The Effect of Talk and Writing on Learning Science: An Exploratory Study

LÉONARD P. RIVARD
Collège universitaire de Saint-Boniface, 200, avenue de la Cathédrale, Saint-Boniface, MB, R2H 0H7, Canada

STANLEY B. STRAW
University of Manitoba, Winnipeg, MB, Canada

Received 30 March 1998; revised 26 May 1999; accepted 19 June 1999

ABSTRACT: This study investigated the role of talk and writing on learning science. The purpose was to explore the effect of talk, writing, and talk and writing on the learning and retention of simple and integrated knowledge, and to describe the mechanisms by which talk and writing mediate these processes. Forty-three students were randomly assigned to four groups, all stratified for gender and ability. At intervals during an instructional unit, three treatment groups received problem tasks that involved constructing scientific explanations for real-world applications of ecological concepts. A control group received simpler descriptive tasks based on similar content. Students in the talk-only treatment group (T) discussed the problem tasks in small peer groups. Students in the writing-only treatment group (W) individually wrote responses for each of the tasks, but without first talking to other students. Students in the combined talk and writing treatment group (TW) discussed the problems in groups prior to individually writing their explanations. Dependent variables included simple, integrated, and total knowledge scores based on multiple-choice tests, essay questions, and concept maps obtained at three timepoints during the study: a pretest; an immediate posttest; and a delayed posttest. Records of student talk and writing were also analyzed to describe the mechanisms involved. The findings suggest that talk is important for sharing, clarifying, and distributing knowledge among peers, while asking questions, hypothesizing, explaining, and formulating ideas together are all important mechanisms during peer discussions. Analytical writing is an important tool for transforming rudimentary ideas into knowledge that is more coherent and structured. Furthermore, talk combined with writing appears to enhance the retention of science learning over time. Moreover, gender and ability may be important mediating variables that determine the effectiveness of talk and writing for enhancing learning. © 2000 John Wiley & Sons, Inc. *Sci Ed* **84:**566–593, 2000.

Correspondence to: L. Rivard; e-mail: lrivard@ustboniface.mb.ca
Contract grant sponsor: Social Sciences and Humanities Research Council of Canada

INTRODUCTION

According to the National Science Education Standards, the goal of science literacy for all will necessitate the creation of active learning environments in which students construct knowledge and achieve understanding of fundamental scientific ideas while engaging in discourse within the classroom learning community (National Research Council, 1996). However, studies over the last decade have suggested that the actual situation may be quite different from this vision of science education for the twenty-first century. The National Assessment of Educational Progress (NAEP) reported that science instruction in the United States is predominantly accomplished by teacher lecturing (Mullis & Jenkins, 1988), which is drastically different from the science teaching envisioned in the National Science Education Standards. Anderson and Roth (1989) stated that the typical science class is taught "without concern for integrating that knowledge with students' personal knowledge and without the rich conceptual coherence needed to make the knowledge useful in explaining real-world phenomena" (p. 273). Brophy and Alleman (1991) suggested that classroom activities, by and large, embrace low-level routine learning tasks and do not place "much emphasis on developing understanding of content or applying it in meaningful ways" (p. 10). Furthermore, evidence from various assessments suggests that students experience considerable difficulty applying science concepts in their explanations of everyday events and phenomena (Cole, 1990; Lapointe, Mead, & Phillips, 1989; Science Council of Canada, 1984). These observations suggest that science classrooms are, for the most part, passive learning environments in which students generally have few opportunities for integrating their knowledge into an elaborate, coherent conceptual system (Newmann, 1988; Reif & Larkin, 1991).

Science educators have argued that science classrooms ought to be active learning environments in which students construct personal meanings within the classroom community (Erickson & MacKinnon, 1991; Roth, 1990). Many research studies have documented science classrooms in which students are active learners, interacting with peers on cognitively engaging problem-solving tasks within a collaborative learning environment (Alexopoulou & Driver, 1996; Audet, Hickman, & Dobrynina, 1996; Keys, 1996; Meyer & Woodruff, 1997; Palincsar, Anderson, & David, 1993; Roth, 1994a, 1994b). The objective of the present process–product study was to determine the effect on science learning of using problem-solving tasks in which students used language, either individually or collaboratively with peers, to explain real-world applications of ecological concepts. Moreover, qualitative data obtained during the study provided rich descriptions of the discourse processes that occurred during peer interactions and while writing.

Theoretical Perspective

One of the assumptions underlying the National Science Education Standards is that "student understanding is actively constructed through individual and social processes" (National Research Council, 1996, p. 28). Constructivism posits that personal knowledge and understanding result from the myriad connections that learners make while integrating new information with prior knowledge. Constructivist ideas about learning have been embraced by scholars in both literacy (Bruner, 1986; Greene & Ackerman, 1995) and science education (Appleton, 1997; Driver, Asoko, Leach, Mortimer, & Scott, 1994; Tobin, 1993). Some constructivist approaches have emphasized the personal construction of knowledge in which the individual's idiosyncratic experiences within the learning environment are paramount, whereas others have underlined the importance of social processes in mediating cognition (Nystrand, 1990; O'Loughlin, 1992; Piaget, 1950; Prawat & Floden, 1994; Vygotsky, 1978). Fosnot (1993) argued that science education would benefit from a synthesis of these two perspectives. We believe that language-based learning strategies can be effectively orchestrated in the science classroom for achieving this synthesis. Classroom activities that feature listening, talking, reading, and writing can all be used to enhance the cognitive processing of information.

The role of talk in learning has been described by scholars both within and beyond the science education community (Barnes, 1976; Britton, 1982; Bruner, 1986; Corson, 1988; Lemke, 1990; Pea, 1993; Prawat, 1993; Schoenfeld, 1989). For instance, Britton (1982) theorized about how understanding might be enhanced through talking:

> We come to an understanding in the course of communicating it. That is to say, we set out by offering an understanding and that understanding takes shape as we work on it to share it. And finally we may arrive co-operatively at a joint understanding as we talk or in some other way interact with someone else. (p. 115)

Prawat (1993) has argued that there is "a dialectical relationship between individual knowledge, arrived at by reflecting on one's own activity, and knowledge that is socially mediated or jointly agreed on" (p. 11). This mediating role of oral language is consistent with instructional models that have embraced the notion that knowledge is coconstructed or socially constructed within a sociocultural context (Brown, Collins, & Duguid, 1989; Hatano & Inagaki, 1991; Resnick, Levine, & Teasley, 1991; Rogoff, 1990; Rogoff & Lave, 1984; Tharp & Gallimore, 1988). Bruffee (1993) put it succinctly: "knowledge is socially constructed and learning socially interdependent" (p. 155).

The use of writing as a learning strategy has received considerable theoretical support from scholars in a variety of disciplines (Applebee, 1984; Barnes, 1976; Britton, 1989; Howard, 1988; Langer, 1986; Resnick, 1987; Scardamalia & Bereiter, 1986; Schumacher & Nash, 1991; Vygotsky, 1962). Studies of writing-to-learn have generally been interpreted using a "depth-of-processing hypothesis," which suggests that different writing tasks prompt learners to invoke different cognitive strategies for processing and encoding information (Langer & Applebee, 1987, p. 92). For instance, tasks such as listing, defining, or describing require that learners simply focus on one or more concepts in isolation, usually one at a time, whereas analytical tasks such as explaining real-world applications of scientific concepts demand that learners connect these into an integrated web of meaning. McGinley and Tierney (1989) argued that writing tasks can provide the "means or 'routes' for 'traversing a topical landscape' . . ." (p. 249), and that "knowledge is best acquired by traversing [a topic of study] from a variety of perspectives" (p. 250).

Research studies support the use of both talk and writing as teaching and learning strategies (Durst & Newell, 1989; Gaskins et al., 1994; Hayes, 1987; Langer & Applebee, 1987; Newell, 1984, 1986; Rivard, 1994). However, talk and writing, used separately, may not be quite as helpful for conceptualizing relationships as a strategy that combines them in order to obtain the benefits of both modalities (Olson, 1994; Thaiss, 1988). Tishman and Perkins (1997) suggested that "written language, stabilized on paper, invites kinds of reflection not so natural to oral exchanges" (p. 371). In comparing talk and writing, Goody (1994) observed that writers demonstrate more abstract thought, are more objective and explicit, elaborate in greater detail, and tend to be more rigorous in their treatment of the topic. Although writing may be a powerful tool for structuring knowledge, talk is still important for generating, clarifying, sharing, and distributing ideas. Halliday and Martin (1993) suggested that:

> Written language is corpuscular and gains power by its density, whereas spoken language is wave-like and gains power by its intricacy. . . . Writing puts language in chains; it freezes it, so that it becomes a thing to be reflected on . . . Writing deprives language of the power to intuit, to make indefinitely many connections in different directions at once, to explore (by tolerating them) contradictions, to represent experience as fluid and indeterminate. . . . But . . . in destroying this potential it creates another one: that of structuring, categorizing, disciplinizing. It creates a new kind of knowledge: scientific knowledge. (p. 118)

Talk and writing are therefore complimentary modalities. The use of writing as an instrument for learning underlines the personal construction of knowledge, whereas the use of talk for learning is consistent with social constructivist thought. An instructional strategy encompassing both should enhance learning more than another using either of these two language modalities alone.

Asking students to explain scientific phenomenon, either orally or in writing, should enhance their content understanding (Fellows, 1994; Martin, 1970; Perkins & Blythe, 1994; Prain, 1995; Schumacher & Nash, 1991). Brown and Campione (1990) argued that "the burden of explanation is often the push needed to make students evaluate, integrate and elaborate knowledge in new ways" (p. 114). In studies of writing to learn, Newell (1986) observed that explaining enhanced content learning more than notemaking or answering study questions.

Although philosophical views about the nature of explanation vary within the science community, three different conceptions have generated much debate (Thagard, 1989). The inferential conception considers explanation to be a logical argument consisting of a series of statements or propositions (Hempel, 1965). The erotetic conception sees explanation as an answer to a why-question (Braithwaite, 1953; Bromberger, 1966; van Fraassen, 1987; Weaver, 1964), whereas the causal conception views the explanation of a phenomenon as "laying bare its inner workings, its underlying causal mechanisms" (Kourany, 1987, p. 24; Salmon, 1987). Moreover, Horwood (1988) decried the way in which explanation has been interpreted in the science classroom. For Horwood:

> Description is purely information and the bits of information are isolated from any network or relatedness. An explanation is given when connections are drawn between and among pieces of information. (p. 41)

Questions that ask students to explain or make explicit the subtle relationships between and among concepts, as well as to make connections between classroom events and daily life, should enhance learning (Raphael & Gavelek, 1984; Sawyer, 1991; Shepardson & Pizzini, 1991).

Some investigators have recommended the use of particular learning and teaching strategies for making the science classroom more gender-inclusive or gender-free (Adams, 1992; Baker & Leary, 1995; Davis & Steiger, 1994; Rosser, 1990; Scantlebury & Kahle, 1993; Smail, 1987). For instance, Adams (1992) recommends the use of learning strategies such as sharing ideas, classroom dialogue, peer discussions, concrete experiences, and journal activities. Baker and Leary (1995) reported that females "expressed strong feelings for more interaction with their peers in their repeated requests for group work, partners, and more discussion" (p. 9). More recently, Burkam, Lee, and Smerdon (1997) suggested that greater use in the science classroom of "problems with practical implications and opportunities for creative solutions, and active open-ended learning situations" ought to enhance learning by girls.

Using both peer discussion and analytical writing together should enhance learning more than using peer discussion or analytical writing alone. However, would the effect of using talk and writing, alone and combined, be different for the learning of simple and integrated knowledge? Would the effect be different for short-term learning and its retention over time? Would these effects be different for girls and boys? The present product–process study was designed to investigate these questions.

PURPOSE

The purpose of this exploratory study was to investigate the role of talk and writing on learning science. We sought to assess the effect of talk and writing, alone and combined, on the learning of simple and integrated knowledge, as measured immediately after completing an instructional unit (initial learning) and again after a delay of 6 wk (retention). We expected that students who used both talk and writing would learn more than students using just talk or just writing. We also expected that the pattern of differences among the various treatment groups would vary according to time of test, the kind of knowledge (simple, integrated, and total), and gender. Furthermore, we expected that the patterns of use and the role of discourse would differentiate between talk and writing. Three hypotheses framed the quantitative part of the study:

1. There will be no significant difference in mean aggregate simple knowledge scores that can be attributed to treatment, gender, and interaction of treatment and gender.
2. There will be no significant difference in mean aggregate integrated knowledge scores that can be attributed to treatment, gender, and interaction of treatment and gender.
3. There will be no significant difference in mean aggregate total knowledge scores that can be attributed to treatment, gender, and interaction of treatment and gender.

The naturalistic component of the study was framed by two questions. The first focused on the role of peer discussion on learning: How did talk within the peer groups influence students' knowledge about the problems being discussed and what mechanisms were operating within the collaborative environment? A second question addressed how writing was used by individual students to consolidate learning: How did students' written responses compare with the discussion of the problem that had preceded the writing task?

METHOD

A quasi-experimental treatment–control design with an embedded interpretative approach was adopted for the study. Students were randomly assigned to the control and three treatment groups. In the naturalistic component, two peer groups from the talk-and-writing treatment groups were selected for in-depth observations that would be used to enrich the discussion of the quantitative component of the study.

Subjects

The subjects for the present study included 43 eighth grade students (27 boys and 16 girls) in two intact classrooms in a French-language school, grades 7–12, situated in a Canadian prairie province. The students were all taught by a teacher who had volunteered to participate in the study. Most of the students were from families with similar cultural and linguistic backgrounds: They were Franco-Manitoban and francophone (group which uses French as a first language). The socioeconomic level of these students ranged from lower middle-class working families to upper middle-class professional families. The school has a good academic reputation with over 70% of students continuing their education at the postsecondary level.

Design

The independent variables of the study included gender and treatment (talk-only group [T] writing-only group [W], talk-and-writing group [TW], and control group [C]), whereas the dependent variables were: (a) knowledge scores (simple, integrated, and total scores) and (b) time (immediate and delayed posttests).

A mixed factorial design ($4 \times 2 \times 2$), with a pretest and two posttests (immediate and delayed), was employed to test the effects of gender and treatment (T, W, TW, and C) on dependent measures which were based on data from multiple-choice tests, essay questions, and concept maps. Simple, integrated, and total knowledge scores were established for all of these instruments. Aggregate scores were calculated for each kind of knowledge (simple, integrated, and total) on the basis of the data obtained from each of these three instruments. The aggregate scores for simple, integrated, and total knowledge were then statistically analyzed.

INSTRUMENTATION

The instruments which were used to measure student learning included a multiple-choice test, a test with short essay questions, and concept maps. Two kinds of content knowledge were assessed in this study: simple or isolated knowledge; and integrated or relational knowledge (Linn & Songer, 1993). Simple knowledge includes knowledge of facts, terminology, and concepts, whereas integrated knowledge focuses on the relationships among these concepts and includes applications and explanations (Alexander, Schallert, & Hare, 1991; Prawat, 1989). The approach taken in this study was to tease out the simple knowledge composed of concepts or building blocks, from the integrated knowledge composed of relations or *threads* that link up these concepts which together make up the knowledge network of students in a particular content domain (Reif & Larkin, 1991). Procedures for determining simple and integrated knowledge scores were developed for each of the three instruments.

Multiple-Choice Tests

Both versions of the multiple-choice test (A and B) included 30 items, with half of the items measuring simple knowledge and the other half integrated knowledge. Most of the questions were adapted from those used in various international, national, and provincial science assessments that are now in the public domain.

The multiple-choice tests were administered 3 weeks before instruction began (pretest), immediately after completing the unit (immediate posttest), and again 6 weeks later (delayed posttest). Class A wrote version A in the pretest, version B in the immediate posttest, and version A again in the delayed posttest. In contrast, class B wrote version B in the pretest, and versions A and B in that order after the intervention. Content validity of the test items was established by a group of experts in science education. The Pearson correlation coefficient was 0.67 for simple knowledge, 0.65 for integrated knowledge, and

0.80 for total knowledge. McMillan and Schumacher (1989) indicated that a coefficient of > 0.70 is acceptable for research purposes. These correlations were deemed acceptable given the exploratory nature of the study.

Essay Questions

Four short essay questions, which required students to provide written explanations for natural phenomena or real-world applications relevant to the unit of study on ecology, were also administered three times during the experiment. Two of these questions had been used in previous assessments in the province of Manitoba, and two of them were constructed by the researcher. Student responses to all of these questions were typed and a code number was assigned to each one to conceal both the identity of students and the time of test (pretest, immediate posttest, or delayed posttest) to reduce the possibility of scorer bias. Two scores were established for each of these essay questions: (i) a score for simple knowledge that was based on the number of target concepts included in the response and (ii) a score for integrated knowledge that was based on a holistic evaluation of the student's response (Cooper, 1977). The target concepts that students should have included in their written response were established using an expert panel and varied from one question to the next. Students received one point for each target concept included in their written response. The total possible simple knowledge score for the four essay questions was 23 points.

The score for integrated knowledge was established using a holistic marking scale. Criteria included the number of conceptual relationships in the written response, the clarity and organization of the text, and the adequateness of the explanation (Davis, Scriven, & Thomas, 1987; Gorman, Purves, & Degenhart, 1988; Purves, 1984). Each of the four questions was evaluated using a four-point marking scheme (0–4), with the total possible score being 16 points.

The validity of the essay questions was established using an expert panel composed of three biology professors and one science curriculum consultant with expertise in environmental education. Student responses to these questions were scored by three different scorers, the researcher and two research assistants who had been trained for the scoring task. Students received the average of the three scores for each question. Interrater reliability was established for the scoring of each question using Pearson correlation coefficients. For all four questions, the average correlation coefficients ranged from 0.85 to 0.92 for scoring simple knowledge and from 0.79 to 0.85 for scoring integrated knowledge. According to Cooper (1977), "a reliability coefficient of 0.80 is considered high enough for program evaluation" (p. 18). As the objective of this study was not to compare individual students, but rather experimental treatments, the reliabilities were considered adequate for making comparisons between the different groups.

Concept Maps

Wandersee (1990) described concept maps as "maps of cognition" (p. 923). Novak and Gowin (1984) defined them as "a schematic device for representing a set of concept meanings embedded in a framework of propositions" (p. 15). Concept mapping is a valid tool for representing conceptual knowledge and a reliable instrument for research purposes (Edwards & Fraser, 1983; Novak & Musonda, 1991; Wallace & Mintzes, 1990). Criteria like the number of concepts and the number of propositions appear to be both valid and reliable as assessments of conceptual understanding (Cronin, Dekkers, & Dunn, 1982; Mason, 1992; Novak & Musonda, 1991; Stuart, 1985).

The teacher taught the students how to construct concept maps several weeks prior to beginning the unit of study. Students also completed concept mapping exercises during this period to ensure that the technique itself did not confound the results. Concept maps were obtained from students at all three testing sessions: pretest, immediate posttest, and delayed posttest.

A concept seeding technique was used to guide student mapping. Concept seeding involves giving students one or more concepts to initiate the mapping process (J. H. Wandersee, personal communication, March 28, 1994). The expert panel had established 13 key concepts for the unit on ecology: ecology, ecosystem, abiotic factors, biotic factors, biosphere, biome, food chain, food web, community, population, environment, habitat, and niche. Students were given five of these as seed concepts for the mapping exercise. The seed concepts were habitat, community, food web, ecosystem, and abiotic factor. Maps were scored for the number of relevant concepts, a measure of simple knowledge, and the number of relevant propositions, a measure of integrated knowledge, that had been included in the map. Relevant concepts and propositions were established by the expert panel.

Students received two points for each of the key concepts which had not been given as seed concepts that were included in the map. Students also received one point for the seed concepts and for each additional correct concept included in the map. The number of key concepts, seed concepts, and other correct concepts thus established the simple knowledge score for the students. Students received points for each scientifically correct proposition that was included in the map. A proposition was defined as a line or arrow linking two concepts. The expert panel identified eight critical propositions for the ecology unit and students received two points for these propositions. Students also received one point for all other scientifically correct propositions. The number of propositions, critical or otherwise, defined the integrated knowledge score for the students.

The concept maps were scored by the researcher and two research assistants who had received prior training. The mean interrater reliability (Spearman correlation coefficient) calculated separately for both simple and integrated knowledge components for all possible pairs of scorers was 0.98 and was therefore considered adequate for the study.

Qualitative Data

Two peer-discussion groups, one from each class, each composed of four students, were videotaped during the five problem-solving sessions scheduled throughout the course of the study. The same groups of students were videotaped each time so that the recording process might become less obtrusive over time. These students discussed the problem tasks in small peer groups prior to individually writing their response to each question. The videotapes were transcribed for later analysis and then destroyed. These transcriptions, which constitute case studies of peer-discussion groups in action, were analyzed to determine how student understanding evolved during these group sessions. Written responses to each of the explanatory tasks, which were completed immediately after discussion with peers, were also collected from the individual students. Pseudonyms have been used for all statements and responses reported in the study. The oral statements and written responses of the students were compared to determine how both talk and writing influenced collective and individual thinking about the problem tasks (Miles & Huberman, 1984). We hoped that interpreting these qualitative data would allow us to clarify some of the mechanisms underlying the role of talking with peers for understanding science content (Peshkin, 1993). The analysis initially focused on the transcripts while looking for patterns of interaction within the peer groups, and later extended to searching for connections between these two forms of discourse, talk during a structured discussion while working on an explanatory task with students' explanations written soon afterwards.

CONTEXT FOR THE STUDY

An ecology unit in the eighth grade science program established the context for the study. The study unit included individual, small-group, and whole-class activities that involved teacher-led and peer group discussions, library research, lab activities, worksheets, simulations, audiovisual presentations, and varied reading and writing tasks. The objective in all of these activities was to give students the requisite conceptual building blocks for constructing a coherent knowledge system about ecology. Once the conceptual groundwork was laid for exploring a particular topic, such as ecological populations and communities, problem-solving sessions were organized in which students were now required to apply this basic knowledge to new situations that they had not encountered during previous instruction. Problem solving can be defined as "thinking [that] is functional, active, and grounded in goal directed action" (Rogoff, 1990, p. 8). Problem solving can involve a variety of different tasks, including writing an explanation for everyday natural phenomena or simply exploring new ideas that apply basic science concepts.

TREATMENT

The students were randomly assigned to four treatment groups, all stratified for gender and ability. The study included a talk-only group, a writing-only group, a talk-and-writing group, and a control group. Students were only separated into these four groups for the problem-solving sessions. Each of the five problem-solving sessions lasted about 50 minutes. With the exception of the control group, the tasks that were assigned in the sessions involved explaining real-world applications of the ecological concepts studied earlier.

Each of the five problem-solving sessions focused on different key concepts in ecology: biomes; adaptation; ecosystems, populations and communities; niche and habitat; and food chains and food webs. For instance, one of the explanatory tasks in the session on ecosystems, populations, and communities asked students to explain whether or not a dead log could be considered an ecosystem. The explanatory tasks required students to make connections among the various ecology concepts being studied.

Students in the talk-only group discussed the assigned problems with peers. The students in the writing-only group individually responded in writing to these same problems, but without talking with peers. Students in the talk-and-writing group discussed the problems with peers before individually responding in writing. Students in all three groups worked on explanatory tasks: two of the groups benefitted from talk (TW and T groups), whereas two groups benefitted from writing (TW and W groups). In comparison, students in the control group individually completed descriptive tasks based on the same ecological concepts. These tasks included fill-in-the-blanks, true-or-false exercises, matching exercises, definitions, and descriptions. When students wrote during the problem sessions, they were instructed to respond to as many of the problem tasks, descriptive or explanatory, as possible while working alone at their desk. Apart from the problem-solving sessions, the students in each of the two classes were together throughout the 6-wk instructional unit, which involved about 16 hr of class time; that is, they read the same materials, did the same assignments, and received essentially the same instruction.

The peer groups for the problem-solving sessions were composed of four students and were heterogeneous for both gender and ability. Students in both the talk-only and the talk-and-writing peer groups were instructed to use a three-step procedure for discussing the assigned problems. The procedure involved: (1) brainstorming possible explanations (Linn & Burbules, 1993); (2) elaborating, clarifying, and asking or answering questions about the proposed explanations (Webb, 1989); and (3) evaluating, criticizing, justifying, and revising these ideas. In addition, students were given a written prompt during these discussions to scaffold metacognitive awareness during the explanatory session (Coleman, 1992; Meloth & Deering, 1994; Palincsar et al., 1993).

RESULTS

Statistical treatment of the data involved a series of analyses of covariance using pretest scores as covariate each time. Hypothesis testing involved analyses based solely on the aggregate scores for each kind of knowledge (simple, integrated, and total). Other exploratory analyses were based on the three separate measures (multiple-choice tests, essay questions, and concept maps), as well as the aggregate scores, for both the immediate and delayed posttest results. These exploratory analyses were used to ensure that instrumentation was not confounding results in this small-group study. *Post hoc* analyses using least-significant-difference (LSD) pairwise comparisons were also examined for trends in the data. In addition, contrasts comparing either the two groups that used talk with the writing-only and control groups (TW + T = W + C), or the talk-and-writing group with the other three groups were also examined for both main treatment effects and gender–treatment interactions (TW = T + W + C). The first group contrast isolated the role of talk, either alone or augmented by writing, for learning science, whereas the second contrast focused instead on the synergistic effect of talk and writing.

Using group and gender as categorical variables, analysis of covariance (ANCOVA) was employed to explore the aggregate scores. Separate analyses were conducted for the immediate and delayed posttests. Table 1 shows the adjusted means and the standard deviations for the aggregate scores by treatment for all students using the pretests scores as covariate each time. An analysis of the adjusted means for the aggregate scores for all students shows that the rank order of the talk-and-writing group and the talk-only group were first and second, respectively. The writing-only and control groups ranked either third or fourth for all knowledge measures at both the immediate and delayed posttests.

Table 2 shows the adjusted means and the standard deviations for aggregate scores by gender and treatment with the pretest scores being used as covariate each time. The trends across treatment groups were quite different for boys and girls. For boys, the talk-and-writing and the writing-only groups generally ranked first and second, respectively, across the different aggregate knowledge measures for both posttests. Writing thus appeared to be more helpful than just talking for boys. In comparison, girls in the talk-only and the talk-and-writing groups always ranked either first or second for all measures and time-of-tests. Talking thus appeared to be more helpful than just writing for girls.

Three null hypotheses guided this study. Hypothesis testing involved analyses based on the aggregate scores for each kind of knowledge (simple, integrated, and total). *Post hoc* analyses using LSD pairwise comparisons were also examined for trends in the data.

Hypothesis One

The first hypothesis stated that there will be no significant difference ($p < 0.05$) in mean aggregate simple knowledge scores that can be attributed to: (a) treatment, (b) gender, and (c) interaction of treatment and gender. At the immediate posttest, no differences were

TABLE 1
Adjusted Means (M_{adj}) and Standard Deviations for Aggregate Scores by Treatment

	Posttest 1		Posttest 2	
	M_{adj}	SD	M_{adj}	SD
Simple knowledge				
Control	28.2	5.47	24.6	4.8
Talk	30.6	6.17	27.3	5.4
Talk-writing	33.3	5.37	30.5	4.71
Writing	29.5	5.77	26.1	5.07
Integrated knowledge				
Control	27.3	5.09	24.8	4.59
Talk	30.1	5.64	25.5	5.11
Talk-writing	30.1	4.95	29.8	4.5
Writing	26.5	5.31	23.6	4.81
Total knowledge				
Control	55.5	9.46	49.3	8.63
Talk	61.2	10.51	53.3	9.62
Talk-writing	63.2	9.25	60.2	8.45
Writing	55.9	9.92	49.5	9.05

TABLE 2
Adjusted Means (M_{adj}) and Standard Deviations for Aggregate Scores by Gender and Treatment

	Posttest 1				Posttest 2			
	Boys		Girls		Boys		Girls	
	M_{adj}	SD	M_{adj}	SD	M_{adj}	SD	M_{adj}	SD
Simple knowledge								
Control	30.1	5.51	26.4	5.42	27.0	4.83	22.3	4.76
Talk	31.4	5.46	29.7	5.40	25.3	4.78	29.2	4.75
Talk-writing	36.8	6.03	29.7	4.94	33.6	5.29	27.4	4.34
Writing	33.2	5.39	25.8	5.56	30.4	4.73	21.7	4.84
Integrated knowledge								
Control	29.4	5.14	25.3	5.0	27.1	4.65	22.5	4.52
Talk	32.3	4.98	27.9	5.02	26.2	4.50	24.8	4.54
Talk-writing	29.7	5.4	30.5	4.58	30.6	4.87	29.0	4.14
Writing	30.8	4.97	22.2	5.0	28.7	4.87	18.4	4.52
Total knowledge								
Control	59.5	9.55	51.5	9.34	54.0	8.72	44.6	8.54
Talk	63.7	9.31	58.7	9.30	51.3	8.51	55.3	8.5
Talk-writing	66.4	10.19	60.0	8.54	64.2	9.31	56.2	7.80
Writing	64.0	10.03	47.8	9.4	59.0	9.15	40.1	8.58

TABLE 3
Analysis of Covariance Summary for Simple Knowledge Aggregate Scores at the Delayed Posttest

Source	SS	*df*	MS	*F*	*p*
Gender	150.490	1	150.490	6.808	.013[a]
Group	201.961	3	67.320	3.046	.042[a]
Gender × group	204.372	3	68.124	3.082	.040[a]
Pretest	835.243	1	835.243	37.787	<.0005
Error	751.538	34	22.104		

[a]$p < .05$.

found for the main effects of treatment or for the interaction of treatment and gender when the aggregate simple knowledge scores were analyzed. However, significant differences were found between boys and girls for this measure with boys showing better immediate recall of simple knowledge than girls, $F(1, 34) = 8.309$, $p = .007$. At the delayed posttest, significant differences were apparent for treatment ($p = .042$), for gender ($p = .013$), and for gender–treatment interaction ($p = 0.04$). Results of the ANCOVA are shown in Table 3.

The *post hoc* analysis for main effects suggested that students who discussed the problems before writing their explanations showed better retention of simple knowledge over time than both the control group (TW > C, $p = .007$) and the group of students who just wrote without discussing the problems with peers (TW > W, $p = .038$). The *post hoc* pair comparisons for the gender–treatment interaction suggested that boys who had discussed the problems before writing their explanations showed better retention of simple knowledge over time than both groups of boys who were in the control group (TW > C, $p = .022$) or who had just discussed the problems without writing their explanations afterwards (TW > T, $p = .004$). The analysis of pair comparisons also suggested that girls who had simply discussed the problems with peers showed better retention of facts over time than those girls who had just written their explanation without any prior discussion (T > W, $p = .05$).

Hypothesis Two

The second hypothesis stated that there will be no significant difference ($p < .05$) in mean aggregate integrated knowledge scores that can be attributed to: (a) treatment, (b) gender, and (c) interaction of treatment and gender. The only significant difference in the data at the immediate posttest was for the effect of gender with boys again outperforming girls, $F(1, 34) = 6.367$, $p = .016$. At the delayed posttest, differences were observed for the separate effects of both gender and treatment. Results of the ANCOVA are shown in Table 4. *Post hoc* analysis of pairwise comparisons suggests that students who discussed problems with peers before individually writing their explanations (TW group) showed better retention of integrated knowledge over time than the other three groups (TW > C, $p = .015$; TW > W, $p = .003$; and TW > T, $p = .039$).

Hypothesis Three

Hypothesis three stated that there will be no significant difference ($p < .05$) in mean aggregate total knowledge scores that can be attributed to: (a) treatment, (b) gender, and

TABLE 4
Analysis of Covariance Summary for Integrated Knowledge Aggregate Scores at the Delayed Posttest

Source	SS	*df*	MS	*F*	*p*
Gender	189.814	1	189.814	9.443	.004[b]
Group	243.835	3	81.278	4.044	.015[a]
Gender × group	122.120	3	40.707	2.025	.129
Pretest	1054.790	1	1054.79	52.475	<.0005
Error	683.429	34	20.102		

[a]$p < .05$; [b]$p < .01$.

(c) interaction of treatment and gender. Similar to the previous analysis for integrated knowledge, the only significant difference at the immediate posttest was for the effect of gender, with boys again outperforming girls, $F(1, 34) = 8.78$, $p = .006$. At the delayed posttest, the analysis indicated that significant differences existed for the effects of treatment and gender. Furthermore, the value for the gender–treatment interaction approached significance ($p = .056$). Results of the ANCOVA are shown in Table 5. Students using the talk-and-writing strategy again outperformed those students who did not benefit from peer discussion (TW > C, $p = .005$; TW > W, $p = .006$).

DISCUSSION

The results suggest that peer discussion combined with analytical writing enhances the retention of science knowledge by students over time, but appears to have little effect on immediate learning. *Post hoc* pair comparisons between the talk-and-writing group and groups in which students worked individually on similar (writing-only group) or related descriptive tasks (control group) were particularly striking as the TW group outperformed these groups on all three knowledge measures (simple, integrated, and total knowledge). However, the differences between those students who simply discussed the problems in peer groups (talk-only group) and those who also used analytical writing (TW group) were significant for the integrated knowledge measure only. Peer discussion may be sufficient for the retention of facts and simple concepts, but may have to be augmented by writing for the retention of more complex integrated knowledge. The depth-of-processing hypoth-

TABLE 5
Analysis of Covariance Summary for Total Knowledge Aggregate Scores at the Delayed Posttest

Source	SS	*df*	MS	*F*	*p*
Gender	625.440	1	625.440	8.774	.006[b]
Group	857.144	3	285.715	4.008	.015[a]
Gender × group	595.280	3	198.427	2.784	.056
Pretest	3707.095	1	3707.095	52.005	<.0005
Error	2423.635	34	71.283		

[a]$p < .05$; [b]$p < .01$.

esis suggests that analytic writing should enhance this kind of learning, regardless of whether or not talk precedes it. However, the use of analytic writing alone did not enhance the learning of integrated knowledge in this study. Talk appears to have augmented the conceptual knowledge available to students prior to the actual processing phase, thereby enhancing their retention of integrated knowledge over time.

The anomaly in the data is the absence of any effects for treatment or for gender–treatment interaction at the immediate posttest, which was a measure of what students actually learned after the instructional unit was just completed. Because significant differences were observed after a delay of 6 wk, differences should also have been observed immediately following the instructional unit at the immediate posttest. Yet, adjusted mean aggregate scores at both immediate and delayed posttests generally favored the talk-and-writing group, followed by the talk-only group, the writing-only group, and the control group, in that order. However, differences may not have been large enough, given the small sample size, to establish significance at the immediate posttest. Nonetheless, significant differences were found at the immediate posttest on two of the multiple-choice measures: integrated knowledge, $F(3, 34) = 5.776$, $p = .003$; and total knowledge, $F(3, 34) = 4.746$, $p = .007$.

Talk for Sharing Knowledge

Peer discussion appeared to be an important mechanism for sharing knowledge. When the two groups that used talk were contrasted with the other two groups that did not benefit from peer interaction (TW + T ≠ W + C), significant differences were observed for the total aggregate score at the immediate posttest, $F(1, 34) = 4.719$, $p = .037$. In addition, the two other aggregate measures at the immediate posttest were marginally significant with alpha set at 0.10: the simple knowledge aggregate score, $F(1, 34) = 3.089$, $p = .088$, and the integrated knowledge aggregate score, $F(1, 34) = 3.895$, $p = .057$. Moreover, all three knowledge scores based on the multiple-choice measures at posttest 1 were significantly different: simple knowledge, $F(1, 34) = 7.754$, $p = .009$; integrated knowledge, $F(1, 34) = 7.830$, $p = .008$; and total knowledge, $F(1, 34) = 11.096$, $p = .002$. Although differences between these groups on the essay questions and the concept maps still favored those students using talk, the contrasts were not significant.

Glaser (1991) argued that small group discussions enhance student understanding by extending "available knowledge" (p. 134). Evidence of extending available knowledge and of distributing this knowledge to the students in the group was overwhelming in the transcripts. In the following transcript, students were asked to explain whether or not a dead log should be considered an ecosystem. One of the students in the group was absent from school when this problem session was held, so only three students participated in the peer discussion:

> Derek: Well, I don't think so, because / there is nothing // living.[1]
>
> Bill: (. . .) but it's biotic.
>
> Kelly: I think that the opposite is true. Because abiotic means dead / but there are all kinds

[1] The following transcription codes are used throughout: ". . .": ellipses indicates speech that trails off into silence; "[ ]": square brackets indicates the addition of words that facilitate comprehension of the transcript; "( )": parentheses indicates nonverbal cues and actions; "_": underlining of text indicates simultaneous or overlapping speech; "," and ".": commas and periods indicate normal breaks in flow of speech; "?": question mark indicates that the context of the speech is interpreted as a question; "(. . .)": indicates that words are undeciphered; "/": indicates a pause of less than 2 seconds; "//": indicates a pause of greater than 2 seconds.

of bugs and all that which . . .

Derek: Which live inside.

Kelly: And, yeah. And there is also . . .

Derek: And that is like their house (. . .)

Kelly: It decomposes and then, I don't know.

Derek: So, is it an ecosystem? It would depend. Because there are mosses all over, and there are parts which are dead already, so . . .

Bill: Yeah.

Kelly: But a log generally contains all kinds of, like, bugs.

Derek: (. . .) insects. So, it's considered an ecosystem?

Kelly: Yes, because . . .

Bill: Well, I think it is.

Kelly: Yes, because both parts are present / there is the abiotic, that's the dead log. And biotic, well, that's all those living organisms that live on the log.

The transcript demonstrates that talk was able to extend the available knowledge for these students. They were obviously grappling with the concepts of ecosystem, and abiotic and biotic factors. The problem task forced them to collectively reexamine their basic knowledge about these concepts and to determine whether or not the definitions that they had learned earlier in class applied to a dead log. All three students correctly differentiated between the abiotic and biotic components of ecosystems in their written responses immediately following the group discussion.

Four mechanisms appeared important during the group discussions: (1) asking questions, (2) hypothesizing, (3) formulating ideas together, and (4) explaining. The role of asking questions or asking for clarification seems to be an important catalyst for moving the discussion along. Two excerpts, both taken from the same discussion, illustrate this point. The students were discussing a question that involved examining various artifacts (deciduous leaves, coniferous needles, and cacti) and explaining how these were well adapted for their respective environments. In this first excerpt, the students had been examining coniferous specimens:

Lisa: Well, I didn't really understand David's point.

David: Okay, well, / I don't know why / like the needles / like those we find in Canada / it's like / it's like the needles are long and are solidly held on the branch. It's like the tree can't afford to grow / to lose its leaves / because / because there isn't enough sun. They don't have enough time to reproduce during the summer. Whereas these trees (holds up a leaf from a deciduous tree) are / it's softer and it's not as hard / and it's / they have enough time to lose their leaves and to regrow them.

A little later during the same discussion, the students had been grappling with the adaptations of the cactus for life in the desert:

Lisa: What was it you said at the beginning?

Melissa: That / there is water inside because there isn't much rainfall.

Lisa: (Points towards the cactus.) In that?

Melissa: Yeah, it conserves water because there isn't much rainfall.

Lisa: So, it doesn't need as much water as the other plants.

Melissa: Yeah, because it conserves water.

David: Bingo! They are adapted to a climate in which there isn't much precipitation. That's great!

Lisa: Uh-huh.

David: Okay.

Lisa's questions and her requests for clarification forced the group to elaborate their explanations until they made sense to her and to the others.

Hypothesizing also served as an engine for discussions by keeping students cognitively involved during these talk sessions. One of the problem tasks asked students to compare and contrast two biomes by analyzing unidentified climatograms: one, a tropical deciduous forest in Oaxaca, Mexico and the other, a desert in Fallon, Nevada. Students had already hypothesized that biome 2 likely represented a desert and were now analyzing the climatogram of the Mexican tropical forest. The discussion that ensued is described as follows:

Melissa: Biome 1 is probably a country /

Elliot: Somewhere in the tropics.

David: A forest.

Melissa: Not enough precipitation, I think.

Elliot: Well no (. . .)

David: Probably near the equator.

Melissa: Yeah.

Elliot: Yes / equatorial.

David: Oh yes / there is a // as you can see (pointing to the bar graph) / there is a / like in biome 1 / more / there is more / there is a / very little precipitation and then, all at once, there is a lot.

Melissa: That could be a desert, that one.

Elliot: That's the rainy season (pointing to the bars showing the precipitation from May through September).

Melissa: That could be the desert because it's quite warm (pointing to the line graph of monthly mean temperatures).

David: No, it's because /

Melissa: A lot of rain all at once.

David: Yes, but in a desert, there is less than one hundred.

Elliot: No, but the desert.

David: Less than 100 mm of precipitation.

Elliot: I think that it's a tropical forest.

David: A tropical forest.

The discussion was very animated, with students arguing back and forth about their various interpretations of the climatograms. Although students did not achieve consensus as to the identity of the climatograms during the discussion, talk served as a catalyst for negotiating and refining their collective understanding. In their written responses after the discussion, all four students correctly identified the biomes represented by these two climatograms. Individual students appeared to have refined their personal understanding while they wrote.

Schoenfeld used the metaphor of "ideas in the air" to describe what often happens when students work in peer discussion groups (1989, p. 71). Many cases of two or more students formulating ideas together were observed in the transcripts. In the following excerpt, the students were asked to compare the size of the ears and the shape of the head in three closely related fox species and to explain how these could be considered adaptations to the environment. The three species included the Arctic fox with small ears and a well-rounded head, the Fennec or desert fox with long, pointed ears and a slender muzzle, and the red fox which appears to be intermediate for these two traits:

> Kelly: Well, maybe (. . .) there is a lot of wind (. . .). I don't know.
>
> Derek: The Arctic fox is white and /
>
> Bill: We're talking about the ears!
>
> Derek: Well, the ears are small /
>
> Bill: So /
>
> Bonnie: They won't freeze.

The last part of the conversation involved three students essentially completing a single sentence among them. Despite the brevity of this sentence when the fragments are combined, the students still managed to come up with an adequate explanation for the small ears of the Arctic fox. The idea appeared to be floating around in the atmosphere and required verbal interactions among the students for its conceptualization. The students' written responses after the discussion suggest that they had developed a satisfactory understanding of this concept. All of them essentially explained that the Arctic fox's small ears could be considered an adaptation for cold climates because they minimized heat loss, thereby reducing the possibility of freezing these extremities. Student understanding seemed to be constructed privately after collectively sharing knowledge, which was diffuse and ill-defined.

Occasionally, one student in the group assumed a more dominant role in explaining ideas to peers. For instance, the student in the following excerpt is talking about the adaptations of coniferous trees to boreal climates:

> David: They are somewhat like cacti. / There is little precipitation so / like / they [the needles] can't fall / because there isn't a lot of // sunlight. Like / winter lasts about 4 or 5 months. Summer lasts about 2 or 3 [months] / so they don't have time to grow and to regrow again.

David's response is an adequate explanation for the adaptation of conifers to boreal climates. He recognizes the limits that climate has imposed on vegetative growth in northern biomes. However, this understanding was not shared by all students in his peer-discussion group.

Our analysis also suggests that there are limits to peer discussion. First, if the group, collectively, does not possess even basic ideas about the problem task, then peer discussion will be generally ineffective as a teaching strategy. Second, incomplete or poor understanding or an inability to apply the knowledge later on can also result despite constructive discussions with peers. Third, an uncritical acceptance of ideas, particularly those suggested by high-status students, can also hinder the effectiveness of talk in groups.

Writing for Consolidating Knowledge

Peer discussion combined with writing appeared to enhance the retention of science knowledge over time. When the talk-and-writing group was contrasted with the other three groups at the delayed posttest (TW $\neq$ T + W + C), significant differences were observed for all three aggregate measures: simple knowledge, $F(1, 34) = 7.47$, $p = .01$; integrated knowledge, $F(1, 34) = 11.167$, $p = .002$; and total knowledge, $F(1, 34) = 10.568$, $p = .003$. Furthermore, all three concept-mapping measures were also significantly different: simple knowledge, $F(1, 34) = 8.845$, $p = .005$; integrated knowledge, $F(1, 34) = 9.483$, $p = .004$; and total knowledge, $F(1, 34) = 9.488$, $p = .004$. Excluding the outliers from the analysis resulted in one other measure, the integrated knowledge measure based on the essay questions, being significantly different, $F(1, 33) = 4.686$, $p = .038$, and another for total knowledge with the essay questions marginally significant, $F(1, 33) = 3.494$, $p = .07$.

The talk-and-writing group responded in writing to each of the questions immediately after the discussion with peers. As such, these written responses are records of how students had appropriated the ideas discussed with peers, and how they had translated these into written text using the discursive tools available to them. Analysis of these written responses framed by the initial peer discussion suggests that talk was used for interpreting the problem task, and for generating, sharing, clarifying, and evaluating ideas. Writing, on the other hand, was used for organizing these ideas into a coherent response that respected grammatical and syntactic conventions. Although the discussions appeared to be rambling and disorganized at times, the written responses that resulted generally appeared to be more coherent and well organized. Oral discourse is divergent, highly flexible, and requires little effort of participants while they collectively explore ideas, but written discourse is convergent, more focused, and places greater cognitive demands on the writer.

For instance, one of the questions on food webs read as follows:

> The great gray owl is the provincial bird of Manitoba. This owl species feeds on small rodents, such as voles and shrews found in the Canadian tundra. Explain how the great gray owl ultimately depends on sunlight for food.

The following excerpt from the transcript on food webs describes the conversation based on this question:

> Kelly: Well, I think that / well / the sun makes the plants grow, and then the rodents and all that /
>
> Bill: Food chain.
>
> Kelly: / they eat the plants / and after that, they eat the other ones.
>
> Derek: Yeah, because the more vegetation there is, the more /
>
> Bill: Voles, and all that.
>
> Derek: Yeah, that will have /
>
> Bill: If there isn't enough vegetation, it will be more difficult for the owl to have some [voles].

The discussion appears to be quite chaotic with many incomplete ideas, interruptions, and vague and uncertain references evident. After discussing this question, the four students individually answered the question:

- They [the owls] depend on the sun so that the vegetation will grow, and the more vegetation [there is], the more rodents they can eat to feed themselves. (Derek)
- They [the owls] depend on the sun because the voles and shrews eat plants which require sunlight, and the less plants [there are] the less voles and shrews [there will be]. (Bill)
- Since vegetation requires sunlight for growth, the prey of the great gray owl eat this vegetation and the owl eats these herbivores. (Kelly)
- Since more sunlight will give more vegetation, and the rodents will be more abundant, and the owl will be able to feed itself. (Bonnie)

Although the peer discussion may appear trivial and incoherent, the written responses of these four students suggest that their conceptual understanding had surpassed these surface language fragments. All four students explicitly described a causal chain of events originating with the sun, to vegetation, to herbivores, then to owls in their written responses to the explanatory task.

The role of writing appeared to be particularly important to students for organizing their ideas. The following excerpt from the transcript on niche and habitat demonstrates that David, a high-ability student, used talk for defining the problem task and for generating and clarifying ideas relevant as solution to the problem. However, the written response that he completed immediately following this discussion suggests that writing was employed to organize these ideas into a logical and coherent response. The students had been asked to explain how a hunter might be able to use his knowledge of bear ecology (habitat, niche, reproduction, etc.):

David: The question is not very clear.

Elliot: Yes, it’s clear.

David: Well, no. It should read: How could a hunter use his knowledge of bear ecology to kill or hunt the bear.

Elliot: Well, hunting [and hunter] is the same thing. //

David: Okay. //

Melissa: Well, habitat / well, to know where to find the bear so you can hunt it.

Elliot: To know which bear he’s hunting.

David: He wouldn’t find it in a desert.

Melissa: Yeah.

Lisa: Habitat. Habitat, what?

Melissa: Habitat is where it lives. If he doesn’t know where it lives, how can he hunt it?

David: Well, its niche / like / if we know who hunts or who kills it / well, we would know that the bear is in the same general area as that animal.

Melissa: Where does it feed? Like if it finds something.

Elliot: He could look at trees because they scratch trees in their territory.//

David: Reproduction // well, reproduction / look for its young.

Elliot: No, there you will / you might run into the mother.

David: Well yes, exactly. The mother will be there, the bear will be there.

Elliot: She is ferocious, you can't hunt that. // Well, that's it.

Melissa: Well, how can he [the hunter] use his knowledge?

David: To observe. // To inform himself about the bear.

After this discussion with peers to define the problem and to generate and clarify ideas, David wrote the following response:

> The hunter could study or do some research on the bear species he wished to hunt. Its habitat: it's probable that the bear lives in a coniferous forest and not in a desert or other unlikely habitat. Its niche: The hunter could observe what the bear eats, it's quite probable that the bear would be in the area. He could observe what feeds on bears, once again the bear might be in the area. Signs left on the trunks of trees would be another factor. Reproduction: Where the young are located it is very probable that the mother bear would be in area to protect them.

Although the student erroneously believes that even large predators like bears are hunted by species other than humans, all of the other ideas can be considered satisfactory responses to the question. This case illustrates Sutton's (1992) view that there are two fundamental ways of using language: "One is exploratory, the other declarative, one is tentative while the other is definite . . ." (p. 49). Talk fulfilled the first function, whereas writing accomplished the second. Interestingly, all of the ideas suggested during the preceding discussion were incorporated into the written response, but with more coherence and structure. Explaining, or analytical writing, requires a reflective logical stance that encouraged students to refine their thinking thereby enhancing their conceptual understanding.

SUMMARY AND IMPLICATIONS

Lacking confidence in the quantitative data because of the small size of the sample, firm conclusions were considered to be inappropriate on the basis of this exploratory study. Nonetheless, trends were evident in the data and tentative conclusions have been proposed. Moreover, the qualitative data were used to inform the study regarding possible mechanisms for the observed effects.

Talk for Distributing Knowledge

Peer discussion appeared to be an important mechanism for sharing knowledge among students. The analyses of covariance and the *post hoc* contrasts both rejected the idea that writing alone enhances learning more than talk or peer discussion. Yet, individual writing is often the only strategy invoking language that is used in many classrooms. The results suggest that talk is important for distributing knowledge. Working individually, many students probably lacked the knowledge base for constructing adequate explanations for the various problem tasks and, consequently, did not learn as much working alone as those students who discussed these problems with peers. Talk or discussion appears to be important for sharing, clarifying, and distributing knowledge among peers. Asking questions, hypothesizing, explaining, and formulating ideas together all appear to be important mechanisms during these discussions.

Writing for Consolidating Knowledge

Peer discussion combined with writing appeared to enhance the retention of science knowledge over time. The analyses and *post hoc* contrasts confirmed that writing promotes the retention of science knowledge over time. However, for writing to be effective, students must already possess certain basic knowledge, and interacting with peers allowed them to share, clarify, and distribute this knowledge. Only then did writing exert a positive effect on learning. The qualitative data suggest that writing is an important discursive tool for organizing and consolidating rudimentary ideas into knowledge that is more coherent and well-structured. These findings suggest that writing is important for the retention of science knowledge over time, but that talk, or peer discussion, is a necessary precursor. Writing only seems to work if talk works with it.

Suggestions for Future Research

Interactions with Gender. Significant differences were observed between boys and girls for every aggregate knowledge measure and time-of-test. When the intact data were analyzed, however, only the simple knowledge measure showed significant differences, supporting an interaction between gender and treatment. On the one hand, *post hoc* pairwise comparisons suggested that boys who wrote after discussing the problems with peers showed better retention of facts and simple concepts over time than other boys who either worked alone on descriptive tasks or just discussed possible explanations with others without writing. On the other hand, girls who had the opportunity to discuss the problems with peers showed better retention of simple knowledge over time than other girls who just wrote in response to the problem tasks. This finding is consistent with the frequent call by many proponents of gender-inclusive classrooms for more discussion and collaboration in science teaching. Although the statistical evidence for interactions between treatment and gender was not very strong, these exploratory analyses nevertheless suggest that this question merits further study.

Rivard (1994) argued that "race, ethnicity, academic ability, and gender may all be variables that interact with the use of writing-to-learn strategies" (pp. 976–977). In this study, analytic writing was used as a heuristic for learning science. Would the use of expressive writing, as found in journals and learning logs, give similar results, or would this type of writing make the science classroom more gender-inclusive by personalizing learning for all students?

Interactions with Ability. Figure 1 compares the adjusted mean total aggregate scores on both immediate and delayed posttests of low- (lower quartile), average-, and high-ability (upper quartile) students, respectively, for the four treatment and control groups. Visual inspection of these graphs suggests that peer discussion may be particularly important for certain students. Students of low and average ability who used talk, either alone or combined with writing, appear to have learned more initially, while also showing better retention of this knowledge over time. These two groups of students seemed to benefit the most from sharing knowledge in peer groups.

In comparison, high-ability students appeared to do better when they individually responded in writing to the problem tasks. The high-ability students who individually wrote in response to the problem tasks not only showed better learning initially, as measured by the immediate posttest, but also showed better retention over time. However, because the talk-and-writing group ranked just behind the writing-only group on both posttests, one

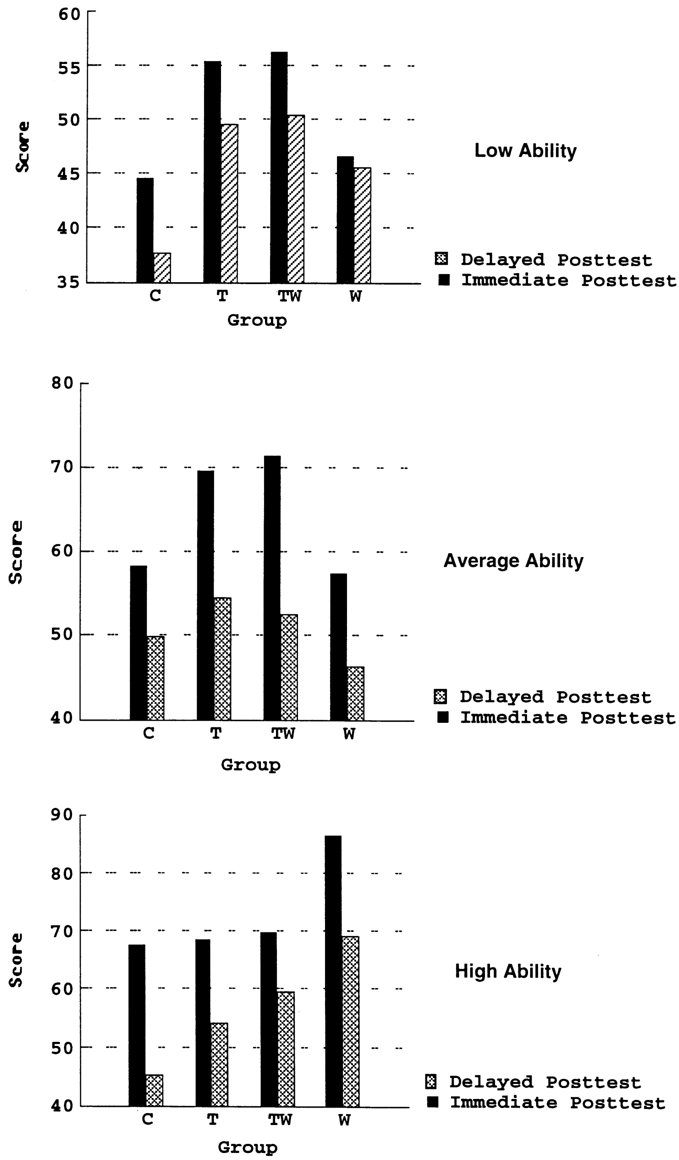

Figure 1. Bar graph of adjusted mean total aggregate scores by treatment and ability.

could argue that talk enhances science learning, even for high-ability students. Talking should enhance the learning of science concepts for the high-ability group of students because they tend to act as peer tutors by often explaining during discussion sessions (Webb, 1989). However, high-ability students are more likely than others to already possess the cognitive tools and the conceptual building blocks necessary for completing the problem tasks alone (Webb, 1992). The issue of whether or not an aptitude–treatment interaction exists between the use of language-based strategies involving talk and writing and student ability certainly merits further study.

Significance of the Study

The present study is significant for several reasons. Significant findings need to be helpful to classroom practitioners instead of just being theoretically or academically interesting (Duschl, 1994). The context for the study was authentic: A traditional instructional unit on ecology taught by an experienced middle school teacher was enhanced by simply integrating five 50-min problem-solving sessions involving peer discussion and analytical writing. Teachers should therefore be able to easily translate these research findings into classroom practice.

Second, the study addressed the issue of inert knowledge by proposing an instructional strategy which can be easily implemented in science classrooms that engages or accesses students' prior knowledge and enhances conceptual understanding (Cohen, McLaughlin, & Talbert, 1993; Perkins & Blythe, 1994). The findings suggest that science teachers should endeavor to include more writing tasks in the classroom, but only after students have had sufficient opportunities for collaborative exploratory talk while being guided by cognitively engaging problem-solving tasks. For instance, students might initially collaborate in small groups on the construction of a concept map, then individually transpose this graphic representation into written text. This classroom strategy would allow students to share and clarify ideas, while talking about the concept-mapping task, before consolidating and refining their knowledge while writing about the map.

Third, few studies have investigated how both talking and writing can influence classroom learning (Parker & Goodkin, 1987). Yet, writing may only work as a heuristic if talk precedes it. The review of the literature cited many studies that separately confirm the role of talk and the role of writing as heuristic strategies. However, no study has addressed the issue of how these two modalities cognitively mesh together during learning.

Conclusions

The results suggest that talk is important for sharing, clarifying, and distributing scientific ideas among peers while asking questions, hypothesizing, explaining, and formulating ideas together all appear to be important mechanisms during discussions. The use of writing appears to be important for refining and consolidating these new ideas with prior knowledge. These two modalities appear to be dialectical: talk is social, divergent, and generative, whereas writing is personal, convergent, and reflective. Moreover, writing appears to enhance the retention of co-constructed knowledge over time. Gender and ability are important variables that may be mediating the effects of talk and writing that should be investigated in a more robust future study.

The authors gratefully acknowledge Larry D. Yore's helpful comments in the preparation of this paper.

REFERENCES

Adams, M. (1992). Cultural inclusion in the American college classroom. In L. L. B. Border & N. V. N. Chism (Eds.), Teaching for diversity (pp. 5–17). San Francisco, CA: Jossey-Bass.

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. Review of Educational Research, 61, 315–343.

Alexopoulou, E., & Driver, R. (1996). Small-group discussion in physics: Peer interaction modes in pairs and fours. Journal of Research in Science Teaching, 33, 1099–1114.

Anderson, C. W., & Roth, K. J. (1989). Teaching for meaningful and self-regulated learning of science. In J. Brophy (Ed.), Advances in research on teaching (Vol. 1, pp. 265–309). Greenwich, CT: JAI Press.

Applebee, A. N. (1984). Writing and reasoning. Review of Educational Research, 54, 577–596.

Appleton, K. (1997). Analysis and description of students' learning during science classes using a constructivist-based model. Journal of Research in Science Teaching, 34, 303–318.

Audet, R. H., Hickman, P., & Dobrynina, G. (1996). Learning logs: A classroom practice for enhancing scientific sense making. Journal of Research in Science Teaching, 33, 205–222.

Baker, D., & Leary, R. (1995). Letting girls speak out about science. Journal of Research in Science Teaching, 32, 3–27.

Barnes, D. (1976). From communication to curriculum. London: Penguin Books.

Braithwaite, R. B. (1953). Scientific explanation: A study of the function of theory, probability and law in science. Cambridge, UK: Cambridge University Press.

Britton, J. (1982). Talking. In G. M. Pradl (Ed.), Prospect and retrospect: Selected essays of James Britton (pp. 112–122). Montclair, NJ: Boynton/Cook.

Britton, J. (1989). Writing and reading in the classroom. In A. H. Dyson (Ed.), Collaboration through writing and reading: Exploring possibilities (pp. 217–246). Urbana, IL: National Council of Teachers of English.

Bromberger, S. (1966). Why-questions. In R. G. Coldny (Ed.), Mind and cosmos: Essays in contemporary science and philosophy (pp. 86–111). Pittsburgh, PA: University of Pittsburgh Press.

Brophy, A. L., & Alleman, J. (1991). Activities as instructional tools: A framework for analysis and evaluation. Educational Researcher, 20, 9–23.

Brown, A. L., & Campione, J. C. (1990). Interactive learning environments and the teaching of science and mathematics. In M. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, A. diSessa, & E. Stage (Eds.), Toward a scientific practice of science education (pp. 111–139). Hillsdale, NJ: Lawrence Erlbaum.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 32, 32–42.

Bruffee, K. A. (1993). Collaborative learning: Higher education, interdependence and the authority of knowledge. Baltimore, MD: John Hopkins University Press.

Bruner, J. (1986). Actual minds, possible worlds. Cambridge, MA: Harvard University Press.

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (1997). Gender and science learning early in high school: Subject matter and laboratory experiences. American Educational Research Journal, 34, 297–331.

Cohen, D. K., McLaughlin, M. W., & Talbert, J. E. (Eds.) (1993). Teaching for understanding: Challenges for policy and practice. San Francisco, CA: Jossey-Bass.

Cole, N. S. (1990). Conceptions of educational achievement. Educational Researcher, 19, 2–7.

Coleman, E. B. (1992). Facilitating conceptual understanding in science: A collaborative explanation-based approach. Unpublished doctoral dissertation, University of Toronto, Toronto, ON, Canada.

Cooper, C. R. (1977). Holistic evaluation of writing. In C. R. Cooper & L. Odell (Eds.), Evaluating writing: Describing, measuring, judging (pp. 3–31). Urbana, IL: National Council of Teachers of English.

Corson, D. (1988). Oral language across the curriculum. Clevedon, UK: Multilingual Matters.

Cronin, P. J., Dekkers, J., & Dunn, J. (1982). A procedure for using and evaluating concept maps. Research in Science Education, 12, 17–24.

Davis, B. G., Scriven, M., & Thomas, S. (1987). The evaluation of composition instruction (2nd ed.). New York: Teachers College Press.

Davis, F., & Steiger, A. (1994). Feminist pedagogy and a new literacy for women. In S. P. Baardman, S. B. Straw, & L. E. Atkinson (Eds.), Social reflections on writing: To reach and realize (pp. 157–175). Winnipeg, MB: Literacy Publications.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23, 5–12.

Durst, R. K., & Newell, G. E. (1989). The uses of function: James Britton's category system and research on writing. Review of Educational Research, 59, 375–394.

Duschl, R. A. (1994). Editorial policy statement and introduction. Science Education, 78, 203–208.

Edwards, J., & Fraser, K. (1983). Concept maps as reflectors of conceptual understanding. Research in Science Education, 13, 19–26.

Erickson, G. L., & MacKinnon, A. M. (1991). Seeing classrooms in new ways: On becoming a science teacher. In D. A. Schön (Ed.), The reflective turn: Case studies in and on educational practice (pp. 15–36). New York: Columbia University Teachers College.

Fellows, N. J. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. Journal of Research in Science Teaching, 31, 985–1001.

Fosnot, C. T. (1993). Rethinking science education: A defense of Piagetian constructivism. Journal of Research in Science Teaching, 30, 1189–1201.

van Fraassen, B. (1987). The pragmatics of explanation. In J. A. Kourany (Ed.), Basic issues in the philosophy of science (pp. 65–67). Belmont, CA: Wadsworth.

Gaskins, I. W., Guthrie, J. T., Satlow, E., Ostertag, J., Six, L., Byrne, J., & Connor, B. (1994). Integrating instruction of science, reading, and writing: Goals, teacher development, and assessment. Journal of Research in Science Teaching, 31, 1039–1056.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. Learning and Instruction, 1, 129–144.

Goody, J. (1994). Entre l'oralité et l'écriture. Paris: Presses universitaires de France.

Gorman, T. P., Purves, A. C., & Degenhart, R. E. (Eds.) (1988). The IEA study of written composition. I: The international writing tasks and scoring scales. Oxford, UK: Pergamon Press.

Greene, S., & Ackerman, J. M. (1995). Expanding the constructivist metaphor: A rhetorical perspective on literacy research and practice. Review of Educational Research, 65, 383–420.

Halliday, M. A. K., & Martin, J. R. (1993). Writing science: Literacy and discursive power. Pittsburgh, PA: University of Pittsburgh Press.

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.

Hayes, D. A. (1987). The potential for directing study in combined reading and writing activity. Journal of Reading Behavior, 19, 333–352.

Hempel, C. G. (1965). Studies in the logic of explanation. In C. G. Hempel (Ed.), Aspects of scientific explanation and other essays in the philosophy of science (pp. 245–290). New York: Free Press.

Horwood, R. H. (1988). Explanation and description in science teaching. Science Education, 72, 41–49.

Howard, V. A. (1988). Thinking on paper: A philosopher's look at writing. In V. A. Howard (Ed.), Varieties of thinking: Essays from Harvard's philosophy of education research center (pp. 84–92). New York: Routledge.

Keys, C. W. (1996). Writing collaborative laboratory reports in ninth grade science: Three case studies of social interactions. School Science and Mathematics, 96, 178–186.

Kourany, J. A. (Ed.) (1987). Scientific knowledge: Basic issues in the philosophy of science. Belmont, CA: Wadsworth.

Langer, J. A. (1986). Learning through writing: Study skills in the content area. Journal of Reading, 29, 400–406.

Langer, J. A., & Applebee, A. N. (1987). How writing shapes thinking (Research Report No. 22). Urbana, IL: National Council of Teachers of English.

Lapointe, A. E., Mead, N. A., & Phillips, G. W. (1989). A world of differences: An international assessment of mathematics and science (Report No. 19-CAEP-01). Princeton, NJ: Educational Testing Service.

Lemke, J. L. (1990). Talking science: Language, learning, and values. Norwood, NJ: Ablex.

Linn, M. C., & Burbules, N. C. (1993). Construction of knowledge and group learning. In K. Tobin (Ed.), The practice of constructivism in science education (pp. 91–119). Washington, DC: AAAS Press.

Linn, M. C., & Songer, N. B. (1993). How do students make sense of science? Merrill-Palmer Quarterly, 39, 47–73.

Martin, J. R. (1970). Explaining, understanding, and teaching. New York: McGraw-Hill.
Mason, C. L. (1992). Concept mapping: A tool to develop reflective science instruction. Journal of Research in Science Teaching, 76, 51–63.
McGinley, G. A., & Tierney, R. J. (1989). Traversing the topical landscape: Reading and writing as ways of knowing. Written Communication, 6, 243–269.
McMillan, J. H., & Schumacher, S. (1989). Research in education: A conceptual introduction (2nd ed.). New York: Harper & Row.
Meloth, M. S., & Deering, P. D. (1994). Task talk and task awareness under different cooperative learning conditions. American Educational Research Journal, 31, 138–165.
Meyer, K., & Woodruff, E. (1997). Consensually driven explanation in science teaching. Science Education, 80, 173–192.
Miles, M. B., & Huberman, A. M. (1984). Qualitative data analysis: A sourcebook of new methods. Beverly Hills, CA: Sage.
Mullis, I. V. S., & Jenkins, L. B. (1988). The Science report card: Elements of risk and recovery. Trends and achievement based on the 1986 National Assessment. National Assessment of Educational Progress. Princeton, NJ: Educational Testing Service.
National Research Council. (1996). National Science Education Standards. Washington, DC: National Academy Press.
Newell, G. E. (1984). Learning from writing in two content areas: A case study/protocol analysis. Research in the Teaching of English, 18, 265–287.
Newell, G. (1986). Learning from writing: Examining our assumptions. English Quarterly, 19, 291–302.
Newmann, F. M. (1988). Can depth replace coverage in the high school curriculum. Phi Delta Kappan, 69, 345–348.
Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. Cambridge, UK: Cambridge University Press.
Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. American Educational Research Journal, 28, 117–153.
Nystrand, M. (1990). Sharing words: The effects of readers on developing writers. Written Communication, 7, 3–24.
O'Loughlin, M. (1992). Rethinking science education: Beyond Piagetian constructivism toward a sociocultural model of teaching and learning. Journal of Research in Science Teaching, 29, 791–820.
Olson, D. R. (1994). The world on paper: The conceptual and cognitive implications of writing and reading. Cambridge, UK: Cambridge University Press.
Palincsar, A. S., Anderson, C., & David, Y. M. (1993). Pursuing scientific literacy in the middle grades through collaborative problem solving. The Elementary School Journal, 93, 643–658.
Parker, R. P., & Goodkin, V. (1987). The consequences of writing: Enhancing learning in the disciplines. Upper Montclair, NJ: Boynton/Cook.
Pea, R. D. (1993). Learning scientific concepts through material and social activities: Conversational analysis meets conceptual change. Educational Psychologist, 28, 265–277.
Perkins, D., & Blythe, T. (1994). Putting understanding up front. Educational Leadership, 51, 4–7.
Peshkin, A. (1993). The goodness of qualitative research. Educational Researcher, 22, 23–29.
Piaget, J. (1950). La construction du réel chez l'enfant. Neuchâtel, Switzerland: Delachaux & Niestlé.
Prain, V. (1995). In B. Hand & V. Prain (Eds.), Teaching and learning in science: The constructivist classroom (pp. 63–81). Sydney, Australia: Harcourt Brace.
Prawat, R. S. (1989). Teaching for understanding: Three key attributes. Teaching and Teacher Education, 5, 315–328.
Prawat, R. S. (1993). The value of ideas: Problems versus possibilities in learning. Educational Researcher, 22, 5–16.
Prawat, R. S., & Floden, R. E. (1994). Philosophical perspectives on constructivist views of learning. Educational Psychology, 29, 37–48.
Purves, A. C. (1984). In search of an internationally-valid scheme for scoring compositions. College Composition and Communication, 35, 426–438.

Raphael, T. E., & Gavelek, J. R. (1984). Question-related activities and their relationship to reading comprehension: Some instructional implications. In G. G. Duffy, L. R. Roehler, & J. Mason (Eds.), Comprehension instruction: Perspectives and suggestions (pp. 234–250). New York: Longman.

Reif, F., & Larkin, J. H. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. Journal of Research in Science Teaching, 28, 733–760.

Resnick, L. B. (1987). Education and learning to think. Washington, DC: National Academy Press.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.) (1991). Perspectives on socially shared cognition. Washington, DC: American Psychological Association.

Rivard, L. P. (1994). A review of writing to learn in science: Implications for practice and research. Journal of Research in Science Teaching, 31, 969–983.

Rogoff, B., & Lave, J. (Eds.). (1984). Everyday cognition: Its development in social context. Cambridge, MA: Harvard University Press.

Rogoff, B. (1990). Apprenticeships in thinking: Cognitive development in social context. New York: Oxford University Press.

Rosser, S. V. (1990). Female-friendly science: Applying women's studies methods and theories to attract students. New York: Pergamon Press.

Roth, K. J. (1990). Developing meaningful conceptual understanding in science. In B. F. Jones & L. Idol (Eds.), Dimensions of thinking and cognitive instruction (pp. 139–175). Hillsdale, NJ: Lawrence Erlbaum.

Roth, W.-M. (1994a). Student views of collaborative concept mapping: An emancipatory research project. Science Education, 78, 1–34.

Roth, W.-M. (1994b). Experimenting in a constructivist high school physics lab. Journal of Research in Science Teaching, 31, 187–223.

Salmon, W. C. (1987). Why ask, "why?"? An inquiry concerning scientific explanation. In J. A. Kourany (Ed.), Scientific knowledge: Basic issues in the philosophy of science (pp. 51–64). Belmont, CA: Wadsworth.

Sawyer, M. H. (1991). A review of research in revising instructional text. Journal of Reading Behavior, 33, 307–333.

Scantlebury, K., & Kahle, J. B. (1993). The implementation of equitable strategies by high school biology student teachers. Journal of Research in Science Teaching, 30, 537–545.

Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 778–803). New York: Macmillan.

Schoenfeld, A. H. (1989). Ideas in the air: Speculations on small group learning, environmental and cultural influences on cognition, and epistemology. International Journal of Educational Research, 13, 71–88.

Schumacher, G. M., & Nash, J. G. (1991). Conceptualizing and measuring knowledge change due to writing. Research in the Teaching of English, 25, 67–96.

Science Council of Canada. (1984). Science for every student: Educating Canadians for tomorrow's world. Ottawa, ON: Ministry of Supply and Services.

Shepardson, D. P., & Pizzini, E. L. (1991). Questioning levels of junior high schools science textbooks and their implications for learning textual information. Science Education, 75, 673–682.

Smail, B. (1987). Organizing the curriculum to fit girls' interests. In A. Kelly (Ed.), Science for girls? Milton Keynes, UK: Open University Press.

Stuart, H. A. (1985). Should concept maps be scored numerically? European Journal of Science Education, 7, 73–81.

Sutton, C. (1992). Words, science and learning. Buckingham, UK: Open University Press.

Thagard, P. (1989). Explanatory coherence. Behavioral and Brain Sciences, 12, 435–502.

Thaiss, C. (1988). The future of writing across the curriculum. In S. H. McLeod (Ed.), Strengthening programs for writing across the curriculum. New directions for teaching and learning (Vol. 36, pp. 91–102). San Francisco: Jossey-Bass.

Tharp, R. G., & Gallimore, R. (1988). Rousing minds to life: Teaching, learning, and schooling in social context. Cambridge, UK: Cambridge University Press.

Tishman, S., & Perkins, D. (1997). The language of thinking. Phi Delta Kappan, 78, 368–374.

Tobin, K. (Ed.) (1993). The practice of constructivism in science education. Washington, DC: American Association for the Advancement of Science.

Vygotsky, L. S. (1962). Thought and language. Cambridge, MA: MIT Press.
Vygotsky, L. S. (1978). Mind in society. Cambridge, MA: Harvard University Press.
Wallace, J. D., & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. Journal of Research in Science Teaching, 27, 1033–1052.
Wandersee, J. H. (1990). Concept mapping and the cartography of cognition. Journal of Research in Science Teaching, 27, 923–936.
Weaver, W. (1964). Scientific explanation. Science, 143, 1297–1300.
Webb, N. M. (1989). Peer interaction and learning in small groups. International Journal of Educational Research, 13, 21–39.
Webb, N. M. (1992). Testing theoretical model of student interaction and learning in small groups. In R. Hertz-Lazorowitz & N. Miller (Eds.), Interaction in cooperative groups: The theoretical anatomy of group learning (pp. 102–119). Cambridge, UK: Cambridge University Press.