The Effects of Argumentation Scaffolds on Argumentation and Problem Solving

Kyoo-Lak Cho
David H. Jonassen

An important skill in solving problems, especially ill-structured problems, is the production of coherent arguments to justify solutions and actions. Because direct instruction in argumentation has produced inconsistent results and cannot effectively support online learning, we examined the use of online argumentation scaffolds to engage and support coherent argumentation. In this study, we showed that providing a constraint-based argumentation scaffold during group problem-solving activities increased the generation of coherent arguments. The same scaffold further resulted in significantly more problem-solving actions during collaborative group discussions. The effects of the scaffold varied for problem type. Groups that solved ill-structured problems produced more extensive arguments. When solving ill-structured problems, students need more argumentation support because of the importance of generating and supporting alternative solutions. The close relationship between argumentation and problem solving, especially ill-structured problem solving, is significant. The effects of the argument scaffold consistently transferred to the production of arguments during individual problem solving. Students used the familiar argumentation scripts while solving problems individually.

Problem solving requires argumentation (Cerbin, 1988). Argumentation is a process of making claims and providing justification for the claims using evidence (Carr, 1999; Toulmin, 1958). Argumentation is an essential kind of informal reasoning that is central to the intellectual ability involved in solving problems, making judgments and decisions, and formulating ideas and beliefs (Kuhn, 1991). Argumentation requires problem solvers to identify various alternative perspectives, views, and opinions; develop and select a preferred, reasonable solution; and support the solution with data and evidence (Voss, Lawrence, & Engle, 1991). Argumentation is a variable that significantly predicts student performance in both well-structured and ill-structured problems (Hong, Jonassen, & McGee, in press).

Despite its importance, students are not adept at constructing cogent arguments (Cerbin, 1988). Specifically, they do not connect evidence to claims via warrants (claim = tax cut will increase savings; warrant = marginal propensity to consume; evidence = decrease in savings last year and increase in sales tax revenue), but reasoning from claims to evidence is essential for problem solving. Bell and Linn (2000) suggest that conjecturing with warrants, as opposed to descriptions, in order to support arguments indicates that students are making scientific conjectures, which enables them to generate better problem solutions.

How then do we facilitate learner development of argumentation skills? Cerbin (1988) proposed direct instruction of reasoning skills based on an explicit model of argumentation. That approach has been the standard method for years. Leeman (1987) and Saunders (1994) advo-

cated using Toulmin's (1958) model of argument in law education class to develop argumentation skills. Several researchers have also advocated direct instruction on the structure and notation of argumentation (Knudson, 1991; Sanders, Wiseman, & Gass, 1994; Yeh, 1998). However, research findings show inconsistent results: Direct instruction does not always improve argumentation skills as expected. Some research indicates that direct instruction enhances argumentation skills (Sanders et al., 1994), whereas other research demonstrates no positive effects for direct instruction on improving argumentation skills (Knudson, 1991).

An alternative approach to developing argumentation skills is to scaffold argumentation skills through the use of cognitive tools. Lajoie and Lesgold (1992) demonstrated the cognitive effects of Sherlock, a computer-based argumentation generation tool in avionics troubleshooting, showing that the group that used Sherlock out-performed the control group in terms of the number of troubleshooting problems solved and in the quality of processes. A qualitative analysis of the treatment group revealed that the quantity and quality of troubleshooting steps became more expert-like as a function of experience with Sherlock (Lajoie, 1993).

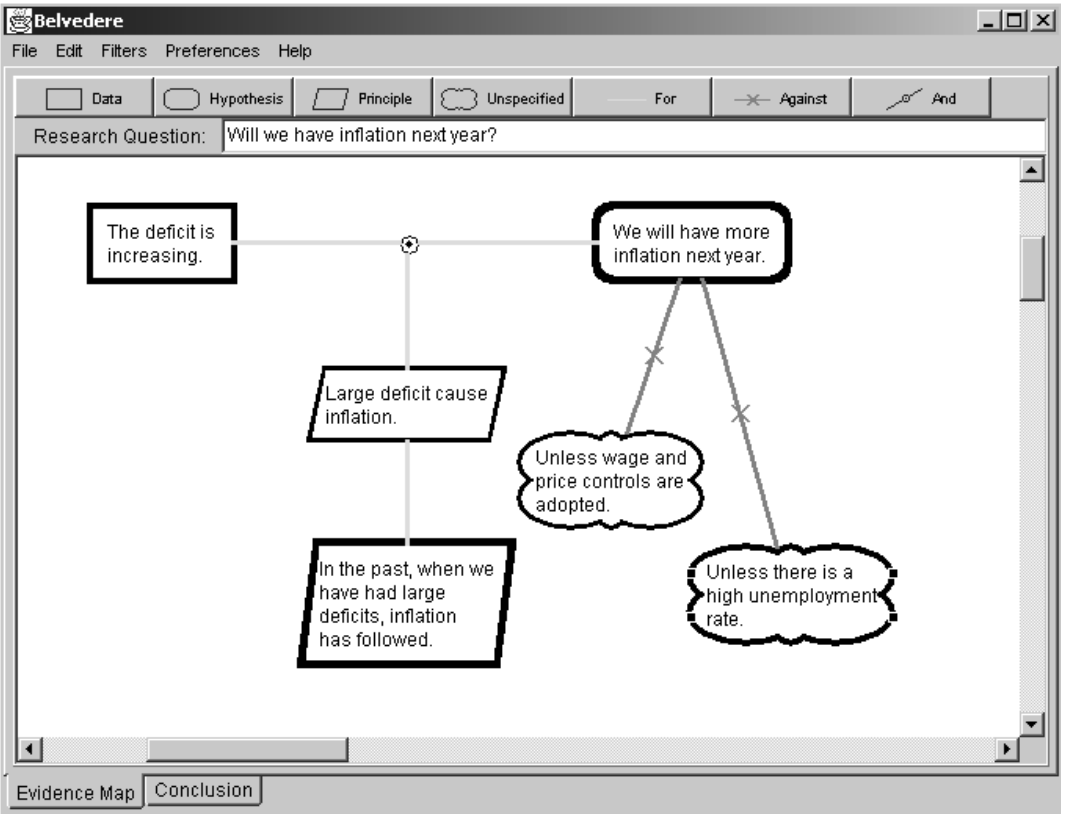
A new class of cognitive tools, computer-supported collaborative argumentation (CSCA) software, has emerged for the purpose of scaffolding student's seeking of warrants and evidence for supporting claims. Scaffolding, for purposes of this study, represents any manipulation by the system of the task itself (Jonassen, 1999). When scaffolding performance, the system performs part of the task for the student, supplants the student's ability to perform some part of the task by changing the nature of the task or imposing the use of cognitive tools that help the learner perform, or adjusts the nature or difficulty of the task. Scaffolds are temporary frameworks to support student performance beyond their capacities, in this case, constructing arguments.

The rationale for using CSCA to scaffold argumentation during conversations is in the theory of constraints. That is, CSCA scaffolds argumentation by *constraining* the nature of student interactions between group members, requiring them to preclassify their contributions to an argumentative discussion. Many other cognitive activities are also constraint based. For example, reading researchers have explored syntactic and lexical constraints on meaning generation while parsing sentences. In problem-solving research, constraints are the set of possible combinations of values between variables in the problem that must be progressively restricted (satisfied) during problem solving (Darses, 1991; Richard, Poitrenaud, & Tijus, 1993). Chi, Slotta, and de Leeuw (1994) described constraint-based interactions between values in the system. Those interactions can be defined canonically, which is what CSCA systems do. Any conceptual system (from a simple sentence to a complex conceptual domain) consists of attributes with values that interact. Those interactions impose constraints on the psychological processes required to operate on that system. Those constraints must be satisfied or eliminated in order for the processes to be completed. In this study, we sought to constrain the nature of interactions among learners in an online discussion forum by requiring them to predetermine the nature of the message they were contributing to the forum in argumentative terms. The constraints were the predetermined message types that modeled a form of argumentation.

So, constraint-based CSCA conversation scaffolds are prestructured forms of conversation systems that impose different conversational ontologies onto the discussion. These ontologies make explicit the constraints involved in the conversation. They supply the explicit statements of the interactions among the attributes in the domain. In this study, we preclassified the types of comments that students could make while solving problems, which constrained the nature of verbal interactions among conversants. The Belvedere environment used in this study, for example (see Figure 1), provides four predefined argumentation constraints: (a) hypothesis, (b) data, (c) principles, and (d) unspecified; and three links: (a) for, (b) against, and (c) and (Suthers, 1998). These constraints form the relations between the ideas that conversants produce.

Limited research has examined the use of

Figure 1 ☐ Illustration of Belvedere argument being constructed.



constraint-based argumentation scaffolds. Tan (2000) reported that students who used QuestMap™, a constraint-based argumentation tool, performed significantly better in stating "grounds" in their argumentation. Grounds are one of the argument components of Toulmin's model of argument (Toulmin, Rieke, & Janik, 1984). However, the research results did not show significant effects of QuestMap™ on problem-solving skills. Suthers and Hundhausen (2001) found no differences between constraint-based systems conveyed in graphical node-link system used in this study versus a tabular layout of the information. Using discourse analysis, Veerman, Andriessen, and Kanselaar (1999) compared and analyzed the content of small group online collaborative discussions carried out by college students. Three online learning tools were provided as a means of communication and discussion, including Netmeeting, an unstructured synchronous online chatting tool, Belvedere, a structured synchronous CSCA tool, and Allaire Forum, an asynchronous bulletin board system (BBS). The content of the group discussions was characterized in terms of their constructive and argumentative contributions and their focus on the meaning and application of concepts. They found that Belvedere discussions were the most argumentative because Belvedere stimulates students to check and counter each other's information most frequently. This finding implies that constraint-based tools are able to scaffold argument development during group problem-solving activities, however the residual effects on content knowledge remain unclear. Buckingham-Shum, MacLean, Bellotti, and Hammond (1997) conducted a series of qualitative studies in order to identify how QOC (Questions, Options, and Criteria), a constraint-based structure and notation system, supports particular kinds of design problem solving and hinders others. They reported that QOC provided the most support when designers dealt with poorly understood (ill-structured)

design problems, unlike the findings of Tan (2000). But the same QOC also distracted the designers' activities when they dealt with well-understood (well-structured) design problems. These research results are not consistent, in part, because their structures and functions vary. Research is needed to clarify these differences.

Purpose of Study

There are many issues that are left unresolved by the limited amount of research on argumentation scaffolds and problem solving. Those issues are restated here in terms of four research questions:

1. How do constraint-based scaffolds affect the nature of argumentation during a problem-solving activity when compared with an unconstrained discussion board? How is the quality of student arguments in the discussion forum affected by the constraint-based tool? We predicted that discussions in the constraint-based groups would include more argumentation than those in the unconstrained discussion groups.
2. How do constraint-based scaffolds affect the nature of problem-solving activity during discussions? While the relationship of argumentation to problem solving is well established conceptually, students' arguments have never been analyzed in terms of their problem-solving activities. We predicted that students in the constraint-based groups would generate more argumentative conversational elements than those in the unconstrained discussion groups.
3. Do constraint -based discussions differentially affect the nature of argumentation and problem solving while solving well-structured versus ill-structured problems? Although we believe that the role of argumentation is more important in solving ill-structured problems than in solving well-structured ones, the very small amount of research has produced mixed results. Because ill-structured problems do not have a single, preferred solution, problem solvers are more obliged to produce arguments for a preferred solution and counterarguments that rebut alternative ones (Kuhn, 1991; Voss, 1988). That is, problem solvers must defend their chosen solution by providing their beliefs or opinions about the problem and presenting data, warrants, and evidence in support of their decision (Meacham & Emont, 1989). Therefore, we predicted that constraint-based groups solving ill-structured problems would generate more argumentation and problem-solving conversational elements than those in the unconstrained discussion groups.
4. Do the effects of constraint-based scaffolds transfer to individual argumentation performance and to individual problem-solving performance following treatment? That is, does the quality of students' arguments and the quality of their problem solutions differ in a transfer task as a function of the prior use of a constraint-based argumentation scaffold? No one has researched the transfer effects of argumentation on problem solving. We believe that the efficacy of any scaffold is determined by how well students can perform a task when the scaffold has been faded. This question has not been tested in the research. We predicted that the effects of quality of argumentation and problem solving would transfer to individual problem-solving performance.

Method

Participants

Participants were college students registered for an undergraduate introductory economics course at a major university in the eastern United States. A total of 69 volunteers agreed to participate in the beginning of the study. However, missing data from some participants reduced the number in the final analyses to 60 participants. Among the participants, 32 were male and 28 were female. There were 4 freshmen, 28 sophomores, 24 juniors, and 4 seniors. This activity was a scheduled course requirement. Participation was voluntary. Students could opt to complete an alternative assignment, however, none did. Participants were assured of the confidentiality of their responses.

Research Design

A 2 × 2 completely randomized factorial design was used for the study. All participants were randomly assigned to work in groups of three students. These small groups were also randomly assigned into one of four treatment groups according to the two independent variables. These between-subject variables included problem type and type of discussion forum used by the groups.

Participants solved two different types of problems. Thirty participants in groups of three solved well-structured problems, while the other 30 solved ill-structured problems. In order to distinguish between problem types, we used a modification of McGrath's (1984) task classification scheme, called the Task Circumplex, as a theoretical framework for selecting two tasks, a well-structured problem (referred to in the Task Circumplex as an *intellective task*) and an ill-structured problem (referred to in the Task Circumplex as a *decision-making task*). Well-structured problems have demonstrably correct answers. Each problem required students to invent, or select, and compute the correct answer. The criterion for successful group performance is the correct answer. The correct answers for well-structured tasks are based either on cultural norms and beliefs, logic and broadly known factors, or on expert consensus (Laughlin, 1980; McGrath, 1984). An example of a well-structured (intellective) task used in the study is:

> In the country of Vietnam, the velocity of money is constant. Real [gross domestic product] grows by 5% per year, the money stock grows by 14% per year, and the nominal interest rate is 11%. What is the real interest rate?

Ill-structured tasks lack demonstrably correct answers. Each problem required students to select and/or resolve by consensus a preferred alternative. The goal for ill-structured problems is to achieve consensus on a collective decision. An example of an ill-structured (decision-making) task used in the study is:

> Suppose you are advising a small country (such as Bermuda) on whether to print its own money or to use the money of its larger neighbor (such as the United States). What are the costs and benefits of a national money? Does the relative political stability of the two countries have any role in this decision? Explain.

A total of 10 three-person groups solved the well-structured problem during online group discussions while 10 groups solved the ill-structured problem.

All students in all groups corresponded during the problem-solving sessions through a BBS that provided a hierarchically structured, threaded discussion. Each participant contributed comments to a discussion board established for each group for each problem. The threaded discussion showed the list of all the messages with headings, so learners did not have to search through old messages unrelated to the discussion topic. In threaded discussions, student messages are organized around topics and subtopics (Klemm & Snell, 1996) that emerge in the discussion. That is, none of the discussions in the BBS was prestructured. Rather, the structure emerged as students responded to each other's messages.

Half of the three-person groups (5 solving well-structured and 5 solving ill-structured problems) used only the BBS to collaboratively solve their problems, while the other 10 three-person groups (5 solving well-structured and 5 solving ill-structured problems) were provided with a constraint-based tool, Belvedere, to structure their discussions while solving their problems. Belvedere was developed by the Learning Research and Development Center at the University of Pittsburgh to support students in the creation of socially constructed arguments (Learning Research and Development Center, 1996). Belvedere provides a framework for organizing, displaying, and recording the argumentation process, so that the students can easily develop their argument toward solving a problem as they work with group members. Specifically, Belvedere provides the four predefined conversation nodes and three links described earlier (see Figure 1; Suthers, 1998). Constraints in this system are determined by the types of nodes and the links between them. Users are required to link their comments to an already existing comment using one of the four message types. For example, data and principles are meant to modify hypotheses. Students use

these predefined boxes and links to develop their arguments during a problem-solving session (see Figure 1 for an example of a problem being solved during the study). Students used the tool to organize their ideas and the messages that they posted to the BBS.

This argumentation support is structurally congruent with Toulmin's model of argument (Toulmin et al., 1984). *Hypothesis* in Belvedere ("We will have more inflation next year," in Figure 1) corresponds to *claim* in Toulmin's model, because stating a hypothesis is like stating a claim (i.e., making a prediction); *data* ("The deficit is increasing," in Figure 1) to *ground,* because stating measurement, observation, or calculation provides evidence (grounds) for the possible answer; *principle* ("Large deficit causes inflation," in Figure 1) to *warrant,* because it describes why some data (grounds) are for or against the possible answer (claim). *Backing* (implying support for an argument) and *rebuttal* (implying rejection of an argument) can be expressed with the links. If a statement is inconsistent with another statement, learners use the *against*. If a statement is consistent with another statement, they use *for*. Learners can also use the *and* relation to link two or more statements that together support or contradict another statement. The links are used to show how statements are related to each other (Learning Research and Development Center, 1996).

Instruments

Every message from the discussion boards was saved while the students collaboratively solved problems. These messages provided the data used as dependent variables in the first three research questions described earlier. Following the collaborative group problem-solving activities, each participant individually completed a problem-solving activity (an essay describing a solution and justification). These individual essays provided the data used as the dependent variable in the final research question described before. The following assessments were made of the data generated in those two activities.

Assessment of quantity of student arguments during group discussion. The posted messages stored in the BBS database for each of three problems were analyzed for which components of argumentation were present in each posting using a coding scheme adapted from Toulmin's model of argument. The model (Toulmin et al., 1984) identifies five major components of argument, including (a) claims, (b) grounds, (c) warrants, (d) backings, and (e) rebuttals. The total numbers of messages coded in each category for each group comprised one of the dependent variables.

Each message produced by every group was classified by two coders without knowing the identity of the group in order to determine the number (frequency of occurrence) of each component of argumentation according to the coding scheme. Cohen's (1968) kappa (0.70) was calculated to establish inter-rater reliability of the coding. The assessment was determined based on consensus between coders. Decisions about messages with different codes were reconciled in discussions between the coders.

Assessment of the problem-solving process during group discussion. All of the posted messages for each group for all problems stored in the BBS database were reanalyzed for their problem-solving function by classifying each message based on a coding scheme adapted from Poole and Holmes's (1995) Decision Function Coding System (DFCS). The DFCS consists of seven categories, including (a) problem definition (PD), (b) orientation (OT), (c) criteria development (CD), (d) solution development (SD), (e) solution approval (SA), (f) solution critique (SC), and (g) nontask statement (NS). Jonassen and Kwon (2001) used a similar coding scheme to compare problem-solving processes used in computer-mediated communication versus face-to-face communication. They reported greater use of nontask, simple agreement (corresponding to solution approval), and simple disagreement (corresponding to solution critique) categories for both well-structured and ill-structured tasks in the computer-mediated group relative to the face-to-face group.

Each message was coded by two coders without knowing the identity of the group in order to determine the number of messages (frequency of occurrence) in each category during

the problem-solving process. Inter-rater reliability of the coding was fairly high (k = .714). Decisions about messages with different codes were reconciled between the coders.

Assessment of the quality of argumentation. Following the experiment, each participant completed an individual problem-solving assignment. Participants solved new problems that were similar in structure and function to the problems they solved during the treatment period (either well-structured or ill-structured). During this assignment, students wrote an individual essay to describe how they would solve the problem, what their solutions were, and why they used the problem-solving approach that they did. During this assessment, students did not have access to the online learning tools. This assignment was used as a measure of transfer of argumentation and problem-solving skills developed during group problem-solving activities. The individual written reports were scored by two assessors without knowing the identity of the participant, using the scoring rubric in Table 1 in order to determine the quality of argumentation based on Toulmin's model of argument (Toulmin et al., 1984). The individual scores were achieved by summing the number of points achieved in each argumentation category (claims, grounds, warrants, backings, and rebuttal). Coefficient alpha was calculated (0.99) to establish interassessor reliability of the score. Alpha was used because the data were ordinal. The assessment was determined based on consensus between assessors. Decisions about messages with different codes were made in discussions between the assessors.

Assessment of problem-solving performance. Each individual written report was also analyzed for problem-solving performance. The rubric for the well-structured task assessed student analyses of different items of value introduced in the problem for their economic function (store of value, unit of account, or medium of exchange) backed by a warrant (e.g., You can accumulate debt with it; It is accepted in exchange for goods and services; It is a property (fortune) that you can sell to make money and purchase to spend money). The rubric for the ill-structured problems analyzed student agreements or disagreements with the position taken in the problem. The professor, a grader, was looking for student interpretations of the claims, reasons to believe those interpretations, and a prediction of whether inflation is expected or unexpected. The interpretations had to be backed by warrants (e.g., Inflation is a result of the government's refusal to pay its spending; Inflation decreases buying power; Inflation rejects value of money). Each individual written report was scored by the professor and the researcher, without knowing the identity of the participant, using these rubrics (alpha = .964). The assessment was determined based on consensus between assessors. Decisions about messages with different codes were made in discussions between the assessors.

Procedure

During the study, students in small groups worked together to perform the well-structured or ill-structured problem-solving tasks. Group members met online. Although students were not required to use university laboratories to complete these activities, most did. Participants in the threaded discussion groups used only the BBS, while students in the Belvedere group constructed solutions and arguments that they then corresponded to each other using the BBS. Each week for three weeks, the groups solved one economics problem. During the study, each group solved three problems. The week after the treatment period, all students completed the individual problem-solving activity in class.

During class time while the experiment was conducted, the instructor lectured as normal but did not intervene in student group work. The professor answered some questions that students raised, but only if he believed that the questions were not directly related to the performance on the task.

Analysis

To analyze the components of student arguments during group discussion, multivariate analyses of variance (MANOVAs) and univariate analyses of variance (ANOVAs) were

Table 1 ☐ Rubrics for assessing quality of argumentation in student essays

Claims	
Quality	*Criteria*
6	The writer states generalizations that are related to the proposition and which are clear and complete.
4	The writer states generalizations that are related to the propositions, but the assertions are not complete. Enough information is available to figure out the writer's intent, but much is left to the reader to determine.
2	The writer makes generalizations that are related to the proposition, but the assertions lack specificity or offer unclear referents. The writer leaves much for the reader to infer in order to determine the impact of the claim.
0	No claim related to the proposition or unclear assertions.

Grounds	
Quality	*Criteria*
6	The supporting data are compete, accurate, and relevant to the claim.
4	The data offered are relevant but not complete. The writer leaves much for the reader to infer from the data. The writer may have offered the data without the complete citation, which would allow the reader to determine the reliability of the data as evidence. The writer may offer data, which are not complete enough to allow the reader to determine their significance.
2	The data or evidence are weak, inaccurate, or incomplete. E.g. a) an attempt at using a general principle without establishing the truth of the principle; b) the use of examples from personal experience which are not generalizable; c) the citation of data when no source is identified; and d) the use of obviously biased or outdated material.
0	No supporting data are offered or the data are not related to the claim.

Warrants	
Quality	*Criteria*
6	The writer explains the data in such a way that it is clear how they support the claim.
4	The writer explains the data in some way, but the explanation is not linked specifically to the claim.
2	The writer recognizes a need to connect the data to the claim and states some elaboration of data, but the writer fails to make the connection. Or most rules and principles are not valid or relevant.
0	No rules and principles are offered.

Backings	
Quality	*Criteria*
6	The writer states correct, relevant, and specific sources of warrants.
4	The writer states correct, relevant sources of warrants but the sources are very general, not specific.
2	The writer states incorrect, irrelevant sources of warrants.
0	No sources of warrants are given.

Rebuttals	
Quality	*Criteria*
6	The writer states complete and systematic identification of constraints of solutions.
4	The writer identifies constraints of solutions but the constraints are not sufficient.
2	The writer offers few constraints of solutions but the constraints are not elaborated.
0	No recognition of constraints of solutions.

Note. Based on Toulmin's model of argument (Toulmin, Rieke, & Janik, 1984).

conducted. The mean frequency per group per problem of the number of messages classified into each category (claims, grounds, warrants, backings, and rebuttals) was treated as dependent variables. This analysis was conducted to test the first and third research questions.

To analyze the components of the problem-solving process during group discussion, MANOVAs and ANOVAs were calculated. Seven different variables from the coding scheme adapted from Poole and Holmes (1995) discussed earlier were treated as dependent variables: (a) PD, (b) OT, (c) CD, (d) SD, (e) SA, (f) SC, and (g) NS. The mean frequency of each variable was the unit of analysis. This analysis was conducted to test the second and third research questions.

To examine the quality of argumentation, ANOVAs were calculated. The unit of analysis was the score for argumentation quality, assessed using a rubric (scale = 0–30). This analysis was conducted to test the fourth research question.

To investigate the performance in problem solving, ANOVAs were performed. The unit of analysis was the score for individual performance in problem solving, assessed according to the assessment rubric (possible six points per question). The rubrics were specific to the question. For the well-structured problems, the rubrics focused on the store of value, unit of account, and medium of change. For the ill-structured problems, the rubrics focused on reasoned agreement-disagreement with a solution looking for specific economic principles employed, the justification for those principles, and whether inflation is expected or unexpected. This analysis was conducted to test the fourth research question.

RESULTS

Components of Argumentation During Group Problem Solving

The first analysis was of the effects of the Belvedere scaffold on group argumentation. In order to perform this analysis, the contents of the group discussions were classified and the frequency of occurrence of the five components of argumentation based on Toulmin's model of argumentation (Toulmin et al., 1984) were summed. The categories included claims, grounds, warrants, backings, and rebuttals. In order to test the first research question, a 2×2 MANOVA was conducted treating the five argument components as dependent variables. The independent variables were the presence or absence of the argumentation scaffold and the kind of problem-solving task.

Descriptive statistics for student arguments during group discussions are presented in Table 2. The MANOVA showed main effects for the use of the scaffold and the type of problem, but no significant interaction occurred. The MANOVA showed that students using the argumentation scaffold, Belvedere, produced significantly more argument components during group discussions $\Lambda = .769$; $F(5, 52) = 3.132$, $p = .015$, $ES = .231$) than subjects in the unscaffolded discussion groups. Follow-up ANOVAs showed that the groups using the scaffold produced significantly more claims ($F(1) = 4.888$, $MSE = 9.22$, $p = .031$, ES = .08) and grounds ($F(1) = 9.326$, $MSE = 6.43$, $p = .003$, ES=0.14) than groups that did not have access to the scaffold. No other variables differed significantly between groups. Almost no messages stating backings or rebuttals were generated by any of the groups, and very few warrants were provided either. This analysis indicates that the constraint-based scaffold can affect student argumentation during group discussions. Students using the scaffold provided increased claims about how to solve the problem and backed them up with more grounds.

The MANOVA also showed a main effect for type of problem. We wanted to know if the type of problem (well-structured vs. ill-structured) affected the number of student arguments during group discussions. We believed that differences in the processes required to solve ill-structured problems would engage more argumentation than the solution of well-structured problems. The main effect showed that groups solving ill-structured problems did indeed produce more arguments during group discussion $\Lambda = .774$; $F(5, 52) = 3.045$, $p = .017$, $ES = .27$) than students solving well-structured problems. The only significant follow-up

ANOVA was the frequency of stating rebuttals ($F(1) = 9.453$, $MSE = .51$, $p = .003$, $ES = .14$). Groups solving ill-structured tasks produced more rebuttals than those solving well-structured problems. While the difference is statistically significant, it is not substantive, as neither group produced many backings or rebuttals. However, this result provides some evidence that the type of problem affects the nature of student argumentation during group discussion, specifically in terms of their willingness to rebut each other's claims, which is an essential characteristic of solving ill-structured problems.

Components of Problem Solving During Group Problem Solving

In order to determine how the argumentation scaffold affected the components of the problem-solving process during group discussions, the contents of group discussions were classified as one of seven components of the problem-solving process (Poole & Holmes, 1995). Table 3 shows the descriptive statistics for the frequency of components of the problem-solving process used during group discussions. A MANOVA and follow-up ANOVAs were conducted for the seven dependent variables.

Next, we tested the effects of the scaffold on the creation of problem-oriented comments during group discussions. The MANOVA showed a main effect for the constraint-based scaffold ($\Lambda = .495$; $F(7, 50) = 7.285$, $p = .000$, $ES = .51$). The groups using the scaffold produced more problem-oriented comments during their discussion. Specifically, follow-up ANOVAs showed significantly more PD comments ($F(1) = 10.048$, $MSE = 1.49$, $p = .002$, $ES = .15$), OT com-

Table 2 ☐ Descriptive statistics for the components of student arguments during group discussion

	Threaded			*Scaffolded*			*Total*		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Well-structured Problems									
Claims	6.200	2.705	15	6.733	3.105	15	6.467	2.874	30
Grounds	2.200	2.042	15	4.000	3.684	15	3.100	3.067	30
Warrants	1.867	2.066	15	1.000	1.309	15	1.433	1.756	30
Backings	0.200	0.561	15	0.000	0.000	15	0.100	0.403	30
Rebuttals	0.200	0.561	15	0.467	0.640	15	0.333	0.607	30
Ill-structured Problems									
Claims	4.667	2.992	15	7.600	3.312	15	6.133	3.442	30
Grounds	3.067	2.017	15	5.267	1.981	15	4.167	2.260	30
Warrants	2.067	1.710	15	1.467	1.598	15	1.767	1.654	30
Backings	0.333	0.617	15	1.400	3.135	15	0.867	2.285	30
Rebuttals	0.800	0.676	15	1.000	0.926	15	0.900	0.803	30
Total									
Claims	5.433	3.105	30	7.167	3.185	30	6.300	3.148	60
Grounds	2.633	2.042	30	4.633	2.977	30	3.633	2.724	60
Warrants	1.967	1.866	30	1.233	1.455	30	1.600	1.699	60
Backings	0.267	0.583	30	0.700	2.292	30	0.483	1.672	60
Rebuttals	0.500	0.682	30	0.733	0.828	30	0.617	0.761	60

Note. *M* = Mean; *SD* = Standard Deviation; *n* = Number of participants

Table 3 ☐ Descriptive statistics for the components of the problem-solving process during group discussion

	Argumentation Scaffold								
	Threaded			*Scaffolded*			*Total*		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Well-structured Problems									
PD	1.267	1.530	15	1.333	1.04	15	1.30	1.2	30
OT	6.800	2.450	15	9.800	3.29	15	8.30	3.2	30
CD	1.000	1.460	15	2.867	2.09	15	1.93	2.0	30
SD	4.933	2.930	15	4.933	2.96	15	4.93	2.8	30
SA	4.733	2.540	15	4.267	1.62	15	4.50	2.1	30
SC	0.867	1.120	15	1.333	1.29	15	1.10	1.2	30
NS	11.00	5.920	15	4.467	2.77	15	7.73	5.6	30
Ill-structured Problems									
PD	0.800	0.775	15	2.733	1.38	15	1.76	1.4	30
OT	4.200	1.930	15	10.130	3.72	15	7.16	4.1	30
CD	0.600	0.737	15	1.533	1.50	15	1.06	1.2	30
SD	3.200	2.330	15	6.067	1.83	15	4.63	2.5	30
SA	2.667	1.790	15	7.133	3.35	15	4.90	3.4	30
SC	1.133	0.743	15	2.600	1.63	15	1.86	1.4	30
NS	3.200	2.170	15	8.333	3.61	15	5.76	3.9	30
Total									
PD	1.033	1.210	30	2.033	1.40	30	1.53	1.3	60
OT	5.500	2.540	30	9.967	3.45	30	7.73	3.7	60
CD	0.800	1.150	30	2.200	1.91	30	1.50	1.7	60
SD	4.067	2.750	30	5.500	2.48	30	4.78	2.7	60
SA	3.700	2.400	30	5.700	2.97	30	4.70	2.8	60
SC	1.000	0.947	30	1.967	1.58	30	1.48	1.3	60
NS	7.100	5.910	30	6.400	3.72	30	6.75	4.9	60

Note: *M* = Mean; *SD* = Standard Deviation; *n* = Number of participants; PD = Problem Definition; OT = Orientation; CD = Criteria Development; SD = Solution Development; SA = Solution Approval; SC = Solution Critique; NS = Nontask Statement.

ments ($F(1) = 34.702$, $MSE = 8.62$, $p = .000$, $ES = .38$), CD comments ($F(1) = 12.562$, $MSE = 2.34$, $p = .001$, $ES = .18$), SD comments ($F(1) = 4.70$, $MSE = 6.56$, $p = .034$, $ES = .08$), SA comments ($F(1) = 10.153$, $p = .002$, $ES = .15$), and SC comments ($F(1) = 9.085$, $MSE = 5.91$, $p = .004$, $ES = .14$) produced by the groups using the argumentation scaffold. In fact, the scaffolded groups produced more problem solving comments in every category except NS. That is, in every productive category of problem-solving behavior, the argumentation-scaffolded groups produced more problem-solving communications than did the threaded discussion groups. Not only did Belvedere affect the nature of argumentation in groups, it also appeared to have functioned as an affordance for every problem-solving process. The connection between argumentation and problem solving appears quite strong in these results.

Next, we tested the effects of problem type on problem-solving processes manifest in the group discussions. The MANOVA showed a main effect for problem type. That is, the type of

problem significantly affected the nature of the problem-solving process during group discussion ($\Lambda = .694$; $F(7, 50) = 3.153$, $p = .008$, $ES = .31$). Those groups solving ill-structured problems made significantly more OT comments during group discussions than those solving well-structured problems. Follow-up ANOVAs showed that groups solving ill-structured problems made significantly more SC comments ($F(1) = 5.715$, $MSE = 5.91$, $p = .020$, $ES = .09$), while those solving well-structured problems made significantly more CD comments ($F(1) = 4.814$, $MSE = 2.34$, $p = .032$, $ES = .08$). SCs were understandably more important to those resolving the solutions to ill-structured problems, while those solving well-structured problems focused more on the CD for determining the correct answer.

The MANOVA also indicated a significant treatment-by-task interaction for the components of the problem-solving process during group discussion between argumentation scaffold and the type of problem ($\Lambda = .429$, $F(7, 50) = 9.521$, $p = .000$). The frequency of stating PD ($F(1) = 8.753$, $p = .005$, $ES = .14$), SD ($F(1) = 4.700$, $p = .034$, $ES = .08$), SA ($F(1) = 15.444$, $p = .000$, $ES = .22$), and NS all showed significant interactions ($F(1) = 33.649$, $p = .000$, $ES = .38$) in ANOVAs. For the first three variables, the effects of scaffolding were much more pronounced in the groups solving ill-structured problems. For the non-task comments, a disordinal interaction occurred, with the groups without argumentation scaffolding engaging in more non-task interactions while solving well-structured problems and the scaffolded groups engaging in more non-task interactions while solving ill-structured problems. Generally speaking, argumentation scaffolding appears to affect ill- structured–problem solving more than well-structured–problem solving.

Individual Argumentation and Problem Solving on Transfer Task

The fourth research question focused on transfer effects of treatments to individual problem-solving tasks. We wanted to know if students would transfer the argumentation methods that had been scaffolded during the group problem-solving activities to their individual problem-solving activities. In order to answer that question, we created individual well-structured and ill-structured problem-solving activities that students completed following the group activities. We analyzed the student essays both for problem-solving performance and quality of argumentation. Each individual assignment was analyzed according to a rubric (scale: 0–30) to assess the quality of argumentation in the student essays. Descriptive statistics for the quality of argumentation are summarized in Table 4.

A one-way ANOVA showed that students who had used the constraint-based scaffold, Belvedere, created higher quality argumentation in their individual problem-solving exercises than students who had used the threaded discussion board ($F(1, 56) = 4.797$, $MSE = .437$, $p = .033$). This result provides evidence that the argumentation scaffold has a transfer effect on the quality of argumentation during individual problem solving.

Table 4 ☐ Descriptive statistics for the quality of argumentation in individual problem-solving activities

	Threaded			*Scaffolded*			*Total*		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Well-structured problems	8.267	4.71	15	8.933	5.11	15	8.600	4.84	30
Ill-structured problems	14.660	6.03	15	19.600	3.64	15	17.130	5.50	30
Total	11.460	6.23	30	14.260	6.96	30	12.860	6.70	60

Note. *M* = Mean; *SD* = Standard Deviation; *n* = Number of participants

Another one-way ANOVA compared the performance between those solving well-structured and ill-structured tasks. It showed that students faced with ill-structured problems created higher quality arguments in individual problem solving (M = 4.53, SD = 1.57) than students facing well-structured problems (M = 3.87, SD = 1.01, $F(1, 56) = 3.93$, MSE = .030, p = .05).

Finally, we compared students' problem-solving performance on their individual problem-solving tasks using a rubric (scale: 0–6, described earlier). Means are shown in Table 5. A one-way ANOVA showed that students who had used the argumentation scaffold, Belvedere, did not perform significantly better (p =.08) on their individual problem-solving tasks than students who did not have access to the scaffold. This result indicates that there is no conclusive evidence that argumentation scaffold transfers to individual problem-solving performance as determined by the performance criteria we used.

DISCUSSION

Effects of Scaffolds on Argumentation and Problem Solving

The results of this study showed that using a constraint-based argumentation scaffold positively affected the ability of groups to collaboratively construct arguments in an online environment. Although only claims and grounds (evidence) were affected, these are two of the most important components in argument construction and also components of argumentation in which students are weak. Examples of claims produced by the students while solving problems include:

Prices of the goods produced in the year 2010 increased by 52 percent compared to the prices that the goods in the economy sold for in 2000.

According to the traditional view, a debt-financed tax cut will decrease public saving, private saving, and national saving.

The Ricardian view better predicts what will actually happen in the economy than the traditional one when tax cuts.

Examples of grounds produced by the students while solving well-structured problems include:

The implicit price deflator for GDP was calculated, and the value is 1.52.

The tax cut stimulates consumer spending, and government revenues from taxes have decreased.

Even though the government cuts the tax, it does not mean that we have less tax burden. It just reschedules the tax burden to the future.

Surprisingly, the unscaffolded group produced more warrants (e.g., "The GDP can give an estimate of how the economy is doing. Thus, if GDP increases, the state of the economy has gotten better, and if it decreases, the economy has gotten worse") than the scaffolded group, perhaps because the scaffolded groups were focusing on supporting their claims with evidence. Our results were consistent with those

Table 5 ☐ Descriptive statistics for performance in problem solving

	Threaded			*Scaffolded*			*Total*		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Well-structured problems	3.67	.89	15	4.067	1.100	15	3.87	1.08	30
Ill-structured problems	4.13	1.58	15	4.930	1.490	15	4.53	1.59	30
Total	3.90	1.30	30	4.500	1.358	30	4.20	1.35	60

Note. M = Mean; SD = Standard Deviation; n = Number of participants

of Tan (2000), who reported that a CSCA group performed better than a control group in stating grounds. It is also likely that performance in generating warrants was mediated by the instructor in this study, whose lectures focused on the principles that were being studied in the unit. Rebuttals (rejecting others' arguments) are so foreign to college classrooms that students are reluctant to produce them, even in computer-mediated environments.

We also believe that scaffolded students focused more on claims and grounds because those are the most important elements of the solution to the problems. That is, answers to problems are typically stated in the form of claims and justified by grounds, while warrants, backings, and rebuttals play only a supportive role. Premises and evidence are the primary entities in argument construction.

While this study confirmed the effects of the constraint-based argumentation scaffold on the components of argumentation during problem solving, the more important finding related to the effects on components of problem solving. Among seven components of the problem-solving process that were measured, only one component, NS, was not affected by the scaffold. Groups supported by the scaffold produced more PD, OT, CD, SD, SA, and SC. Examples of PD produced by the students while solving problems include:

The first task is a calculation question regarding GDP.

Usually, we calculated to solve them, but this task is a little bit different from the tasks we did the last two weeks.

Examples of OT produced by the students while solving problems include:

What do we begin with? I guess it's GDP.

How about solving step by step? That is, we deal with the traditional view first and then with the Ricardian one.

Examples of CD produced by the students while solving include:

Why don't each of us calculate them and compare each other?

Are these postings hypotheses? I saw you all selected hypotheses. Some are hypotheses and others are not. Check them out.

An example of SD by the students while solving problems is:

Between the years 2000 and 2010, the prices have risen by sixty percent, rising from 1 to 1.6 percentage points. Economists call a price index with a fixed basket of a good Laspeyres index. This can be measured by the [consumer price index] because the CPI does not allow the basket of goods to change.

Examples of SA produced by the students while solving problems include:

I got the same results as you.

It's basically what A's answers were since we all agreed.

An example of SC produced by a student while solving a problem is:

C, I saw what you got for this question but I disagree with part D.

Yoking problem solving outcomes to formal argumentation is a significant finding, especially showing effects of an argumentation scaffold on all forms of problem solving. Jonassen and Kwon (2001) showed effects of computer conferencing only for solution approvals and critiques. However, in this study, the constraint-based argumentation scaffold affected all of the relevant components of problem solving, indicating a more generalized effect of argumentation on problem solving.

Effects of Problem Type on Argumentation and Problem Solving

An extensive research base confirms the differential effects of instruction on different tasks. In this study, the groups solving ill-structured problems produced more arguments than the groups solving well-structured problems both during group discussion and in the quality of argumentation in individual problem solving.

Groups solving ill-structured problems also focused more on problem-solving processes, specifically stating criteria and using those criteria to critique solutions. Examples of CD produced by the students while solving ill-structured problems include:

Important thing is how we explain and justify well.

Before we make a decision, let's provide why we believe so.

Examples of SC produced by the students while solving ill-structured problems include:

I don't think the governor can understand what economists say to him exactly.

I don't completely agree both because everyone will save all of the money; some of it will be spent and some of it will be saved.

The effects of the argumentation scaffold were also more pronounced in stating PD, SD, and SA while solving ill-structured problems. Examples of PD produced by the students while solving ill-structured problems include:

This task seems to require us to make a decision. As I remember, no absolute answer exists.

There are two important conditions we have to take a close look at. Check them out.

An example of an SD produced by a student while solving an ill-structured problem is:

With an increase in taxes equal to an increase in government purchase, the interest rate and investment will remain the same. The explain is that an increase in taxes will shift the curve to the right while an increase in government purchases by the same amount to the left bringing it back to the equilibrium interest rate and investment. One would expect the explanation to depend on the marginal propensity to consume, and it should, but in this case, when the increase in taxes and government purchases are the same, they cancel each other out.

Examples of SAs produced by the students while solving ill-structured problems include:

I will accept B's answer. It also looks OK to me.

I definitely agree with B. Don't you think now is the time when you make a decision?

We conclude that ill-structured problems are more affected by argumentation than well-structured problems. This result is consistent with many studies contending that when solving an ill-structured problem, students are required to choose a preferred solution and reject alternative ones (Voss, 1988); develop their own argument (Jonassen, 1997); and defend their own solution (Meacham & Emont, 1989). Because no single, correct convergent solution exists in an ill-structured problem (Kitchener, 1983), a student in a group problem-solving situation must argue to persuade group members that his or her claims are more reasonable and valid than others. Ill-structured problems provide students with more opportunities to make arguments than well-structured problems.

Not surprisingly, the effects of problem type were more pronounced on problem solving than they were on argumentation. Although solving ill-structured problems produced more arguments during group discussions than solving well-structured problems, the only significant univariate effect was for stating rebuttals. The use of rebuttals is an requirement for solving ill-structured problems. The purpose of argumentation in ill-structured problems is to persuade or convince people to accept one's claims (Fisher and Sayles, 1966) and to reject other solution claims. Successful argumentation thus depends on the degree of how well one's argument makes others' opinions change. While solving ill-structured problems, multiple solutions and solution paths may exist (Kitchener, 1983), and there is no universal agreement on the appropriate solution (Voss, 1988), so personal opinions or beliefs about the problem are required in the process of interpretation (Meacham & Emont, 1989). It is easy to understand why solving ill-structured problems requires students to use rebuttals more than solving well-structured problems. The use of rebuttals in solving well-structured problems is not as common because there is a single best solution in well-structured problems (Simon, 1978), so when students in a group problem-solving situation reach an agreement about how

to solve the problem, no more counterarguments are needed.

Transfer Effects to Individual Problem Solving

We also assessed any potential transfer of the processes used during group problem solving to individual problem solving. Students solved problems individually following their group experiences. While the scaffold clearly affected quality of argumentation during individual problem solving, it did not significantly affect the individual problem-solving performance. The transfer of argumentation skills from group to individual performances was understandably affected by the scaffold. However, because problem-solving performance was not directly scaffolded, it was less likely to appear in individual performances. The rubrics used to assess problem-solving performance were more consistent with the argumentation model than they were with the Poole and Holmes (1995) problem-solving model. Finally, it is very likely that individual problem solving was mediated by individual differences. Without the support of peers during group activities, more of the variance in performance was accounted for by a variety of individual differences.

CONCLUSION

The situated learning movement in education presupposes the importance of solving more authentic problems during learning. Authentic problems tend to be ill structured, as opposed to the well-structured problems that are encountered in most K–16 education. Because ill-structured problems are dialectical in nature, with two or more opposing conceptualizations of the problem or its solution result, the production of arguments to support those differing conceptualizations is essential. Supporting those activities in online environments has proven problematic. If it is important that students learn to articulate the differing assumptions in support of arguments for whatever solution they recommend, then online tools should be made available in online environments to support argumentation. This study has provided evidence that argumentation can be supported effectively by online argumentation scaffolds and that the production of better arguments directly affects the problem-solving activities that students use. Belvedere is but one tool for supporting argumentation. Because the nature of argumentation varies with the type of problem or task, it is probable that the nature of the argumentation support tool should be tailored to the task. The online environments developed to date (e.g., Belvedere) have used the same kind of argument structure. Jonassen and Remidez (2002) have described the development of a constraint-based discussion environment onto which different rhetorical structures can be mapped to support different kinds of conversations (e.g., argumentation, hermeneutic, literary analysis). This type of environment can be tailored to the nature of the task that is being supported by conversations. To the degree that educators expect learners to engage in authentic learning in online environments, they must provide learners with more sensitive and sophisticated tools for supporting the different kinds of cognitive tasks implied by those problems. □

Kyoo-Lak Cho is with the Korean Research Institute for Vocational Education & Training (KRIVET) in Seoul, South Korea.

David H. Jonassen is Distinguished Professor of Learning Technologies at the University of Missouri.

REFERENCES

Bell, P., & Linn, M.C. (2000). Scientific arguments as artifacts: Designing for learning from the Web with KIE. *International Journal of Science Education, 22*(8), 797–817.

Buckingham Shum, S.J., MacLean, A., Bellotti, V.M.E., & Hammond, N.V. (1997). Graphical argumentation and design cognition. *Human-Computer Interaction, 12*, 267–300.

Carr, C.S. (1999). *The effect of computer-supported collaborative argumentation (CSCA) on argumentation skills in second-year law students.* Unpublished doctoral dissertation. The Pennsylvania State University, Pennsylvania.

Cerbin, B. (1988). *The nature and development of informal*

reasoning skills in college students. (ERIC Document Reproduction Service No. ED 298 805)

Chi, M.T.H., Slotta, J.D., & de Leeuw, N.A. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27–43.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provisions for scales disagreement of partial credit. *Psychological Bulletin, 70*, 213–220.

Darses, F. (1991). The constraint satisfaction approach to design: A psychological investigation. *Acta Psychologica, 78*, 307–325.

Fisher, W.R., & Sayles, E.M. (1966). The nature and functions of argument. In G.R. Miller & T.R. Nilsen (Eds.), *Perspectives on argumentation* (pp. 2–22). Chicago, IL: Scott, Foresman and Company.

Hong, N.S., Jonassen, D.H., & McGee, S. (in press). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching.*

Jonassen, D.H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65–94.

Jonassen, D.H. (1999). Designing constructivist learning environments. In C.M. Reigeluth, (Ed.). *Instructional design theories and models: A new paradigm of instructional technology,* Vol 2. Mahwah, NJ: Lawrence Erlbaum Associates.

Jonassen, D.H., & Kwon, H.I. (2001). Communication patterns in computer-mediated vs. face-to-face group problem solving. *Educational Technology Research and Development, 49*(10), 35–52.

Jonassen, D.H., & Remidez, H. (2002, January). *Mapping alternative discourse structures onto computer conferences.* Paper presented at the annual meeting of Computer Support for Collaborative Learning, Boulder, CO.

Kitchner, K.S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 26*, 222–232.

Klemm, W.R., & Snell, J.R. (1996). Enriching computer-mediated group learning by coupling constructivism with collaborative learning. *Journal of Instructional Science and Technology, 1*(2). [Online]. Available: http://www.usq.edu.au/electpub/e-jist/vol1no2/article1.htm.

Knudson, R.E. (1991). Effects of instructional strategies, grade and sex on students' persuasive writing. *Journal of Experimental Education, 59*(2), 141–152.

Kuhn, D. (1991). *The skills of argument.* Cambridge, UK: Cambridge University Press.

Lajoie, S.P. (1993). Computer environments as cognitive tools for enhancing learning. In S.P. Lajoie & S.J. Derry (Eds.), *Computers as cognitive tools* (pp. 261–288). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lajoie, S.P., & Lesgold, A. (1992). Apprenticeship training in the workplace: A computer coached practice environment as a new form of apprenticeship. In M. Farr & J. Psotka (Eds.), *Intelligent instruction by computer: Theory and practice* (pp. 15–36). New York, NY: Taylor & Francis.

Learning Research and Development Center. (1996). *Advanced cognitive tools for learning.* [Online]. Available: http://advlearn/lrdc.pitt.edu/belvedere/index.html.

Leeman, R.W. (1987). *Taking perspectives: Teaching critical thinking in the argumentation course.* (ERIC Document Reproduction Service No. ED 292 147)

McCann, T.M. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English, 23*(1), 62–76.

McGrath, J.E. (1984). *Groups: Interaction and performance.* Englewood Cliffs, NJ: Prentice-Hall.

Meacham, J.A., & Emont, N.C. (1989). The interpersonal basis of everyday problem solving. In J.D. Sinnott (Ed.), *Everyday problem solving: Theory and application* (pp. 7–23). New York, NY: Praeger.

Poole, M.S., & Holmes, M.E. (1995). Decision development in computer-assisted group decision making. *Human Communication Research, 22*(1), 90–127.

Richard, J.F., Poitrenaud, S., & Tijus, C. (1993). Problem-solving restructuration: Elimination of implicit constraints. *Cognitive Science, 17*, 497–529.

Sanders, J.A., Wiseman, R.L., & Gass, R.H. (1994). Does teaching argumentation facilitate critical thinking? *Communication Reports, 7*(1), 27–35.

Saunders, K.M. (1994). Law as rhetoric, rhetoric as argument. *Journal of Legal Education, 44*(4), 566–578.

Simon, D.P. (1978). Information processing theory of human problem solving. In D. Estes (Ed.), *Handbook of learning and cognitive processes.* Hillsdale, NJ: Lawrence Erlbaum Associates..

Suthers, D. (1998). *Representations for scaffolding collaborative inquiry on ill-structured problems.* Paper presented at the 1998 AERA Annual Meeting, San Diego, California.

Suthers, D.D., & Hundhausen, C.D. (2001). Learning by constructing collaborative representations: An empirical comparison of three alternatives. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning* (pp. 577–584). Universiteit Masstricht, Maastricht, Netherlands.

Tan, S.C. (2000). *Supporting collaborative problem solving through computer-supported collaborative argumentation.* Unpublished doctoral dissertation. The Pennsylvania State University, Pennsylvania.

Toulmin, S. (1958). *The uses of argument.* Cambridge, England: Cambridge University Press.

Toulmin, S.E., Rieke, R.D., & Janik, A. (1984). *An introduction to reasoning* (2nd Ed.). New York, NY: Macmillan.

Veerman, A.L., Andriessen, J.E.B., & Kanselaar, G. (1999, December). *Collaborative learning through computer-mediated argumentation.* Paper presented at the conference on Computer Supported Collaborative Learning (CSCL 99), San Francisco, California.

Voss, J.F. (1988). Problem solving and reasoning in ill-

structured domains. In C. Antaki (Ed.), *Analyzing everyday explanation: A casebook of methods* (pp. 74–93). London, UK: SAGE Publications.

Voss, J.F., Lawrence, J.A., & Engle, R.A. (1991). From representation to decision: An analysis of problem solving in international relations. In R.J. Sternberg & P.A. French (Eds.), *Complex problem solving* (pp. 119–157). Hillsdale, NJ: Lawrence Erlbaum Associates.

Yeh, S.S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English, 33*(1), 49–83.