaborative and critical statements, taken from the control condition, was contrasted with the average number of total elaborative and critical statements separately for the elaboration condition and critical condition.

For the elaboration condition, there was not a statistically reliable difference between the elaborative and control condition for the number of elaborative statements, $t(14) = 1.33$, $p = 0.20$, $d = 0.66$; however, the effect size could be interpreted as a "large" effect. Furthermore, there was a significant decrease in the number of critical statements generated, $t(14) = -2.08$, $p = 0.05$, $d = 1.08$. While the elaboration condition was not instructed to *avoid* critical statements, it turned out that they supplanted critical dialog with elaborative statements.

For the critical condition, the instructions did not have their intended effects (the reasons why will be addressed in the Discussion section). The critical dyads produced the same number of critical statements as the control condition, $t(14) = -0.42$, $p = 0.67$. Furthermore, the critical dyads produced the same number of elaborative statements as the control condition, $t(14) = -0.10$, $p = 0.91$.

Table 5. *Manipulation check: average frequencies for elaborative and critical dialogs*

	Condition		
	Control Dyads (n = 8)	*Critical Dyads* (n = 8)	*Elaborative Dyads* (n = 8)
Elaborate Suggestion	8.12 (6.94)	8.12 (5.17)	12.12 (4.97)
Provide Reason	0.12 (0.35)	0.12 (0.35)	0.25 (0.71)
Provide Implication	0.87 (1.13)	0.50 (0.53)	1.00 (1.41)
Total Elaborative	**9.12 (7.72)**	**8.75 (5.47)**	**13.38 (5.76)**
Counter-suggestion	11.37 (6.78)	9.00 (2.78)	6.25 (5.01)
Clarification Questions	28.37 (10.90)	24.37 (12.12)	18.00 (9.71)
Request Reason	0.50 (0.76)	2.62 (2.13)	2.00 (1.69)
Evaluation	6.25 (3.24)	7.37 (3.85)	4.87 (3.27)
Total Critical	**46.50 (15.17)**	**43.37 (15.98)**	**31.12 (13.08)**

Note. Standard deviations are in parentheses.

The critical dyads performance on all of the performance and learning measures revealed non-reliable differences (all Fs < 1). Due to the lack of differences in interaction patterns, problem-solving performance, and learning between the critical and control conditions, they were collapsed into a single control condition (*n* = 39). Both the nominal and elaborative dyads were compared against this control condition.

2.2.3. Preliminary results: Nominal dyads analysis

To test the effects of interacting with a partner, the interacting dyads that did not receive communication training (i.e., the control dyads) were compared against the nominal dyads. Nominal dyad analyses control for the number of ideas produced by each individual; hence, the only difference between nominal and real dyads is the verbal interactions between partners. Hence, nominal dyads analysis is considered the "gold standard" for making comparisons between groups and individuals (Hill, 1982).

For the present study, nominal dyads were created by randomly pairing individuals, with the constraint that they were of the same gender. Gender was added as a constraint because the interacting dyads were of the same gender. Once the nominal pairs were created, the pair was assigned the highest value from the individual in the dyad for the savings and optimization scores. For example, suppose Peter and Brian were randomly paired together in the Individual condition. Peter's optimization score was 0.71 while Brian's was 0.73, thus Brian's optimization score was assigned to the pair. The same was true for the savings measure; however, the average number of iterations was assigned to the pair. The average, instead of the highest value, was selected because the number of tests does not *ipso facto* lead to better designs.

There are several methods for statistically aggregating individual performance (Lorge, Fox, Davitz, & Brenner, 1958), but the best performance was taken under the assumption that real, interacting dyads may only rely on the best member's ideas. To accommodate this possibility, the best individual member's performance was used to set a high criterion for finding an advantage for interacting with a collaborative partner.

To test the effects of interaction, the control dyads were compared against the nominal dyads. The results for the problem-solving and learning measures for nominal and control dyads are summarized in Table 6.

Table 6. *Average problem-solving and learning gains for nominal and control dyads*

	n	*Iterations*	*Savings*	*Optimization*	*Text-Explicit*	*Inferential*
Nominal Dyads	10	49.55	62.13	0.63	21.43	10.56
		(15.13)	(14.15)	(0.09)	(15.72)	(15.07)
Control Dyads	39	36.79	46.29	0.55	27.17	7.69
		(13.35)	(22.68)	(0.11)	(14.62)	(17.39)

Note. Standard deviations are in parentheses.

There were marginal differences between the nominal and control dyads for savings ($F(1, 65) = 3.81$, $p = 0.06$) and the optimization score ($F(1, 65) = 3.11$, $p = 0.08$), both favoring the nominal dyads. There was strong evidence that nominal pairs were much more rapid in testing their designs. There was a main effect of condition on iterations ($F(1, 65) = 5.58$, $p = 0.02$) reflecting a greater number of tests for the nominal dyads ($M = 49.55$, $SD = 15.13$) than the control dyads ($M = 36.79$, $SD = 13.35$). Working with a partner thus reduced the number of tests that were conducted. In turn, the number of tests conducted was strongly correlated with both performance measures (savings, $r(66) = 0.72$; optimization score, $r(66) = 0.65$).

Given that iterations differed for nominal and control dyads, the number of iterations was included as a covariate for both of the problem-solving measures. Iterations were not, however, used as a covariate for either of the learning analyses because there was no *a priori* reason for suspecting that iterations would lead to better learning. This was supported by the weak correlations between iterations and text-explicit ($r(66) = -0.044$) and inferential ($r(66) = -0.035$) learning for the full dataset.

When iterations were used as a covariate, the marginal differences all became non-significant. Controlling for the number of iterations, there were no differences in savings or optimization scores between nominal and real dyads, Fs(1, 64) < 1.00 (see Figure 2a & 3b). As reported above, there were no differences in learning text-explicit or inferential knowledge, Fs(1, 133) < 2.45, ps > 0.12 (see Figure 2c & 3d). This suggests that the interacting control dyads were at least as good as the best nominal individuals, after controlling for the speed at which the individuals were able to test their designs.

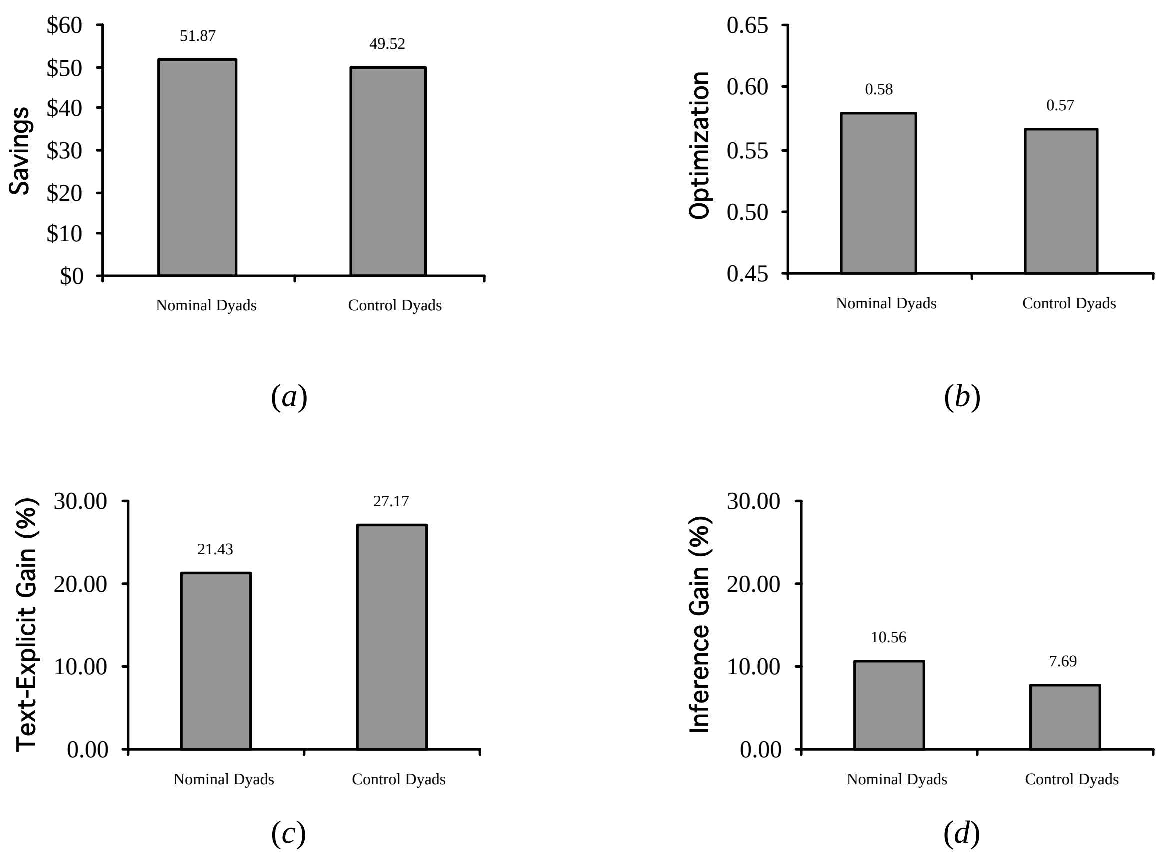


Figure 2. Adjusted means for (a) savings and (b) optimization scores, and average (c) text-explicit and (d) inference gains for the nominal and real dyads.

Collaborative work incurs the cost of coordinating and checking with a partner. The upshot of this is that the group may ultimately produce a better product. Furthermore, the penalty of time or activity may decrease over repeated collaborative events. In other words, the next time the same group gets together, the cost of interacting may be reduced. To test this conjecture, the number of iterations that the control condition produced during the first half of the experiment was contrasted with the number of iterations from the last half of the experiment. There were significantly more tests for the second half (M = 22.57, SD = 9.39) than the first half (M = 14.88, SD = 7.07), $t(64) = 3.761$, $p < 0.0004$. This suggests that the control dyads became better able to coordinate with each other over time.

2.2.4. The effect of elaborative interactions

To test the effectiveness of receiving instructions to elaborate, the elaborative condition was compared to the control dyads. In terms of problem solving, there was no effect of condition on the number of iterations, suggesting the elaborative (M = 41.00, SD = 18.62) and control dyads (M = 36.79, SD = 13.35) tested their designs an equal number of times, $F(1, 65) < 1$. There was no effect of condition on savings, suggesting both conditions constructed similarly priced bridges, $F(1, 65) = 2.08$, $p = 0.15$. On the other hand, there was a main effect of condition on optimization score ($F(1, 65) = 5.36$, $p = 0.02$) reflecting a higher score for the elaborative dyads (M = 0.63, SD = 0.15) than the control dyads (M = 0.55, SD = 0.11). Elaborating a partner's ideas and suggestions increased the dyads' ability to optimize their particular designs (see Figure 3b).

In terms of learning, there was no effect of condition on text-explicit knowledge, suggesting both conditions learned roughly the same amount of shallow knowledge, either from the text or through interacting with the simulation, $F(1, 133) < 1.00$. On the other hand, there was a main effect of condition on inferential learning gains ($F(1, 133) = 8.41$, $p < 0.01$) reflecting a higher score for the elaborative dyads ($M = 17.84$, $SD = 19.42$) than the control dyads ($M = 7.69$, $SD = 17.39$). Elaborative dialog seemed to help the dyads construct better and deeper representations of the task domain (see Figure 3d).

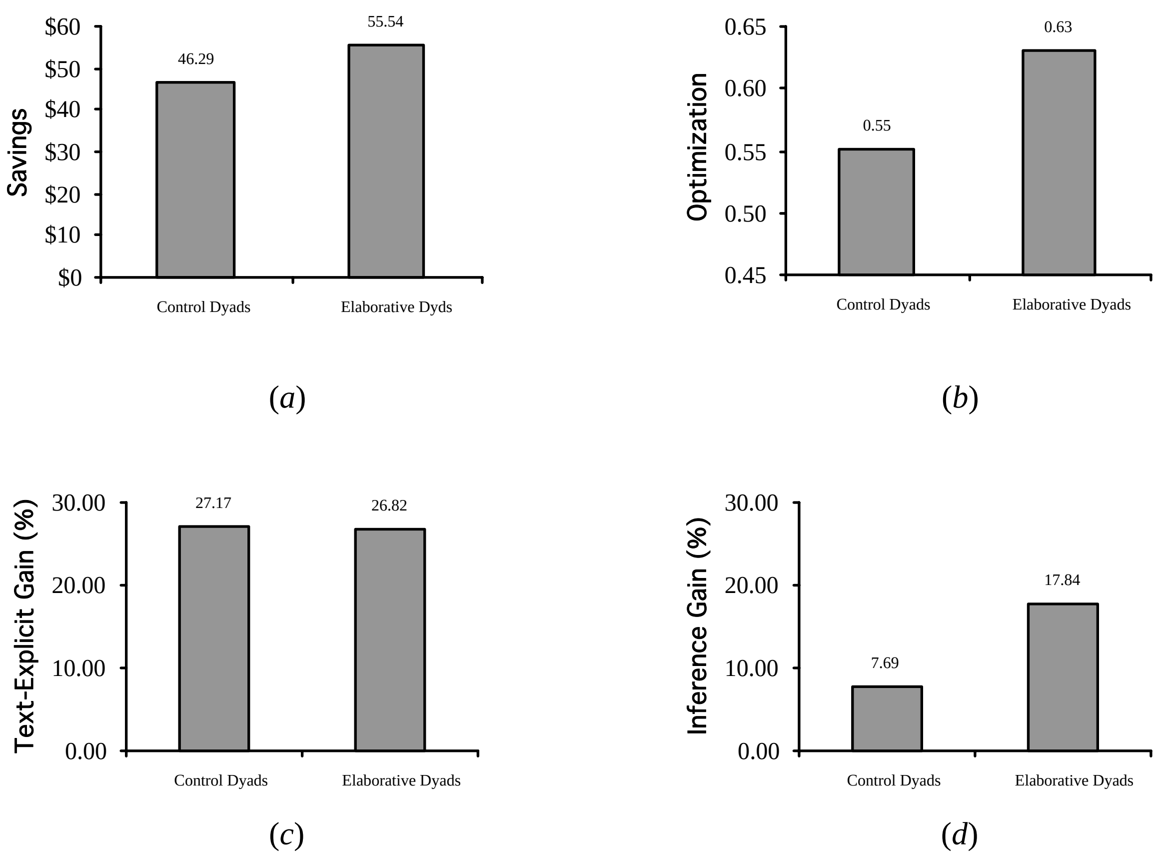


Figure 3. Average (a) savings, (b) optimization scores, (c) text-explicit, and (d) inference gains for the control and elaborative dyads.

Why did the elaboration condition perform better and learn more than the control condition? There are three interrelated hypothesized explanations. The first is related to the way in which the elaborative partners interacted. This is an obvious choice because the only dimension in which the two conditions differed was their communication patterns. Elaborating upon another person's idea may increase the specificity of each suggested modification. As it has already been argued, dyads devoted several turns to negotiate which action to take. In the present task, there are at least three variable assignments that need to be made. First, a change needs to be specified (e.g., change the type of steel), then a specific value needs to be specified (e.g., from carbon steel to high strength low alloy steel), and finally the location needs to be identified (e.g., member #12). Each of these variables can take multiple turns to establish. An elaborative sequence may come to establish the variable assignments more quickly than one in which the partner asks the other to specify the variable assignments.

To test the efficiency in communication patterns, the number of clarification questions was coded in the elaborative and control conditions. Clarification questions were selected because the goal of the question is to establish the variable assignments, which is illustrated in the following exchange (see Table 7). Beth makes the suggestion that they make some members hollow (turn 82). Abby asks which solid members should be changed to hollow (turn 83), to which Beth replies that the top member should be changed (84).

Table 7. *Clarification question example*

Turn	**Speaker**	**Contribution**	
82	Beth:	So these are ah solid tube. Alright, so we can make some of these hollow.	
83	Abby:	The top ones or the…	clar_q
84	Beth:	The top.	
85	Abby:	Like every other one?	clar_q
86	Beth:	Try it…	

In contrast, consider an elaborative exchange (see Table 8). In this brief exchange, Mike proposes that they make the diagonal members a smaller diameter (turn 82). Dan accepts Mike's proposal, and elaborates it by suggesting a location (i.e., the middle, turn 83). They are able to make a suggestion and implement it relatively quickly because they can avoid the need to ask for clarification. In addition, they are building off one-another's ideas, instead of fleshing out a single partner's idea.

Table 8. *Elaboration example*

Turn	**Speaker**	**Contribution**	
82	Mike:	Cause usually, I don't know, do you want to try making the cross members smaller?	
83	Dan:	Um, we could,- just the ones in the middle not the ones on the end.	elab
84	Mike:	Yeah, right.	

The elaboration condition (M = 18.00, SD = 9.71) generated marginally fewer clarification questions than the control condition (M = 26.37, SD = 11.32), $F(1, 22) = 3.19$, p = 0.09, d = 0.79. Because the elaboration condition asked marginally fewer clarification questions, this suggests that the communication instructions had a direct effect on their dialog, which, in turn had an indirect effect on their problem-solving performance.

The second hypothesized explanation is in the way in which the dyads used the feedback from the simulation. The most useful feedback for the present task was found in two different sources. The first is the color-coded feedback, which was superimposed over the individual members, when the user tests a bridge. Recall that the magnitude of the tensile and compressive forces is encoded in the simulation as a continuous change in color intensity in the drawing (for an example, refer back to the screenshot in Appendix H. The members in the center of the bottom cord are darker blue than the members at the ends, indicating higher levels of tension.). The second source is the Member List in which the information is presented as a ratio of the member's strength to impressed force. An example of the use of the feedback in dialog can be found in the following exchange (see Table 9).

Table 9. *Simulation feedback example*

Turn	Speaker	Contribution
168	Ben:	Let's see if we can break it. Then we can change them all, some of these.
169	Nathan:	Hm-mmm
170	Ben:	Find one that's getting really red.
171	Nathan:	Still not that red, right?
172	Nathan:	This vertical one too, and this vertical one. They're still pinkish.
173	Nathan:	We could make them thinner.
174	Ben:	Alright. That's...second

In this example, Ben sets the goal to find a member that is not experiencing much compression (i.e., "red", turn 170). They use this information to select a member to change. Once the change was made, they observed the effect by looking at the intensity change (turn 171). This pair used this information to suggest a location (i.e., a particular member), as well as to make specific changes (i.e., the cross-sectional diameter, turn 173).

To test the hypothesis that the elaborative condition was more effective in exploiting the feedback provided by the simulation, the explicit mention and use of the feedback was coded for the elaborative and control conditions. The elaborative condition made explicit mention and use of the simulation feedback more times than the control condition (see Table 10), $F(1, 22) = 4.69$, $p = 0.04$, $d = 0.80$. Because the feedback from the simulation was an effective cue for redesigning a cheaper and more optimized bridge design, and considering that the elaborative condition explicitly mentioned the feedback at a rate of three times the control condition, this may help explain why the elaboration condition performed and learned better than the control condition.

Table 10. *Average frequency of explicit mentioning of simulation feedback*

	Condition	
	Control Dyads (n = 16)	*Elaborative Dyads* (n = 8)
Color-coded FB	3.50 (4.10)	8.75 (11.12)
Member-list FB	0.00 (0.00)	1.12 (2.47)
Total Feedback	**3.50 (4.10)**	**9.87 (10.45)**

Note. Standard deviations are in parentheses.

The third explanation is derived from the types of strategies each condition employed. Dyads were categorized as employing either an effective or ineffective strategy. Strategy effectiveness was derived from a thorough task analysis. The companion text for the software made a few recommendations for optimizing a truss, which was then supplemented with protracted interactions with the simulation. The following strategies were coded (see Table 11). Positive strategies included: (1) identifying individual members with either a high or low amount of stress, which is helpful because it can guide the dyad toward (or away) from modifying certain members; (2) only manipulating one feature or property during a single iteration; (3) finding the point at which a member fails to help find each member's optimized load, and (4) making a single change and observing the resulting change in price (which also helps with the task of making a bridge as cheap as possible). Negative strategies included: (5) removing members or joints; (6) adding superfluous members or joints to the existing structure (this is only effective in very few instances); and (7) using longer members (because strength decreases as members increase in length, especially for members under compression).

Table 11. *Average frequency of explicit mention of positive and negative strategies*

	Condition	
	Control Dyads (n = 16)	*Elaborative Dyads* (n = 8)
1. Identify members with high/low stress	0.12 (0.50)	0.25 (0.46)
2. Vary one thing at a time	0.31 (0.60)	0.25 (0.46)
3. Find failure point for individual member	0.37 (0.72)	0.75 (1.75)
4. Set up & conduct experiments: change property & observe price change	0.44 (0.73)	1.00 (2.07)
Positive strategies	**1.25 (1.69)**	**2.25 (4.37)**
5. Removing members & joints	5.87 (5.12)	2.62 (3.50)
6. Adding several members & joints	1.75 (1.84)	1.87 (1.64)
7. Using longer members	0.69 (1.35)	0.75 (1.75)
Negative strategies	**8.31 (5.61)**	**5.25 (5.73)**

Note. Standard deviations are in parentheses.

Dyads were categorized based on their explicit mention of the strategies. Each dyad was rank ordered for their use of positive strategies. Each dyad in the upper half of the rank was assigned a gestalt grade for their search strategy effectiveness. That is, those who frequently used positive strategies were categorized as "high;" whereas, dyads in the bottom half of the ranked list were categorized as "low." The frequencies of the dyads classified as using effective strategies can be found in the upper half of Table 12. The same method was used for negative strategies. The frequencies of the dyads classified as using ineffective strategies can be found in the lower half of Table 12.

The control dyads were more likely to be classified using a negative strategy than the elaborative or control dyads, $\chi^2(1) = 4.11$, $p = 0.04$. Additionally, the elaborative dyads were more likely to be classified using a positive strategy, $\chi^2(1) = 3.00$, $p = 0.08$. Therefore, it seems that, at the minimum, the elaborative dyads were most likely to avoid the use of negative strategies and potentially engage in positive strategies.

Table 12. *Frequencies of dyad classification for positive and negative search strategies*

	Condition	
	Control Dyads (n = 16)	*Elaborative Dyads* (n = 8)
Positive strategies		
High	6	6
Low	10	2
Negative strategies		
High	11	2
Low	5	6

2.3. DISCUSSION

The field of cognitive psychology is slowly evolving to include under its purview interactions between two cognitive agents. Ericsson and Simon (1980) laid the groundwork for this evolution by justifying verbal interactions between people from a theoretical position. However, in that same paper, the authors cautioned scientists to be weary of verbal data that included verbalizations that provided explanations and justifications for the individual's actions. It is precisely these types of verbalization that are of educational interest, as it has been shown that providing explanations can alter an individual's understanding of a domain (Chi et al., 1989a; Coleman, 1998; Webb, 1989).

When an individual develops an explanation for herself, this type of verbalization is referred to as *self-explaining* (Chi et al., 1989a). The target of the explanation is typically oneself. However, in a collaborative problem-solving context, the speaker must consider her audience (i.e., her partner). When students attempt to co-generate an explanation together, this process might be referred to as a collaborative interaction.

The learning effects due to collaboration seem to come from two different modes of interaction. On the one hand, each individual operates on their partner's reasoning by elaborating one-another's ideas. On the other hand, the partners can also operate on each other's reasoning by challenging the validity of their partner's suggestion or idea. These two dialog patterns were labeled *elaborative* and *critical* interactions in the present study. The background literature on elaboration and argumentation suggest that each dialog has its own strengths and weaknesses. The question is, which will lead to stronger learning gains relative to a control condition?

In an effort to answer this question, the current study investigated two interrelated research questions. Before testing the first research question, problem solving and learning was contrasted between real, interacting dyads and nominal, non-interacting dyads. The results suggest, after controlling for activity (i.e., iterations), interacting dyads were as good as the best individual in a nominal pair. This finding is congruent with literature showing that groups are equivalent to the best individual.

The first research question asked if specific types of interaction can be trained. That is, can individuals be taught to be elaborative or critical? Evidence for the first research question was mixed. On the one hand, the critical dyads were not able to assimilate the experimental instructions such that their interacts were altered. There was no evidence that they were able to produce critical statements at a rate that was different from the control condition. On the other hand, it was found that dyads who were asked to elaborate were able to assimilate the instructions and interact with each other in an elaborative way. Moreover, the elaborative dyads were more successful in both problem solving and learning compared to dyads who were not given special communication instructions. This finding suggests that it is easier to train college undergraduates to elaborate than evaluate ideas.

The second research question asked if receiving interaction training has a measurable effect relative to a control condition that did not receive communication instructions. Again, the evidence was mixed. Because the critical condition was not able to implement the instructions, their performance was similar to that of the control on all measures of problem solving and learning. The elaboration condition, on the other hand, was able to outperform the control condition on the more sensitive dependent measure (i.e., the optimization score) and the deeper measure of learning.

It is encouraging that a short communication intervention can have a measurable impact on at least one type of interaction style (i.e., elaboration). Further research needs to be conducted to isolate exactly how elaboration has an impact on problem solving and learning. For instance, are there different types of elaboration that are more or less effective for problem solving (i.e., precise versus imprecise elaborations, or elaborative completions versus complete statements)? Furthermore, are there differences in subsequent learning from different types of elaboration?

2.3.1. The costs associated with working in a pair

The results from the preliminary analyses that showed nominal individuals performing better than the control conditions warrants further inquiry. Two findings are relevant. First, nominal dyads tested more bridges than the interacting dyads. Why were the nominal dyads able to work at a faster pace? One possible reason is the need to coordinate ideas and actions when working with a partner. One particularly clear example of the cost associated with coordination can be found in the following example (see Table 13). The excerpt was taken from the critical interaction condition. The first few turns are dedicated to establishing *which* idea to implement (turns 1-6). Then they decide *where* to implement the idea (turns 7-9). Once the idea and location are decided, they must then ground their actions within the simulation (H. H. Clark & Brennen, 1991; H. H. Clark & Schaefer, 1989) (turns 10-17). All of these processes take time, which is the main reason why individuals were able to test more bridges than dyads.

Table 13. *Example of idea and action coordination in real dyads*

Turn	**Speaker**	**Contribution**	**Code**
1	Anne:	Okay. So maybe we should make like all this stronger?	Idea
2	Beth:	Hm-mmm.	Idea
3	Anne:	I don't know.	Idea
4	Beth:	So should we put that back to what it was? And make that, or…	Idea
5	Anne:	Maybe. Do you want to try that?	Idea
6	Beth:	Okay.	Idea
7	Anne:	Should we do all of them? Or…should we do…I don't know. I have no idea if it's gonna work.	Idea
8	Beth:	Try doing just this triangle in here.	Idea
9	Anne:	This? And this? Or just those three?	Idea
10	Beth:	Maybe just this	Action
11	Anne:	Like that?	Action
12	Beth:	Yeah.	Action
13	Anne:	To…this one?	Action
14	Beth:	Hm-mmm.	Action
15	Anne:	Same thing over here?	Action
16	Beth:	Hm-mmm.	Action
17	Anne:	‘kay.	Action

The observed slowdown for this collaborative task has implications for other studies conducted on collaborative problem solving. Generally, experiments that investigate the effects of collaboration on problem-solving performance control for time-on-task (Stroebe & Diehl, 1994) or test it as a possible explanation of the collaborative results (Okada & Simon, 1997). However, the results from the present experiment suggest that it is not time that is the crucial variable, but something closer to the amount of activity or task engagement. Instead of controlling for time, experimenters may consider controlling for activity (however conceived).

The second relevant finding was that there were at least marginal differences between the nominal and real dyads’ performance on both measures of problem solving (before iterations was factored out statistically). Why might nominal dyads performed better than interacting dyads? There are potentially two interrelated explanations for the lack of group learning differences. Before reviewing the explanations it should be noted that, while each dyad had learning gains that were significantly greater than zero, there were several items of information that were yet unlearned (range: 64% – 80%). Thus, the lack of differences cannot be explained by a simple ceiling effect.

The first explanation is based on a cognitive load hypothesis (Sweller, 1994). Interacting dyads are faced with several simultaneous tasks: (1) They are learning how to interact with a partner in a novel situation, (2) attempting to solve a complex problem, (3) engaging in a dialog, which requires the speaker to produce relevant contributions (and, in the critical condition the contributions were supposed to conform to a particular style), and (4) trying to learn the relationships between bridge price and member strength by manipulating the properties of the individual members and their configuration. Because each task in and of itself may require some cognitive resources, all of them combined may be a bit overwhelming and does not leave many resources available for deep processing of the material. The cognitive load hypothesis suggests a simple modification to the procedure. The dyads could be prompted to take periodic breaks and reflect on what they have learned so far (Katz, Allbritton, & Connelly, 2003). The cognitive load hypothesis would predict that including reflection breaks might produce stronger learning gains. Imposing short reflection breaks could be easily implemented in a replication of the current study.

The second possible explanation, which is an implication of the first, suggests that real dyads might show stronger learning gains if they are given a delayed posttest. Any additional learning, above and beyond the information encoded during the one and a half-hour experiment, might occur during post-mortem reflection of experimental content. That is, the individual might learn the task-relevant information after the experiment is over, when he or she has the cognitive resources to reflect on what was said and done during collaborative problem solving (Azmitia, 1996).

2.3.2. Effective interaction training

Another issue raised by the results is the lack of expected performance by the critical interaction condition. The literature on argumentation suggests training groups to evaluate suggestions and ideas should lead to superior problem solving and learning. According to the results, it appears that providing college undergraduates with instructions to elaborate upon one another’s ideas has a positive impact on both their problem-solving performance and deep learning. Training an undergraduate population to challenge one another, however, did not have the same effect as the elaboration instructions. Instead, the pattern of data suggests that the critical dyads were not able to employ the communication strategy. There was very little in the dialogs that suggested the dyads challenged each other in a deep way. Twenty-three percent of their dialog was coded as deep, whereas 77% was coded as shallow. The critical condition produced mostly clarification questions and counter-suggestions, both of which are shallow styles of critical interaction.

Because the critical dyads were unable to implement the interaction instructions, one may be left wondering why. To explain why, it is useful to differentiate between *internal* and *external* collaboration scripts (O'Donnell & Dansereau, 1992). An internal collaboration script is one’s own personal style of communication, which is adopted over years of interacting with people in a given society or culture. Thus, internal collaboration scripts are culturally bound and idiosyncratic, given a person’s past experiences. An external collaboration script is one that is provided to the participants during an experiment. An external script is a way of interacting that is explicitly taught to an individual.

One hypothesis why the critical condition did not engage in evaluative behaviors is because the undergraduates’ internal collaboration scripts were in conflict with the external collaboration scripts provided to them (Dillenbourg, 2002). We can assume the interaction patterns displayed by the control condition serve as representative of the internal collaboration script that most undergraduates bring with them to any given situation. Being asked to be critical may impose a style of interaction that is “unnatural.” In so doing, the experiment may have imposed an additional load on the dyad. Evidence from two studies supports the conflicting-scripts hypothesis.

First, Moreland (2005, May) found an analogous result in a study on transactive memory. In one condition, groups were explicitly told to develop a transactive memory system, while the control group was not instructed to do so. He found the control group outperformed the instruction group. To explain the surprising results, Moreland postulated that the instructions to form a transactive memory system disrupted the default group processes.

The second piece of evidence for the conflicting-scripts hypothesis, which is more germane to the issue of argumentation, is a study by Kollar, Fischer, and Slotta (2005). They found a main effect for high-structured internal scripts for two of their three dependent measures. This suggests that individuals who already have a well-established, internal script for argumentation tend to learn more than the students who have a weakly established internal script for argumentation. Their interpretation of the main effect, which was effectively a conflicting-scripts hypothesis, conjectured that learning a well-structured, external collaboration script concurrent with the domain-general and domain-specific knowledge reduced the likelihood of making deep elaborations of the learning materials.

The two hypotheses outlined above suggest a few possible directions for future research. To test the conflicting-scripts hypothesis, the present experiment could be replicated in a cultural population that has a well-established, internal collaboration script that encourages disagreements and challenges (e.g. Setlock, Fussell, & Neuwirth, 2004).

2.3.3. How elaboration led to increased problem solving and learning

As stated in the introduction, elaborative activities enhance learning by increasing an individual's ability to recall information, comprehend a text, and learn conceptual material. A similar finding was observed in the present experiment. The elaborative condition answered more questions correctly on the inference items than the control condition. How did instructions to elaborate lead to better problem solving performance and therefore learning?

An analysis of the interactions may offer a potential explanation. Specifically, as was argued in the introduction, elaboration may serve to more quickly assign variables to unfilled slots (e.g., when represented propositionally). The results suggested that elaboration may have been effective in filling unassigned variables because the elaborative dyads asked fewer clarification questions than the control condition. An example from the protocol showed one person making a recommendation for a change, and the other person suggesting specifically where to implement the change. Additionally, the results indicate that the elaborative dyads were better able to use the color-coded feedback from the simulation. The color-coded feedback is helpful in deciding where to implement the changes.

Combining these two results, making faster variable assignments and better use of the simulation's feedback, we might interpret an exchange between two dyads in the elaboration condition in the following way. One person suggests that they modify the member properties by changing solid members to hollow beams. The second person may elaborate the suggestion by looking at the color-coded feedback and making a recommendation based on the color-coded feedback. If the second person makes explicit how she made her recommendation, then the use of the color-coded feedback is now available to the dyad for future use. The finding that the elaborative dyads were more likely to be classified as using positive strategies and less likely to be categorized as using negative strategies, supports this interpretation.

How, then, does effective interactions and use of simulation feedback translate into better learning. One interpretation is that the better performance in the simulation allowed the students to answer the questions correctly on the posttest. The questions were designed to mirror some of the issues the dyads face while solving the design problem. For example, the second short-answer question asked, *Why does adding supports, which themselves experience no internal member forces, increase the overall strength of the bridge?* The question was targeted at the idea that members under compression can be strengthened by bisecting their length with a joint and connecting the joint to another member. This knowledge was helpful in optimizing the initial design.

Provided the dyads are exposed to successful problem solving, the elaboration literature suggests that the dyads may be more likely to remember and transfer that information to the posttest. What is unclear from the current project is if the information that was correctly used on the posttest was jointly constructed, and then reproduced by both individuals, or if the elaborations only helped increase the memory strength for the individual who uttered the elaboration.

2.4. EDUCATIONAL IMPLICATIONS

There is increasing awareness in our nation's schools that collaborative learning is an important component to joining the modern workforce (S. G. Cohen & Baily, 1997). One challenge classrooms face is how to implement collaborative learning into the curriculum effectively. Collaborative work is not an easy task because the teacher surrenders a modest amount of control when allowing the class to work on their collaborative problem-solving assignments. What does the present research recommend for smoothing the transformation of a teacher-driven lecture to a more collaborative classroom environment?

First, it seems clear that groups benefit from a certain amount of structure in their conversations. If left alone, students must discover their own interaction styles. While discovery works for some groups, it may not benefit all groups equally. Encouraging students to build off one another's ideas is a simple intervention that most college students are willing to adopt. Elaborative dialog is an effective method of interaction in terms of both problem solving and deep learning.

However, not all ideas are of equal quality and some level of evaluation will ultimately become necessary. If the instructor values evaluation, she must make a few assumptions. First, if evaluation is going to arise naturally, it will probably be a rare phenomenon. Second, students may be somewhat hesitant to interact critically with one another (Browne & Hausmann, 1998; Browne, Hausmann, & Ostrowski, 2002), even when evaluation may have been a part of a student's college education (Keeley, Browne, & Kreutzner, 1982). Thus, argumentation training may be best suited if it is not framed using intimidating terms, such as: *critical*, *critique*, *argue*, or *evaluate*. Instead, the teacher should motivate her class to ask specific types of questions (i.e., clarification, request reason, etc.) or to challenge one another. The metaphors teachers use to instruct critical dialog are extremely important.

Conflict-driven dialogs might also be best taught using certain computer-mediated scaffolds (Fischer, Bruhn, Graesel, & Mandl, 2002). Presenting arguments over the internet, or with a computer tutor (Ashley, Desai, & Levine, 2002) might also help students become comfortable with this style of interaction. Instructing undergraduate populations to argue will require more than a single, brief instructional intervention. Furthermore, critical instruction may be more effective when friends, rather than strangers, are asked to critique one-another (Azmitia & Montgomery, 1993).

2.5. CONCLUDING THOUGHTS

Two themes emerge from the study presented here. The first theme is the extension of cognitive theory to include an interacting set of individuals. Similar arguments have been made with respect to including artifacts found in the environment (Norman, 1993; Zhang & Norman, 1994), as well as the system-wide interaction of artifacts and multiple agents (Hutchins, 1995). What remains unclear is how the proposed mechanisms for knowledge generation within the individual (i.e., inferential mechanisms) can be mapped on to a collection of interacting individuals. It would be inappropriate to remove individual cognition from the equation, but it would be equally inappropriate to ascribe individual cognitive mechanisms to the collective (Salomon, 1993). Thus, there are probably emergent cognitive mechanisms operating within a group.

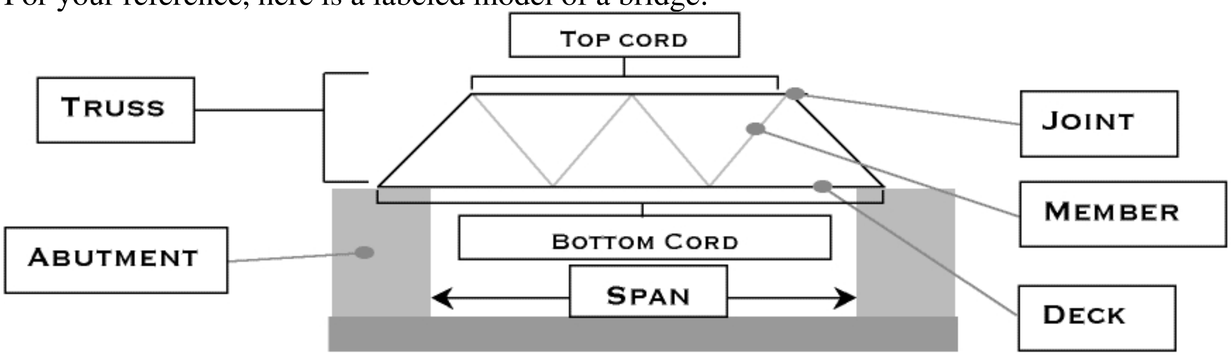
The second theme is related to the first. Building dialog into our theories of cognitive psychology represents a huge challenge. At this point, it is unclear how utterances are produced, or even understood (see Bock & Levelt, 1994; Levelt, 1999 for notable exceptions). Making the connections between cognitive theory and dialog production and comprehension, with connections to learning, represents an exciting frontier for cognitive science.

APPENDIX A

Background Knowledge Questionnaire: Pretest

Participant ID Number:

For your reference, here is a labeled model of a bridge:



Provide a definition for the following terms:

- Load:
- Tension:
- Compression:
- Failure:

Rank order each type of steel, from weakest (1) to strongest (3):

Strength Rank Order	Steel Type
pull down	Carbon Steel
pull down	High-strength Low-alloy Steel
pull down	Quenched and Tempered Steel

Rank order each type of steel, from cheapest (1) to most expensive (3):

Cost Rank Order	Steel Type
pull down	Carbon Steel
pull down	High-strength Low-alloy Steel
pull down	Quenched and Tempered Steel

Indicate which material is more expensive (>) or less expensive (<).

1. Hollow Tube	pull down	Solid Bar
2. Longer Member	pull down	Shorter Member
3. Smaller Cross-section	pull down	Larger Cross-section

Multiple Choice (*Choose the best answer for each question*)

1. As cross-sectional dimension increases, member strength ____________.
   (a) increases
   (b) stays the same
   (c) decreases

2. Hollow tubes have ____________ tensile strength than solid bars.

   (a) higher
   (b) same
   (c) lower

3. Tensile strength is always ____________ the maximum compressive strength.

   (a) greater than
   (b) the same as
   (c) weaker than

4. Under compression, longer members are ____________ shorter members.

   (a) stronger than
   (b) the same as
   (c) weaker than

5. Suppose we took a bridge and made it taller (in the vertical plane), without changing the overall configuration. The internal member forces of the top and bottom cords will ____________.

(a) increases
(b) stays the same
(c) decreases

6. The top of a bridge is always under ____________, while the bottom is always under ____________.

   (a) compression; compression
   (b) compression; tension
   (c) tension; compression
   (d) tension; tension

7. Strength-to-force ratio: a value ____________ one means the member has failed, while a value ____________ one means the member can safely carry the load.

   (a) greater than; greater than
   (b) greater than; less than
   (c) less than; greater than
   (d) less than; less than

8. Under tension, longer members are ____________ shorter members.

   (a) stronger than
   (b) the same strength as
   (c) weaker than

9. Where does the bridge experience the most stress?

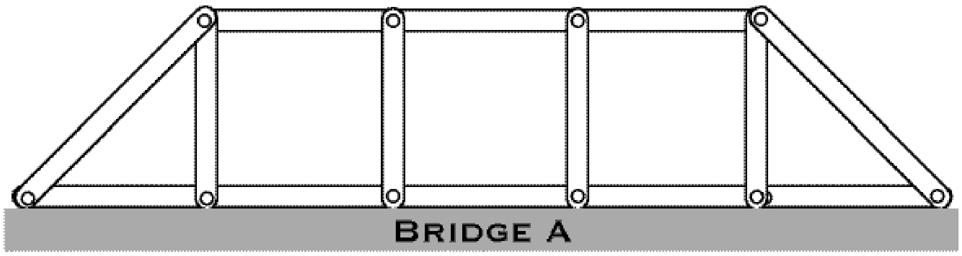
   (a) The top cord
   (b) The middle
   (c) The bottom cord
   (d) The two ends

10. Using members of several different sizes ____________ the overall cost of the bridge.

   (a) increases
   (b) does not change
   (c) decreases

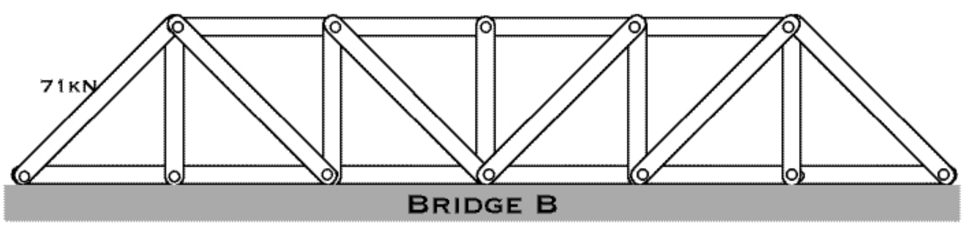
Short Answer

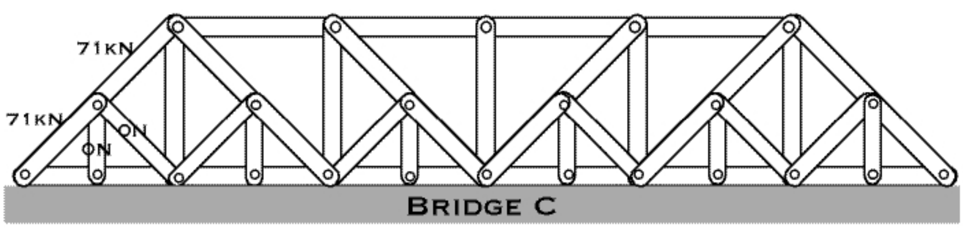
1. The bridge below (Bridge A) is flawed in a significant way. What might be done to the bridge to allow it to carry a load? (Describe your modifications to the bridge in the textbox below.)



<type answer here>

2. Consider two bridges shown below (Bridge B and Bridge C). Their configurations are similar, with one exception. Bridge C has 12 additional members. Although none of the internal member forces in the diagonals (71kN) change, the bridge is stronger. Why does adding supports, which themselves experience no internal member forces, increase the overall strength of the bridge?





<type answer here>

3. Since we have considered the members under compression, describe what you might do to the members under tension (Bridge C), assuming they are made of solid carbon steel bars (160mm)?

<type answer here>

Submit Reset

APPENDIX B

Argumentation Questionnaire

Participant ID Number:

This questionnaire contains statements about arguing controversial issues. Indicate how often each statement is true for you personally by placing the appropriate number in the blank to the left of the statement.

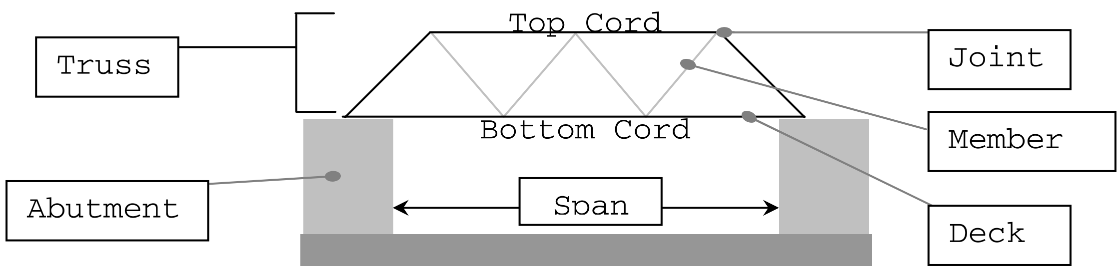
1	2	3	4	5
Almost never true	rarely true	occasionally true	often true	always true

pull down	1. While in an argument, I worry that the person I am arguing with will form a negative impression of me.
pull down	2. Arguing over controversial issues improves my intelligence.
pull down	3. I enjoy avoiding arguments.
pull down	4. I am energetic and enthusiastic when I argue.
pull down	5. Once I finish an argument I promise myself that I will not get into another.
pull down	6. Arguing with a person creates more problems for me than it solves.
pull down	7. I have a pleasant, good feeling when I win a point in an argument.
pull down	8. When I finish arguing with someone I feel nervous and upset.
pull down	9. I enjoy a good argument over a controversial issue.
pull down	10. I get an unpleasant feeling when I realize I am about to get into an argument.
pull down	11. I enjoy defending my point of view on an issue.
pull down	12. I am happy when I keep an argument from happening.
pull down	13. I do not like to miss the opportunity to argue a controversial issue.
pull down	14. I prefer being with people who rarely disagree with me.
pull down	15. I consider an argument an exciting intellectual challenge.
pull down	16. I find myself unable to think of effective points during an argument.
pull down	17. I feel refreshed and satisfied after an argument on a controversial issue.
pull down	18. I have the ability to do well in argument.
pull down	19. I try to avoid getting into arguments.
pull down	20. I feel excitement when I expect that a conversation I am in is leading to an argument.

APPENDIX C

Text Materials

Anatomy and Vocabulary of a Bridge

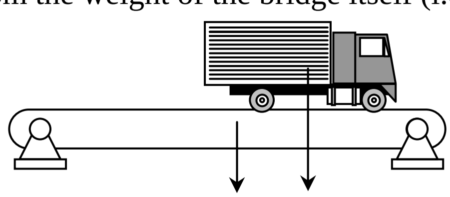


Truss: the supporting structure that distributes the load across several members.

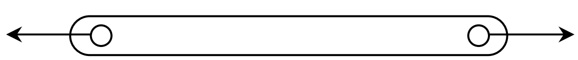
Abutment: the endpoints of the bridge that hold up the structure.
Top Cord: the top members of the bridge.
Bottom Cord: the bottom members of the bridge.
Span: the distance the bridge covers.
Joint: the point at which members are joined together.
Member: an individual metal beam.
Deck: the roadway or surface that carries traffic over the bridge.

Internal Member Forces

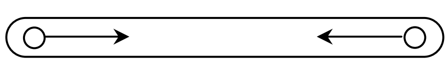
1. Load – the force from the weight of the bridge itself (i.e., all the members), plus cars and pedestrians.



2. Tension – stretching a material in opposite directions. Tension makes a member *longer*.



3. Compression – squeezing a material together in the same direction. Compression makes a member *shorter*.



Material Science

The steel that makes up each individual member is based on a cost/strength analysis. In general, steel strength and cost have a linear relationship. In other words, the following rule of thumb applies: *the stronger the steel, the more expensive it is*. There are three types of steel used in bridge construction:

Carbon Steel

Carbon steel only adds one alloy to the metal, namely carbon. An alloy is an element that is added during fabrication, which serves to strengthen naturally occurring iron. Carbon steel is the weakest and cheapest of the three metal alloys.

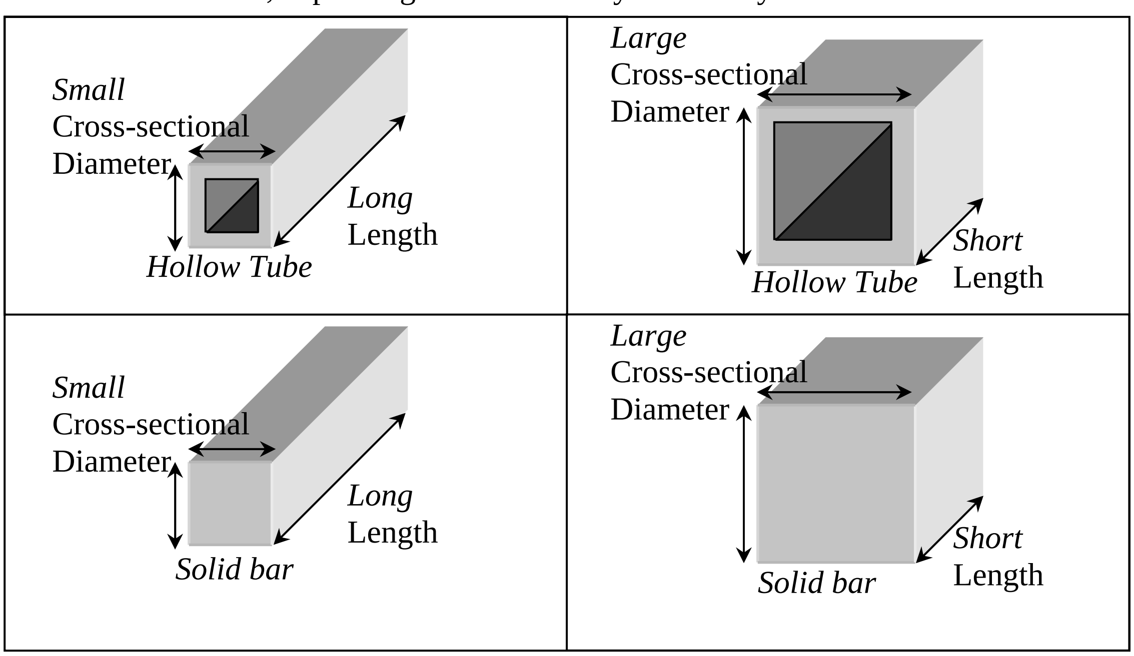
High-Strength Low-Alloy Steel (HSLA)

High-strength low-alloy steel also adds a low amount of carbon to the iron, but it takes the fabrication one step further by adding other alloys. High-strength low-alloy is stronger and more expensive than carbon steel.

Quenched and Tempered Low-Alloy Steel (QTCS)

Quenched and tempered low-alloy steel share similar properties with high-strength low-alloy steel in that both use a small percentage of alloys. *Tempered* refers to the process of heat treatment, which helps improve the strength and resistance against corrosion. *Quenching* refers to a cooling process that further hardens the steel by immersing it in oil or water. This is the strongest and most expensive of the three types of steel.

Aside from the type of steel, individual members are also made in different shapes, lengths, and cross-sectional diameters, depending on the load they must carry.



The shape (hollow/solid, length, cross-sectional diameter) of an individual member affects both the strength and expense. The relationships are summarized below:

	Inexpensive		*Expensive*
1	Hollow Tube	<	Solid Bar
2	Shorter Member	<	Longer Member
3	Smaller Cross Section	<	Larger Cross Section

1. Hollow tubes are cheaper to produce than solid bars. It costs less to produce a hollow tube because there is less material needed to make a hollow beam.

2. Shorter members are cheaper than longer members because less material is needed for a shorter beam.

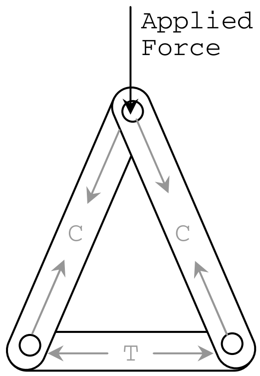
3. Similarly, members with a smaller cross-section are cheaper than members with a large cross-section.

The strength of a member depends, in large part, to the stress it is experiencing. Thus, the table above can be summarized for each type of stress (i.e., *compression* & *tension*).

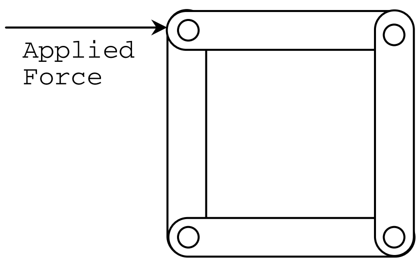
	Compression		
	Weaker		*Stronger*
1	Hollow Tube	<	Solid Bar
2	Longer Member	<	Shorter Member
3	Smaller Cross Section	<	Larger Cross Section
	Tension		
	Weaker		*Stronger*
4	Hollow Tube	<	Solid Bar
5	Shorter Member	=	Longer Member
6	Smaller Cross Section	<	Larger Cross Section

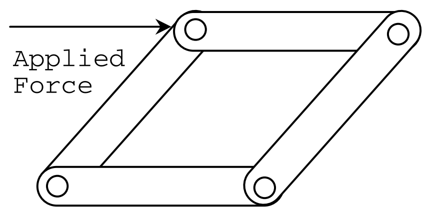
Configuration

In addition to the individual member characteristics, their configuration also has an impact on a bridge's ability to efficiently carry a load. If a load is applied to the simple configuration shown below, the diagonal bars will be under compression (C), and the bar that connects them will be under tension (T).

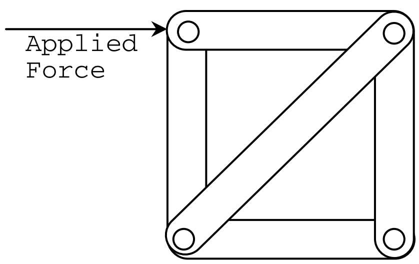


Consider the very simple example of four members, pinned together in the configuration shown below. If we place a load on the side of the top member, the shape will easily come out of a square configuration.





To make the square structure less susceptible to losing its shape, we can add another member.



By adding a cross-brace, we have effectively transformed the square into two interlocking triangles. Because a triangle does not come out of configuration when a side load is applied, it is the basis for all bridge designs.

APPENDIX D

Instructions: Idea-Challenge-Respond [ICR]

In this experiment, we want you to interact with your partner in a specific way. When appropriate, we want you to (politely) challenge your partner's ideas. To make this style of interaction concrete, let's call it the ICR style of interaction. ICR stands for **I**dea, **C**hallenge, and **R**espond.

IDEA	The *Idea* phase of the interaction is to present ideas to your partner. You could also think of it as making a suggestion. The idea phase doesn't need to include complete ideas or suggestions, but they can be built up over several conversational turns.

CHALLENGE	The second phase, *Challenge*, starts when the listener raises a question or objection about an idea. Challenges could be of several forms, some of which might include "*How do you know* ____________?" or "*That might work, but let's try* ____________ *instead*." Challenges might consist of: **asking for clarification**, **requesting more information**, or to **disagree** with an Idea. If something doesn't seem quite right to you, definitely challenge your partner.

RESPOND	The *Respond* phase addresses the challenge or question. This is the speaker's opportunity to clarify his or her initial idea, or search for new information. If the challenger has an idea in mind, you might want to explore that as well.

The idea behind the ICR style of interaction is to motivate you to evaluate each other's ideas. I realize that ICR may not be the most natural way to interact, but it can still be a friendly exchange. Below is an example of three people, using the ICR method to make a prediction about which object will go faster:

1	Nicola:	Now press "ready". The top weight will go faster.	Idea
2	Paul:	Would it?	Challenge
3	Luke:	Yes, because it's smooth.	Response
4	Nicola:	Yes, because it's slippery, it'll go faster. Yes it does.	Response
5	Luke:	Why?	Challenge
6	Paul:	Because if there was a rough surface and the bottom one was on ice…	Response
7	Nicola:	If there was a rough surface, there's more friction, it would slow it down.	Response
8	Luke,Paul:	Yes.	

Obviously, a pre-requisite to this style of communication is to directly address each other, then challenge one another.

You will have an opportunity during the warm-up task to practice ICR. I will try to listen and encourage you to use ICR. When I hear it, then I will point it out.

APPENDIX E

Instructions: Idea-Elaborate-Respond [IER]

In this experiment, we want you to interact with your partner in a specific way. When appropriate, we want you to elaborate upon your partner's ideas. To make this style of interaction concrete, let's call it the IER style of interaction. IER stands for **I**dea, **E**laborate, and **R**espond.

IDEA	The *Idea* phase of the interaction is to present ideas to your partner. You could also think of it as making a suggestion. The idea phase doesn't need to include complete ideas or suggestions, but they can be built up over several conversational turns.

ELABORATE	The second phase, *Elaborate*, starts when the listener builds upon an idea in a significant way. Elaborations extend or make explicit another person's idea. You might elaborate an idea when you hear something that sounds incomplete, in which you feel you could make a significant contribution.

RESPOND	The *Respond* phase addresses the challenge or question. This is the speaker's opportunity to clarify his or her initial idea, or search for new information. If the challenger has an idea in mind, you might want to explore that as well.

The idea behind the IER style of interaction is to motivate you to elaborate each other's ideas. I realize that IER may not be the most natural way to interact, but you probably do it all the time. Below is an example of two people, using the IER method to solve an electric circuit problem:

1	Ben:	An electric circuit has got a voltage source too, hasn't it?	
2	Adam:	Yes, actually it has.	
3	Adam:	(draws)	
4	Adam:	And it consists of (writes)	
5	Adam:	The voltage source has, gives, gives…	Idea
6	Ben:	The voltage sources gives voltage.	Elaboration
7	Adam:	And energy.	Elaboration
8	Ben:	Yes also.	Response
9	Adam:	And current isn't it?	Idea
10	Adam:	The voltage source also gives current.	Elaboration
11	Ben:	And due to this current, there is energy.	Response

Obviously, a pre-requisite to this style of communication is to directly address each other, then challenge one another.

You will have an opportunity during the warm-up task to practice IER. I will try to listen and encourage you to use IER. When I hear it, then I will point it out.

APPENDIX F

Warm-up Task

In the first part of this study, you will get a chance to practice your style of interaction. To get accustomed to the style, we are going to use the following logic problem as a warm-up exercise. Try to collaborate with your partner to solve the following logic problem.

Curtis and four of his friends all went on vacation with their families last year. Each vacationed in a different state and each enjoyed participating in a different activity (see table below). From the clues, determine each child's name, the state where each spent their vacation, and the activity each enjoyed while away.

Clues:

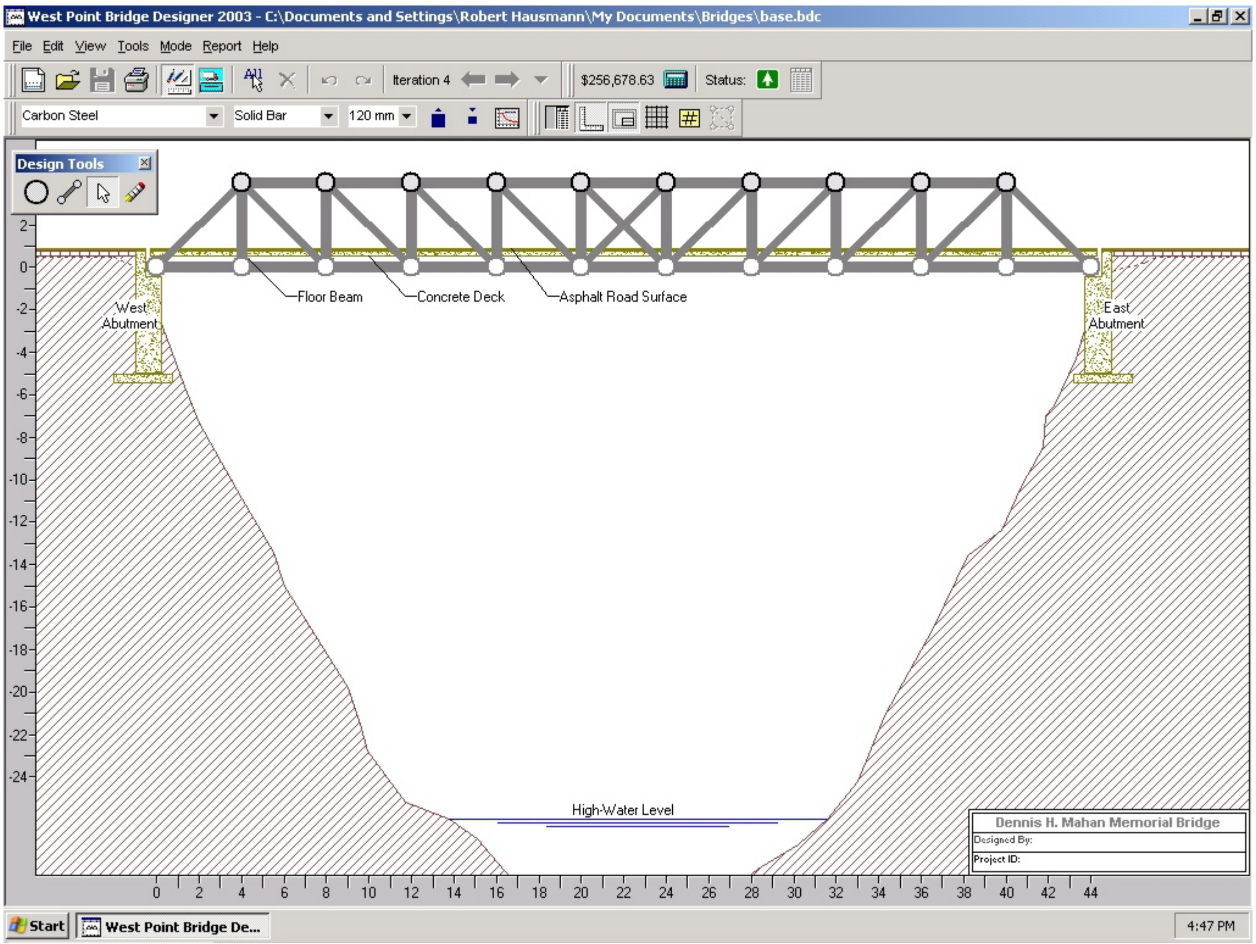
1. Susie and her family vacationed in Arkansas.
2. Reneé spent most her time horseback riding.
3. Michael, who didn't spend his vacation in Oklahoma, enjoyed spending time at the ice skating rink.
4. The girl who vacationed in Missouri with her family enjoyed hiking in the Ozark Mountains.
5. Colorado was the destination of the family who intended to spend most of their time skiing.

Solution:

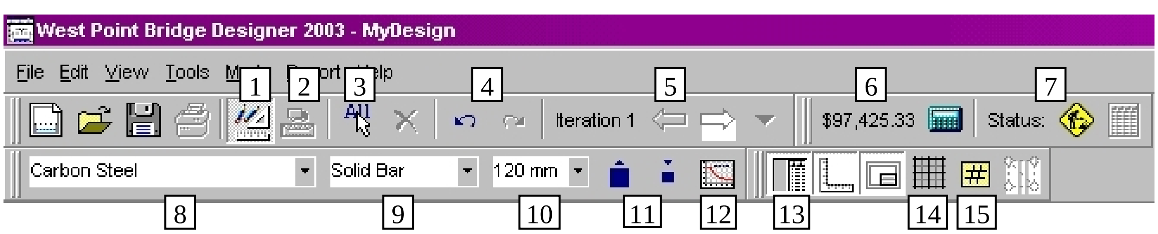
		Name					Activity				
		Amber	Susie	Reneé	Michael	Curtis	canoeing	ice skating	hiking	skiing	horseback
Location	Arkansas										
	Oklahoma										
	Missouri										
	Colorado										
	Nebraska										
Activity	canoeing										
	ice skating										
	hiking										
	skiing										
	horseback										

APPENDIX G

West Point Bridge Designer 2003 Screenshot: Design Mode



Toolbar: Buttons and Commands

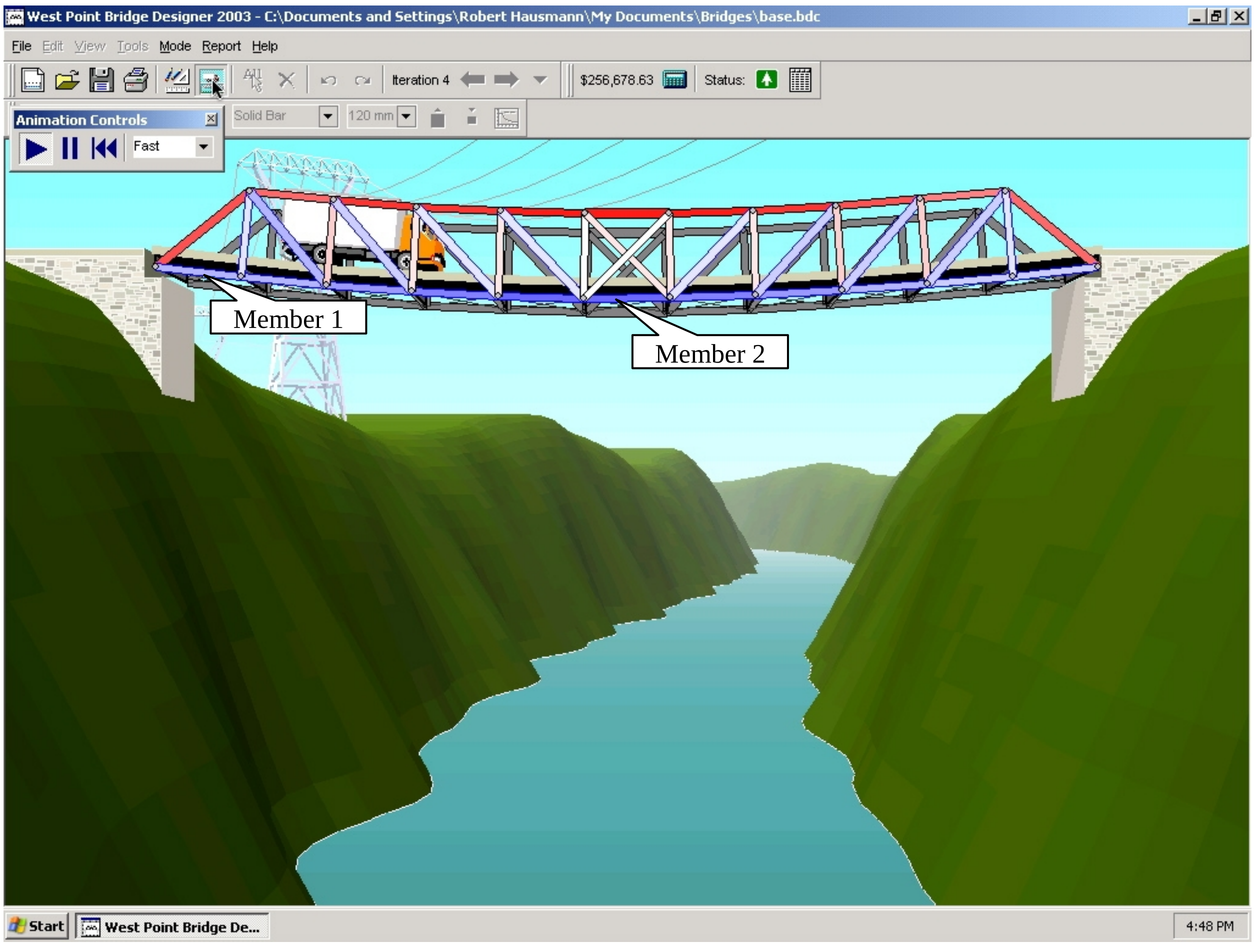


1. **Drawing Board –** the design mode allows the user to place joints and members.
2. **Load Test Mode –** this mode allows the user to simulate a load, which then generates.

3. **Select All –** selects all members in the current design.
4. **Undo –** undo user's last command.
5. **Iteration Counter –** allows the user to go back to previous designs.
6. **Current Bridge Price –** the current total price of the bridge.
7. **Bridge Status –** after a load test, this will symbolize if it withstood the test.
8. **Material –** use the drop-down menu to select a type of metal.
9. **Cross-Section –** use the drop-down menu to select the form (solid bar or hollow tube).
10. **Size –** use the drop-down menu to select the cross-section diameter.
11. **Increase/Decrease Member Size –** quickly increase or decrease the member's cross-sectional dimension.
12. **Member Properties Report –** this information shows member strength as a function of length in graphical form.
13. **Member List –** a table of numbered members which shows the tension and compression results from the load test.

APPENDIX H

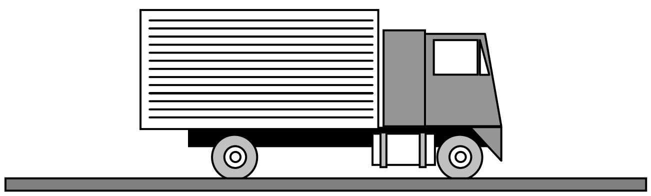
West Point Bridge Designer 2003 Screenshot: Load Test Mode



APPENDIX I

Bridge Optimization Instructions

During this part of the experiment, I want you to optimize the design of a preexisting bridge. The main issue you will face as an engineer is the **total cost** of your bridge. As in most real world designs, there is a finite budget that you must consider. The program automatically displays the price of your configuration in the upper right-hand corner of the screen. The ideal bridge is one that can carry a load (represented by the truck), while minimizing the cost. Your challenge is to build the cheapest bridge that can withstand the load of the truck.



The simulation uses the following color-coded feedback:

Red = compression
Blue = tension

The initial cost of the bride is: $256,678.63. Your goal is to **minimize** the cost.

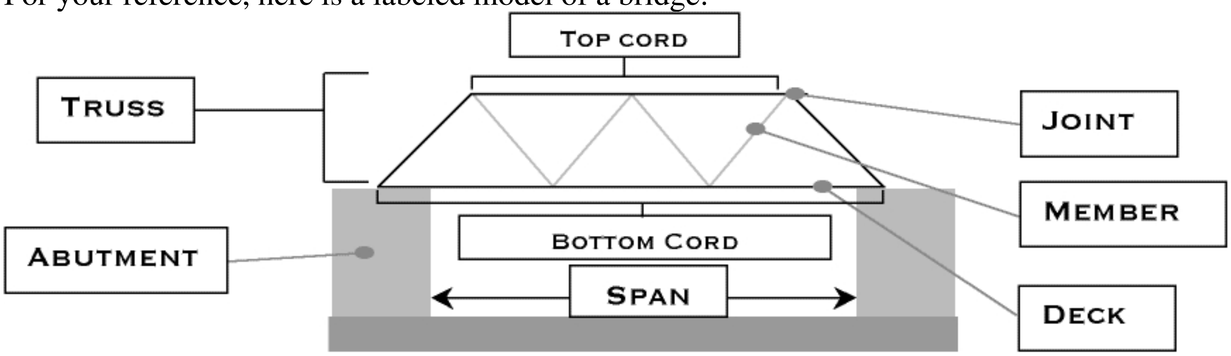
Do you understand what I am asking you to do? Do you have any questions for me?

APPENDIX J

Background Knowledge Questionnaire: Posttest

Participant ID Number:

For your reference, here is a labeled model of a bridge:



Provide a definition for the following terms:

- Load:
- Tension:
- Compression:
- Failure:

Rank order each type of steel, from weakest (1) to strongest (3):

Strength Rank Order	Steel Type
pull down	Carbon Steel
pull down	High-strength Low-alloy Steel
pull down	Quenched and Tempered Steel

Rank order each type of steel, from cheapest (1) to most expensive (3):

Cost Rank Order	Steel Type
pull down	Carbon Steel
pull down	High-strength Low-alloy Steel
pull down	Quenched and Tempered Steel

Indicate which material is more expensive (>) or less expensive (<).

1. Hollow Tube	pull down	Solid Bar
2. Longer Member	pull down	Shorter Member
3. Smaller Cross-section	pull down	Larger Cross-section

Multiple Choice (*Choose the best answer for each question*)

1. As cross-sectional dimension increases, member strength ____________.

    (a) increases
    (b) stays the same
    (c) decreases

2. Hollow tubes have ____________ tensile strength than solid bars.

    (a) higher
    (b) same
    (c) lower

3. Tensile strength is always ____________ the maximum compressive strength.

    (a) greater than
    (b) the same as
    (c) weaker than

4. Under compression, longer members are ____________ shorter members.

    (a) stronger than
    (b) the same as
    (c) weaker than

5. Suppose we took a bridge and made it taller (in the vertical plane), without changing the overall configuration. The internal member forces of the top and bottom cords will ____________.

(a) increases
(b) stays the same
(c) decreases

6. The top of a bridge is always under ____________, while the bottom is always under ____________.

(a) compression; compression
(b) compression; tension
(c) tension; compression
(d) tension; tension

7. Strength-to-force ratio: a value ____________ one means the member has failed, while a value ____________ one means the member can safely carry the load.

(a) greater than; greater than
(b) greater than; less than
(c) less than; greater than
(d) less than; less than

8. Under tension, longer members are ____________ shorter members.

(a) stronger than
(b) the same strength as
(c) weaker than

9. Where does the bridge experience the most stress?

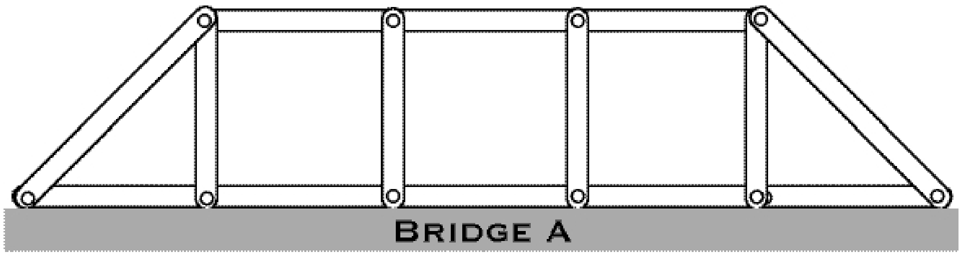
(a) The top cord
(b) The middle
(c) The bottom cord
(d) The two ends

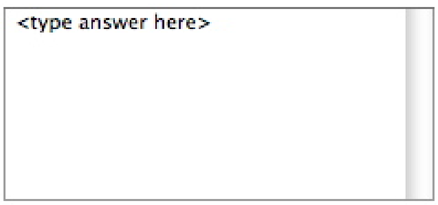
10. Using members of several different sizes ____________ the overall cost of the bridge.

(a) increases
(b) does not change
(c) decreases

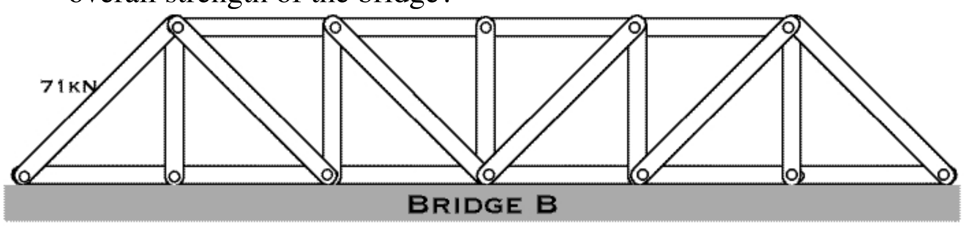
Short Answer

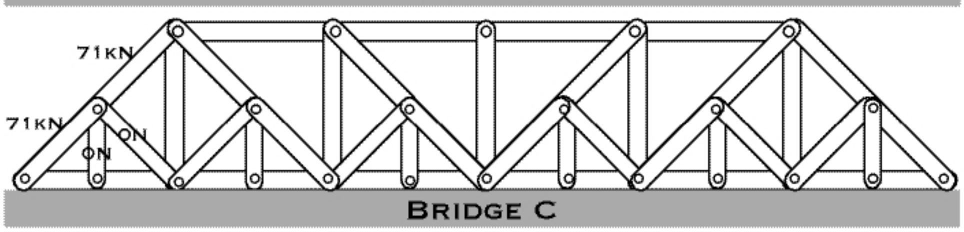
1. The bridge below (Bridge A) is flawed in a significant way. What might be done to the bridge to allow it to carry a load? (Describe your modifications to the bridge in the textbox below.)

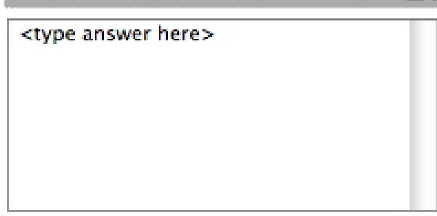




2. Consider two bridges shown below (Bridge B and Bridge C). Their configurations are similar, with one exception. Bridge C has 12 additional members. Although none of the internal member forces in the diagonals (71kN) change, the bridge is stronger. Why does adding supports, which themselves experience no internal member forces, increase the overall strength of the bridge?







3. Since we have considered the members under compression, describe what you might do to the members under tension (Bridge C), assuming they are made of solid carbon steel bars (160mm)?

<type answer here>

4. Have you ever built a bridge before, either for a class project or just for fun?

   (a) yes
   (b) no

   If so, how long ago did you build your ______ bridge? years

5. Have you ever entered a bridge building contest?

   a) yes
   b) no

If so, what was the outcome?

<type answer here>

Submit Reset

APPENDIX K

Content Analysis of West Point Bridge Designer 2003 Simulation

Concept	*Sim*	*Text*	*Pretest*	*Posttest*
Strength				
1. Members with a larger cross-section are **stronger** than members with a smaller cross-section.	mMPR	p.3-B	MC1	MC1
2. Solid members are **stronger** than hollow members (holding cross-section constant).	mMPR	p.3-B	MC2	MC2
3. Tensile **strength** is always greater than compressive strength.	MPR	**Inf**	**MC3**	**MC3**
4. Shorter members carry a **higher load** than longer members under compression.	MPR	p.3-B	MC4	MC4
5. Tensile **strength** is constant across member length.	MPR	p.3-B	MC8	MC8
6. Compressive **strength** is variable across member length (sigmoid function: tanq).	MPR	**Inf**		
7. **Strength**: CS < HSS < QTS	mMPR	p.2-T	RO1	RO1
			RO2	**RO2**
			RO3	RO3
Price				
8. Using fewer bars of different size decreases the fabrication **cost**.	Δsim; CBP	p.3-T	MC10	MC10
9. Hollow members are **cheaper** than solid members (holding cross-section constant).	mMPR	p.3-T	GTLT1	GTLT1
10. Shorter members are **cheaper** than longer members.	mMPR	p.3-T	GTLT2	GTLT2
11. Members with a larger cross-section are more **expensive** than members with a smaller cross-section.	mMPR	p.3-T	GTLT3	GTLT3
12. **Cost**: CS < HSS < QTS	mMPR	p.2-T	RO4	RO4
			RO5	**RO5**
			RO6	RO6
13. To minimize cost, bring the compression(tension) /strength ratio as close to 1.00 as possible without going over (p. 4-23).	Δsim; infer #17.c.	**Inf**		
Stress (Tension & Compression)				
14. Making a bridge taller (vertically), without changing the overall configuration, will decrease the internal member forces in the top and bottom cords (p. 4-30).	Δsim	**Inf**	**MC5**	**MC5**
15. The top of a bridge is always under compression.	CFB; ML	**Inf**	**MC6a**	**MC6a**

16. The bottom of a bridge is always under tension.	CFB; ML	**Inf**	**MC6b**	**MC6b**
17. Values for the compression(tension)/strength ratio:				
a. < 1.00 safely carries load	ML	**Inf**	**MC7a**	**MC7a**
b. > 1.00 member failure	ML	**Inf**	**MC7b**	**MC7b**
c. = 1 optimal design (within factor of safety)	Δsim	Inf		
18. The center of the bridge is under more stress than the ends.	CFB; ML	**Inf**	**MC9**	**MC9**
19. To strengthen a member under compression, add a joint somewhere along the beam (p. 4-30).	Δsim; infer #4	**Inf**	**SA2**	**SA2**
20. Increase the strength of steel and reduce the size of the members under tension.	Δsim	**Inf**	**SA3**	**SA3**
a. The same is NOT true for members under compression.	Δsim			
Configuration				
21. Squares are easily deformed.	EFB	p.4	SA1	SA1
22. Squares can be transformed into rigid structures by forming triangles using cross bracing.	EFB	p.4	SA1	SA1
23. Triangles hold their shape (i.e., not easily deformed).	EFB	p.4	SA1	SA1
24. Triangles are the basis of any bridge design.	Infer 21	p.4	SA1	SA1

Source of information from the simulation Key:

MPR = Member Property Report (corresponds to button #12 in APPENDIX G)

mMPR = Multiple Reports

CFB = Color Coded Feedback (corresponds colors in the test mode, see APPENDIX H)

ML = Member List (corresponds to button #13 in APPENDIX G)

EFB = Error Message Feedback (error feedback from the simulation)

CBP = Current Bridge Price (corresponds to item #6 of APPENDIX G)

Text Key (refers to pages of text found in APPENDIX C):

T = Top of page

B = Bottom of page

Inf = Inference concept

Pre/Posttest Key (# refers to item on pre/post-test):

MC = Multiple Choice

SA = Short Answer

RO = Rank Order

GTLT = Greater Than, Less Than

SA = Short Answer

BIBLIOGRAPHY

Aleven, V., & Koedinger, K. (2000). The need for tutorial dialog to support self-explanation. In C. P. Rose & R. Freedman (Eds.), *Building dialogue systems for tutorial applications. Papers from the 2000 aaai fall symposium* (pp. 65-73). Menlo Park, CA: AAAI Press.

Anderson, A., Mayes, J. T., & Kibby, M. R. (1995). Small group collaborative discovery learning from hypertext. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 23-38). Berlin: Springer-Verlag.

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ashley, K. D., Desai, R., & Levine, J. M. (2002). *Teaching case-based argumentation concepts using dialectic arguments vs. Didactic explanations.* Paper presented at the Intelligent Tutoring Systems Conference, ITS ’02.

Azmitia, M. (1996). Peer interactive minds: Developmental, theoretical, and methodological issues. In P. B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 133-162). Cambridge: Cambridge University Press.

Azmitia, M., & Crowley, K. (2001). The rhythms of scientific thinking: A study of collaboration in an earthquake microworld. In K. Crowley, C. D. Schunn & T. Okada (Eds.), *Designing for science: Implications for everyday, classroom, and professional settings* (pp. 47-77). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Barron, B. (2000a). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*(4), 403-436.

Barron, B. (2000b). Problem solving in video-based microworlds: Collaborative and individual outcomes of high-achieving sixth-grade students. *Journal of Educational Psychology, 92*(2), 391-398.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307-359.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of robert glaser* (pp. 395-451). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Browne, M. N., & Hausmann, R. G. (1998). The friendly sound of critical thinking. *The Korean Journal of Thinking & Problem Solving*, *8*(2), 47-53.

Browne, M. N., Hausmann, R. G. M., & Ostrowski, N. L. (2002). Reframing argument as an act of friendship: Overcoming student reluctance to apply critical thinking. *The Korean Journal of Thinking and Problem Solving, 12*(2), 85-94.

Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, *15*(1), 1-40.

Chan, C., Burtis, J., Scardamalia, M., & Bereiter, C. (1992). Constructive activity in learning from text. *American Educational Research Journal*, *29*(1), 97-118.

Chan, C. K. K. (2001). Peer collaboration and discourse patterns in learning from incompatible information. *Instructional Science, 29*, 443-479.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The Journal of the Learning Sciences*, *14*(2), 161-199.

Chi, M. T. H., & Bassok, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of robert glaser* (pp. 251-282). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989a). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M. T. H., DeLeeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989b). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly, 35*(1), 27-62.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471-533.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences, 1*(1), 69-105.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18*(3), 257-273.

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching, 35*(6), 623-654.

Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Experimental Education, 69*(1), 77-97.

Clark, H. H., & Brennen, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-119). Washington, DC: American Psychological Association.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294.

Clark, S. E., Hori, A., Putnam, A., & Martin, T. P. (2000). Group collaboration in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1578-1588.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1-35.

Cohen, S. G., & Baily, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management, 23*(3), 239-290.

Coleman, E. B. (1998). Using explanatory knowledge during collaborative problem solving in science. *The Journal of Learning Sciences, 7*(3&4), 387-427.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology, 497-509.*

Dillenbourg, P. (1999). What do you mean "collaborative learning"? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier.

Dillenbourg, P. (2002). Over-scripting cscl: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of cscl. Can we support cscl* (pp. 61-91). Heerlen: Open Universiteit Nederland.

Druyan, S. (2001). A comparison of four types of cognitive conflict and their effect on cognitive development. *International Journal of Behavioral Development, 25*(3), 226-236.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397-434.

Fischer, F., Bruhn, J., Graesel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction, 12*, 213-232.

Forman, E. A., & Cazden, C. B. (1985). Exploring vygotskian perspectives in education: The cognitive value of peer interaction. In J. W. Wertsch (Ed.), *Culture, communication, and cognition* (pp. 323-347). New York: Wiley.

Gould, O., Kurzman, D., & Dixon, R. A. (1994). Communication during prose recall conversations by young and old dyads. *Discourse Processes, 17*, 149-165.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Hamilton, R. J. (1997). Effects of three types of elaboration on learning concepts from text. *Contemporary Educational Psychology, 22*, 299-318.

Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 331-348). Washington, DC: American Psychological Association.

Hausmann, R. G. M., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology, 7*(1), 4-14.

Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 21-41). New York: Lawrence Erlbaum Associates.

Hayes, J. R., & Simon, H. A. (1979). Understanding written problem instructions. In H. A. Simon (Ed.), *Models of thought* (Vol. 1, pp. 451-476). New Haven, CT: Yale University Press.

Hill, G. W. (1982). Group versus individual performance: Are n + 1 heads better than one? *Psychological Bulletin, 91*(3), 517-539.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), 379-432.

Hutchins, E. (1995). How a cockpit remembers its speed. *Cognitive Science, 19*, 265-288.

Infante, D. A., & Rancer, A. S. (1987). Argumentativeness scale. In K. Corcoran & J. Fischer (Eds.), *Measures for clinical practice: A sourcebook* (pp. 36-37). New York: Free Press.

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.

Johnson, D. W., & Johnson, R. T. (1992). Positive interdependence: Key to effective cooperation. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 174-199). Cambridge, England: Cambridge University Press.

Karmiloff-Smith, A. (1988). The child is a theoretician, not an inductivist. *Mind & Language, 3*(3), 183-195.

Katz, S., Allbritton, D., & Connelly, J. (2003). Going beyond the problem given: How human tutors use post-solution discussions to support transfer. *International Journal of Artificial Intelligence in Education, 13*, 79-116.

Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogs. *Cognition and Instruction, 18*(1), 53-81.

Keeley, S., Browne, M. N., & Kreutzner, J. (1982). A comparison of freshmen and seniors on general and specific essay tests of critical thinking. *Research in Higher Education, 17*(2), 139-154.

Kim, S.-I., & Van Dusen, L. M. (1998). The role of prior knowledge and elaboration in text comprehension and memory: A comparison of self-generated elaboration and text-provided elaboration. *American Journal of Psychology, 111*(3), 353-378.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1-48.

Kollar, I., Fischer, F., & Slotta, J. D. (2005). Internal and external collaboration scripts in webbased science learning at schools. In T. Koschmann, D. D. Suthers & T. W. Chan (Eds.), *Computer supported collaborative learning 2005: The next 10 years!* (pp. 331-340). Mahwah, NJ: Lawrence Erlbaum Associates.

Kruger, A. C. (1993). Peer collaboration: Conflict, cooperation, or both? *Social Development, 2*(3), 165-182.

Larson, C. O., Dansereau, D. F., O'Donnell, A. M., Hythecker, V. I., Lambiotte, J. G., & Rocklin, T. R. (1985). Effects of metacognitive and elaborative activity on cooperative learning and transfer. *Contemporary Educational Psychology, 10*, 342-348.

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*, 605-620.

Laughlin, P. R., Zander, M. L., Knievel, E. M., & Tan, T. K. (2003). Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective strategies. *Journal of Personality and Social Psychology, 85*(4), 684-694.

Lee, G., Kwon, J., Park, S.-S., Kim, J.-W., Kwon, H.-G., & Park, H.-K. (2003). Development of an instrument for measuring cognitive conflict in secondary-level science classes. *Journal of Research in Science Teaching, 40*(6), 585-603.

Leitao, S. (2003). Evaluating and selecting counterarguments. *Written Communication, 20*(3), 269-306.

Levelt, W. J. M. (1999). Models of word production. *TRENDS in Cognitive Sciences, 3*(6), 223-232.

Light, P., Littleton, K., Bale, S., Joiner, R., & Messer, D. (2000). Gender and social comparison effects in computer-based problem solving. *Learning and Instruction, 10*, 483-496.

Lorge, I., Fox, D., Davitz, J., & Brenner, M. (1958). A survey of studies contrasting the quality of group performance and individual performance. *Psychological Bulletin, 55*, 337-372.

Markman, A. B. (1997). Constraints on analogical mapping. *Cognitive Science, 21*(4), 373-418.

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359-377.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25*(1), 95-111.

Meudell, P. R., Hitch, G. J., & Boyle, M. M. (1995). Collaboration in recall: Do pairs of people cross-cue each other to produce new memories? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 48A*(1), 141-152.

Miller, S. A., & Brownell, C. A. (1975). Peers, persuasion, and piaget: Dyadic interaction between conservers and nonconservers. *Child Development, 46*, 992-997.

Moreland, R. L. (2005, May). *The strategic use of transactive memory systems in groups.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 57-84). Thousand Oaks, CA: Sage Publications, Inc.

Moreland, R. L., & Levine, J. M. (1992). Problem identification by groups. In S. Worchel, W. Wood & J. A. Simpson (Eds.), *Group process and productivity* (pp. 17-47). Newbury Park, CA: Sage.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, N.J.: L. Erlbaum Associates.

Norman, D. A. (1993). Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science, 17*, 1-6.

O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). London: Cambridge University Press.

O'Donnell, A. M., Dansereau, D. F., Rocklin, T. R., Hythecker, V. I., Lambiotte, J. G., Larson, C. O., et al. (1985). Effects of elaboration frequency on cooperative learning. *Journal of Educational Psychology, 77*(5), 572-580.

O'Reilly, T., Symons, S., & MacLatchy-Gaudet, H. (1998). A comparison of self-explanation and elaborative interrogation. *Contemporary Educational Psychology, 23*(4), 434-445.

Okada, T., & Simon, H. A. (1997). Collaboration discovery in a scientific domain. *Cognitive Science, 21*(2), 109-146.

Oyster, C. K. (2000). *Groups: A user's guide*. Boston: McGraw Hill.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117-175.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81*, 639-646.

Pirolli, P., & Bielaczyc, K. (1989). Empirical analyses of self-explanation and transfer in learning to program. In C. M. Olson & E. E. Smith (Eds.), *Proceedings of the 11th* *annual conference of the cognitive science society* (pp. 450-457). Hillsdale, NJ: Erlbaum Associates, Inc.

Pontecorvo, C., & Girardet, H. (1993). Arguing and reasoning in understanding historical topics. *Cognition and Instruction, 11*(3-4), 365-395.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 221-227.

Reder, L. M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. *Review of Educational Research, 50*(1), 5-53.

Reimann, P., & Neubert, C. (2000). The role of self-explanation in learning to use a spreadsheet through examples. *Journal of Computer Assisted Learning, 16*, 316-325.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*(1), 1-29.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*(1), 90-108.

Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction, 11*(3 & 4), 347-364.

Ressler, S. J. (2002). *Designing and building file-folder bridges: A problem-based introduction to engineering*. Defense Dept., Army: United States Military Academy.

Rips, L. J. (1990). Reasoning. *Annual Review of Psychology, 41*, 321-354.

Salomon, G. (1993). No distribution without individuals' cognition: A dynamic interactional view. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 111-138). Cambridge: Cambridge University Press.

Scanlon, E. (2000). How gender influences learners working collaboratively with science simulations. *Learning and Instruction, 10*, 463-481.

Schwartz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right.If they argue! *Cognition and Instruction, 18*(4), 461-494.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences, 4*(3), 321-354.

Seifert, T. L. (1993). Effects of elaborative interrogation with prose passages. *Journal of Educational Psychology, 85*(4), 642-651.

Setlock, L. D., Fussell, S. R., & Neuwirth, C. (2004). *Taking it out of context: Collaborating within and across cultures in face-to-face settings and via instant messaging.* Paper presented at the 2004 ACM conference on Computer supported cooperative work, Chicago, Il.

Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science, 26*, 469-501.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 592-604.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.

Stark, R., Mandl, H., Gruber, H., & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning and Instruction*, *12*(1), 39-60.

Stasser, G., Vaughan, S. I., & Stweart, D. D. (2000). Pooling unshared information: The benefits of knowing how access to information is distributed among group members. *Organizational Behavior and Human Decision Processes, 82*(1), 102-116.

Stein, B. S., & Bransford, J. D. (1979). Constraint on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18*, 769-777.

Stroebe, W., & Diehl, M. (1994). Why groups are less effective than their members: On productivity losses in idea-generating groups. *European Review of Social Psychology, 5*, 271-304.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295-312.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 51*, 207-220.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *The computer as a cognitive tool* (pp. 229-258). Hillsdale, NJ: Erlbaum.

Toulmin, S. E. (1958). *The uses of argument*. London/New York: Cambridge University Press.

Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19*, 123-138.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 381-391.

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*, 311-330.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527-579). Cambridge: MIT Press.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45-69.

Voss, J. F., Greene, T. R., Post, T. A., & Penner, B. C. (1983). Problem-solving skill in the social sciences. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research theory* (Vol. 17, pp. 165-213). New York: Academic Press.

Voss, J. F., Tyler, S. W., & Yengo, L. A. (1983). Individual differences in the solving of social science problems. In R. F. Dillon & R. R. Schmeck (Eds.), *Individual differences in cognition* (Vol. 1, pp. 205-232). New York: Academic.

Vroom, V. H., Grant, L. D., & Cotton, T. S. (1969). The consequences of social interaction in group problem solving. *Organizational Behavior and Human Decision Processes, 4*(1), 77-95.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21-39.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York, NY, USA: Macmillan Library Reference Usa; London, England UK: Prentice Hall International.

Wong, B. Y. L. (1985). Self-questioning instructional research: A review. *Review of Educational Research, 55*(2), 227-268.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87-122.