THE EFFECT OF TALK AND WRITING, ALONE AND COMBINED, ON LEARNING IN SCIENCE: AN EXPLORATORY STUDY

BY

LÉONARD P. RIVARD

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

Ph.D.

Faculty of Education
University of Manitoba
Winnipeg, Manitoba

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-612-13481-4

Canada

Name ______________________________

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

LANGUAGE & LITERATURE | 0 2 7 9 | U·M·I

SUBJECT TERM | SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
- Architecture 0729
- Art History 0377
- Cinema 0900
- Dance 0378
- Fine Arts 0357
- Information Science 0723
- Journalism 0391
- Library Science 0399
- Mass Communications 0708
- Music 0413
- Speech Communication 0459
- Theater 0465

EDUCATION
- General 0515
- Administration 0514
- Adult and Continuing 0516
- Agricultural 0517
- Art 0273
- Bilingual and Multicultural 0282
- Business 0688
- Community College 0275
- Curriculum and Instruction 0727
- Early Childhood 0518
- Elementary 0524
- Finance 0277
- Guidance and Counseling 0519
- Health 0680
- Higher 0745
- History of 0520
- Home Economics 0278
- Industrial 0521
- Language and Literature 0279
- Mathematics 0280
- Music 0522
- Philosophy of 0998
- Physical 0523
- Psychology 0525
- Reading 0535
- Religious 0527
- Sciences 0714
- Secondary 0533
- Social Sciences 0534
- Sociology of 0340
- Special 0529
- Teacher Training 0530
- Technology 0710
- Tests and Measurements 0288
- Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
- Language
  - General 0679
  - Ancient 0289
  - Linguistics 0290
  - Modern 0291
- Literature
  - General 0401
  - Classical 0294
  - Comparative 0295
  - Medieval 0297
  - Modern 0298
  - African 0316
  - American 0591
  - Asian 0305
  - Canadian (English) 0352
  - Canadian (French) 0355
  - English 0593
  - Germanic 0311
  - Latin American 0312
  - Middle Eastern 0315
  - Romance 0313
  - Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
- Philosophy 0422
- Religion
  - General 0318
  - Biblical Studies 0321
  - Clergy 0319
  - History of 0320
  - Philosophy of 0322
- Theology 0469

SOCIAL SCIENCES
- American Studies 0323
- Anthropology
  - Archaeology 0324
  - Cultural 0326
  - Physical 0327
- Business Administration
  - General 0310
  - Accounting 0272
  - Banking 0770
  - Management 0454
  - Marketing 0338
- Canadian Studies 0385
- Economics
  - General 0501
  - Agricultural 0503
  - Commerce-Business 0505
  - Finance 0508
  - History 0509
  - Labor 0510
  - Theory 0511
- Folklore 0358
- Geography 0366
- Gerontology 0351
- History
  - General 0578
  - Ancient 0579
  - Medieval 0581
  - Modern 0582
  - Black 0328
  - African 0331
  - Asia, Australia and Oceania 0332
  - Canadian 0334
  - European 0335
  - Latin American 0336
  - Middle Eastern 0333
  - United States 0337
- History of Science 0585
- Law 0398
- Political Science
  - General 0615
  - International Law and Relations 0616
  - Public Administration 0617
- Recreation 0814
- Social Work 0452
- Sociology
  - General 0626
  - Criminology and Penology 0627
  - Demography 0938
  - Ethnic and Racial Studies 0631
  - Individual and Family Studies 0628
  - Industrial and Labor Relations 0629
  - Public and Social Welfare 0630
  - Social Structure and Development 0700
  - Theory and Methods 0344
- Transportation 0709
- Urban and Regional Planning 0999
- Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
- Agriculture
  - General 0473
  - Agronomy 0285
  - Animal Culture and Nutrition 0475
  - Animal Pathology 0476
  - Food Science and Technology 0359
  - Forestry and Wildlife 0478
  - Plant Culture 0479
  - Plant Pathology 0480
  - Plant Physiology 0817
  - Range Management 0777
  - Wood Technology 0746
- Biology
  - General 0306
  - Anatomy 0287
  - Biostatistics 0308
  - Botany 0309
  - Cell 0379
  - Ecology 0329
  - Entomology 0353
  - Genetics 0369
  - Limnology 0793
  - Microbiology 0410
  - Molecular 0307
  - Neuroscience 0317
  - Oceanography 0416
  - Physiology 0433
  - Radiation 0821
  - Veterinary Science 0778
  - Zoology 0472
- Biophysics
  - General 0786
  - Medical 0760

EARTH SCIENCES
- Biogeochemistry 0425
- Geochemistry 0996
- Geodesy 0370
- Geology 0372
- Geophysics 0373
- Hydrology 0388
- Mineralogy 0411
- Paleobotany 0345
- Paleoecology 0426
- Paleontology 0418
- Paleozoology 0985
- Palynology 0427
- Physical Geography 0368
- Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
- Environmental Sciences 0768
- Health Sciences
  - General 0566
  - Audiology 0300
  - Chemotherapy 0992
  - Dentistry 0567
  - Education 0350
  - Hospital Management 0769
  - Human Development 0758
  - Immunology 0982
  - Medicine and Surgery 0564
  - Mental Health 0347
  - Nursing 0569
  - Nutrition 0570
  - Obstetrics and Gynecology 0380
  - Occupational Health and Therapy 0354
  - Ophthalmology 0381
  - Pathology 0571
  - Pharmacology 0419
  - Pharmacy 0572
  - Physical Therapy 0382
  - Public Health 0573
  - Radiology 0574
  - Recreation 0575
  - Speech Pathology 0460
  - Toxicology 0383
- Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
- Chemistry
  - General 0485
  - Agricultural 0749
  - Analytical 0486
  - Biochemistry 0487
  - Inorganic 0488
  - Nuclear 0738
  - Organic 0490
  - Pharmaceutical 0491
  - Physical 0494
  - Polymer 0495
  - Radiation 0754
- Mathematics 0405
- Physics
  - General 0605
  - Acoustics 0986
  - Astronomy and Astrophysics 0606
  - Atmospheric Science 0608
  - Atomic 0748
  - Electronics and Electricity 0607
  - Elementary Particles and High Energy 0798
  - Fluid and Plasma 0759
  - Molecular 0609
  - Nuclear 0610
  - Optics 0752
  - Radiation 0756
  - Solid State 0611
- Statistics 0463

Applied Sciences
- Applied Mechanics 0346
- Computer Science 0984
- Engineering
  - General 0537
  - Aerospace 0538
  - Agricultural 0539
  - Automotive 0540
  - Biomedical 0541
  - Chemical 0542
  - Civil 0543
  - Electronics and Electrical 0544
  - Heat and Thermodynamics 0348
  - Hydraulic 0545
  - Industrial 0546
  - Marine 0547
  - Materials Science 0794
  - Mechanical 0548
  - Metallurgy 0743
  - Mining 0551
  - Nuclear 0552
  - Packaging 0549
  - Petroleum 0765
  - Sanitary and Municipal 0554
  - System Science 0790
- Geotechnology 0428
- Operations Research 0796
- Plastics Technology 0795
- Textile Technology 0994

PSYCHOLOGY
- General 0621
- Behavioral 0384
- Clinical 0622
- Developmental 0620
- Experimental 0623
- Industrial 0624
- Personality 0625
- Physiological 0989
- Psychobiology 0349
- Psychometrics 0632
- Social 0451

THE EFFECT OF TALK AND WRITING, ALONE AND COMBINED, ON LEARNING IN SCIENCE: AN EXPLORATORY STUDY

BY

LÉONARD P. RIVARD

A Thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

EXPANDED ABSTRACT

The ultimate purpose of the study was to serve as a template for new research. To this end, trends in the data were investigated in order to formulate hypotheses and questions, and to recommend procedures that might be used in a follow-up study. The more immediate or proximate purpose of this exploratory study was to investigate the role of talk and writing, alone and combined, on learning in science. Moreover, gender and ability were considered in the design and analyses in order to identify any possible aptitude-treatment interactions.

The study was conducted in a Franco-Manitoban school, grade seven to twelve, in a major city in the Canadian prairies. Two intact grade eight science classes, both taught by a teacher who volunteered to participate in the study, served as the sample. The 43 students were randomly assigned to three treatment and one control groups, all stratified for gender and ability. All three treatment groups received the same problem tasks, once per school cycle, during the teaching of a science unit on ecology. The control group received simpler descriptive tasks based on similar content. Apart from these sessions, the instruction was identical for all students. The tasks in the problem sessions involved constructing scientific explanations for real-world applications of ecological concepts. The tasks required that students integrate simple knowledge about ecology into more complex and elaborate knowledge.

Students in the talk-only group (T) discussed the problem tasks in small peer groups. Students in the writing-only group (W) individually wrote responses for each of the tasks but without talking to other students. Students in the combined talk and writing group (TW) discussed the problems in small peer groups prior to individually writing a response for each question. The control group (C) completed simple learning

---
Abstract, Dissertation - disk 33A (95-11-27)

tasks based on the same material during the problem sessions.

The independent variables included treatment (T, W, TW & C), gender and ability. Dependent variables included simple, integrated and total knowledge scores based on multiple choice tests, essay questions, concept maps, as well as aggregate measures, at three times during the study (pretest , immediate posttest and delayed posttest). In addition, some of the discussions were videotaped, transcribed and analyzed to determine how student understanding evolved in these peer groups.

Statistical treatment of the data involved a series of analyses of covariance (ANCOVA) using pretest scores as covariates each time. Post hoc analyses using pairwise comparisons and group contrasts were also employed to examine trends in the data. The findings suggest that peer discussion appears to be an important mechanism by which students construct knowledge. Furthermore, peer discussion 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. Questions and hypotheses have been formulated for future research using revised instruments and methods with a larger sample size.

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to Dr. Stanley Straw for his support and guidance throughout my doctoral studies. His keen intellect, incisive comments, and constructive suggestions made my doctoral apprenticeship a period of significant personal and intellectual growth. Appreciation is also expressed to members of the committee, Dr. Arthur Stinner and Dr. Gordon Robinson, and to the external examiner, Dr. Donna Alvermann, for their insightful comments.

In addition, I would like to thank the participating students at Collège Louis-Riel for their complete cooperation from beginning to end of the study. I would also like to extend my gratitude, particularly to Mr. Mayur Raval, teacher, and also to Mr. Léo Robert and Mrs. Monique Fisette, principal and assistant principal. Their enthusiasm, professionalism, and commitment to excellence made the research project an enjoyable and rewarding experience for me. I would also like to express my appreciation to the administration of the *Division scolaire franco-manitobaine,* particularly Mr. Raymond Bisson, superintendent, for allowing me to conduct the study in this newly created school division during the hectic first year of its existence. I am also endebted to the Social Sciences and Humanities Research Council of Canada for providing financial support for this investigation.

Finally, my deepest appreciation goes to my wife Linda and my three sons, Christian, Nathan and Benjamin. To them I am ever grateful for their unfailing support, understanding and encouragement during this journey together.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.0 Statement of the Problem

The problem that was addressed in this study is the effect of writing alone, talking alone, and writing and talking together on science learning when students work on problems requiring them to explain real-world applications of scientific concepts (explanatory task).

1.1 Background of the Problem

1.1.0 *Covering the Material*

In its national study of science education, the Science Council of Canada observed that science at the middle school level "is often presented as a catalogue of facts for the students to assimilate as quickly as possible" (1984a, p. 31). Students are rarely required to construct explanations of natural phenomena or of real-world applications of scientific concepts. Much of the curriculum is theoretical and abstract with teachers viewing students as passive recipients of scientific knowledge (Science Council of Canada, 1984b). In summarizing case studies of science teaching in Canadian schools that were conducted for the Science Council, Orpwood and Souque (1984) reported that:

> In the middle years, the emphasis is on covering a considerable body of material in the time available. "Covering the material" means that the "correct" explanation must be included in students' notes. Teachers stress the specialized vocabulary of science, access to which is controlled through notes and activity sheets designed by teachers . . . . Middle-years teachers emphasize routines, standards of accuracy and thoroughness. For them, accuracy is at the heart of what they believe to be a scientific approach to problems. This emphasis on approved explanations and the right answer is at odds with the process of inquiry and the conceptual and tentative status of knowledge in science. Yet, such predictable activities as note-taking, copying activity sheets and lab procedures are valued because the accumulated information provides a base for work in the next grade, and because they control and

> channel energies by keeping students busy with routine, unambiguous work. Teachers appear reluctant to introduce into their well-ordered and coherent system any activity that might upset the smooth running of things. These teachers seem to make very restricted use of the potential that science has for general education. (pp. 21-22)

This transmission-absorption model of science teaching is not unique to the Canadian educational system (Hendry & King, 1994; Scardamalia & Bereiter, 1989). In their report for the National Assessment of Educational Progress (NAEP) in the U.S., Mullis and Jenkins (1988) had indicated that science instruction in the U.S. was predominantly accomplished by teacher lecturing. Anderson and Roth (1989) have commented that the typical science class in the U.S. is generally 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) suggest 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). Reif and Larkin (1991) have stated that "science courses taught in schools often do not adequately foster the scientific goal of understanding, or may inadvertently even pervert it" (p. 744).

1.1.1 *Evidence from Assessments*

In Canada and Abroad

The deficiencies in this transmission-absorption conception of teaching-learning are readily apparent judging from the findings of science assessments, some from within Canada and others beyond its boundaries. The assessments, unanimously, show that many students "can repeat science facts and principles, but in explanations of events fail to use them" (Cole, 1990, p. 4). An international assessment of mathematics and science in which thirteen-year-olds from four Canadian provinces participated (British Columbia, New Brunswick, Ontario, and Québec) indicated that student performance for applying scientific knowledge and principles ranged from less than 10% (Ontario and New Brunswick's French population) to 31% in British Columbia (Lapointe, Mead, & Phillips, 1989). Gender differences favouring boys were observed in all but two populations, the United Kingdom and the United States. The other ten populations, including all four Canadian provinces, showed performance differences between female and male students.

In the United Kingdom

In the United Kingdom, the Assessment of Performance Unit (APU), which is affiliated with the Department of Education and Science, has been very active in monitoring student strengths and weaknesses in the sciences. One APU survey of science, which was based on a representative sample of all fifteen-year-olds in England, Wales and Northern Ireland, examined student strengths and weaknesses in a number of different categories like using graphs, tables and charts, making observations, interpreting presented information or applying concepts. On the basis of their findings, the APU reported that the three biggest weaknesses involved applying biology concepts (mean performance level of 35.5%), applying chemistry concepts (36.5%) and applying physics concepts (31.0%) (Department of Education and Science, 1985, p. 34). Comparing the profile of performance for different ability groups showed that low-ability students have considerable difficulty with interpreting data and applying science concepts. The authors also noted gender differences in favour of males for most of the categories. Moreover, the largest gender difference involved applying physics concepts, particularly those dealing with electricity.

In Manitoba

In Manitoba, provincial science assessments were conducted in science at grades 3, 6, and 9 in 1986, and more recently in 1990, in almost all of the secondary science subjects. In the 1986 assessment, content items were grouped under six categories: Facts, terminology, concepts, theories, explanations and trends. Performance on the explanatory items was considered marginal for sixth-graders and unsatisfactory for the ninth-graders. The authors of the report recommended that greater emphasis should be placed on applying scientific knowledge to real-world situations and that teachers should use classroom activities "to integrate concepts into larger, more encompassing theories" (Manitoba Education, 1986, p. 122).

Results of the 1990 survey of physics teachers revealed that only about 30% of them frequently ask students to apply information to novel problem situations (Manitoba Education and Training, 1993a). It was further recommended that teachers establish a program that better combines theory with applications: "Let learning be related to everyday life as much as possible" (Manitoba Education and Training, 1993b).

Some of the recommendations in the 1990 Science Assessment report have even addressed the issue of how teachers might achieve the goal of relating science to everyday life. Based on the assessment of grade ten science, the authors recommended that students be given more opportunities in the classroom for writing about, and for discussing, applications of science concepts (Manitoba Education and Training, 1993b, pp. 11-12). The results of the 1990 Biology Assessment showed that students generally do well on simple knowledge or recall questions, but that questions requiring an integration of knowledge, or "linking ideas together and relating them to a general process," present considerable difficulty for them (Manitoba Education and Training, 1993b, p. 14). The authors recommended that "students need to be given specific training in writing short and long answers on an ongoing basis" using application or other higher-order types of questions (Manitoba Education and Training, 1993b, p. 18).

1.1.2 *Difficulties in Applying Knowledge*

Problem of Inert Ideas

The foregoing findings suggest that science students have considerable difficulty interpreting data and applying science concepts. Students seem to perform adequately when assessments evaluate simple knowledge or *verbatim* recall of information like definitions and descriptions or problems which can be solved using simple algorithms. Students probably have not sufficiently integrated their knowledge into an elaborate, coherent cognitive structure. Rather, student knowledge seems to be composed of "inert ideas," an amorphous mass of bits of information with few connecting links (Whitehead, 1929). Case studies of middle school classes depict learners as passive consumers of science knowledge where the "emphasis is on the certain, the exact, the 'right' answer . . . [and] 'approved' explanations" (Science Council of Canada, 1984b, p. 20). Moreover, student alienation from the sciences has been recognized by a number of scholars and government agencies as a potentially serious problem facing Western industrialized countries (Glasser, 1990; Hurd, 1990; Linn, 1990; Rutherford & Ahlgren, 1989; Science Council of Canada, 1984a; Yager, 1989).

Making Connections

Discrete knowledge should not be learned for its own sake, as appears to be the case in many science classrooms (Newmann, 1988). According to the Science Council of Canada, "ensuring excellence . . . should not mean that students are subjected to abstractions they can scarcely comprehend or relate to their personal lives" (1984a, p. 14). The Council believes that "it is important for students to learn that they can understand and deal with the world by means of their own observations and constructed explanations" (1984a, p. 17). Students must be more actively involved in constructing their personal knowledge: Glaser (1991) has written that "learners should become mindful architects of their own knowledge" (p. 131). Roth (1990) has argued that learning should be "an active process in which the learner takes information from the environment and constructs personal interpretations and meanings" (p. 143). Erickson and MacKinnon (1991) view "learners as being purposeful sense-makers — constantly engaged in the task of constructing ideas to make sense out of the situations they encounter" (p. 16). Science students must have ample opportunities for making meaningful connections between classroom learning and the concrete situations which they encounter in daily life (Anderson, 1987; Stinner, 1992).

1.1.3 *The Use of Questions*

Questions, either through direct questioning by the teacher during a lesson, or as adjuncts to whole-class instruction in textbooks and worksheets, can be used for helping the student to make connections between the classroom and daily life. Questions can enhance student learning by forcing them to "confront, link, elaborate, extend, or delimit the concepts into a growing web of science knowledge" (Rivard & Yore, 1992, p. 19). Tharp and Gallimore (1988) have described questioning as a potential strategy for enhancing student learning: They suggest that "questions call up the use of language and in this way assist in thinking" (p. 59). The critical issue here is the kind of question which is being asked of the student. Tharp and Gallimore distinguish between questions which are traditionally used to assess student learning and those which can be used to assist student learning.

Questions can be classified according to the cognitive demands required for answering them (Shepardson & Pizzini, 1991). Input questions require students to simply recall information. Processing questions require students to establish relationships among the information bits, whereas output questions require students to transform the information in meaningful ways. In their analysis of questions in middle school science textbooks, Shepardson and Pizzini suggested that "an overabundance of input-level questions . . . [may] inhibit the students' cognitive level of interaction" with the content, thereby inhibiting meaningful learning (p. 679). On the basis of their review, Raphael and Gavelek (1984) concluded that teacher questions can enhance learning when used to review important ideas or to clarify and consolidate subtle relationships among concepts. Sawyer (1991) supports this conclusion, and indicated that meaningful learning questions, or those requiring high cognitive processing like application, analysis, synthesis and evaluation, facilitate intentional learning, whereas explicit questions facilitate only rote recall.

Questions can be used for controlling the student through the use of routine mechanical tasks or for engaging students with science content in meaningful cognitive tasks. The use of alternate learning strategies for engaging students with the science content, while being guided by a question which requires processing an array of facts and ideas, have much to offer classroom teachers.

1.1.4 *Learning Strategies*

Learning strategies can be defined as "behaviors that the learner engages in during learning that are intended to influence affective and cognitive processing during encoding" (Weinstein & Mayer, 1986, p. 316). Classroom activities which feature listening, talking, reading and writing can all be organized so that the cognitive processing of information is enhanced. Knoblauch and Brannon (1984) have argued that using language can be:

> a heuristic process whereby the effort to assert connections and array them as integrated verbal patterns—texts—yields new understanding: In effect, new knowledge. Discourse, then, far from having the restrictive presentational function that the ancient rhetoricians supposed actually has a central, generative role in the pursuit of knowledge. (p. 53)

Using Writing

The use of writing as a learning strategy has received considerable theoretical support from scholars in a variety of disciplines (Applebee, 1984a; Barnes, 1976; Beyer, 1982; Britton, 1989; Elbow, 1983; Emig, 1977; V. A. Howard, 1988; Langer, 1986a; Lemke, 1990; Odell, 1980; Resnick, 1987; Scardamalia & Bereiter, 1986; Shanahan, 1988; Van Nostrand, 1979; Vygotsky, 1962). Vygotsky (1962), for instance, has referred to writing as the "deliberate structuring of the web of meaning," which suggests that writing might be useful as a learning strategy. Moreover, the use of short explanatory writing tasks for enhancing science learning has been recommended by many educators (Comley, Hamilton, Klaus, Scholes, & Sommers, 1984; Kirkpatrick & Pittendrigh, 1984; Stanley, 1991; and Strenski, 1984). According to these authors, using explanatory writing tasks in the science classroom is a useful strategy for encouraging students to link concepts, thereby enhancing meaningful learning (Novak & Gowin, 1984). The research base supporting the use of writing as a learning strategy is extensive and, along with the underlying theoretical arguments, will be reviewed in chapter two.

Using Talking

However, writing, by itself, may not be quite as helpful for conceptualizing relationships as a strategy which combines writing and talking. Thaiss (1988) has suggested that "the writing part works only if reading, talking, and listening work with it" (p. 99). While acknowledging all four of these modes, Fillion (1985) has argued that "informal, 'exploratory' language plays an important role in moving pupils toward understanding and appreciation of the information and ideas presented to them in curriculum" (p. 2881). The use of talking as a learning strategy has received theoretical support from scholars both within and beyond the science education community (Barnes, 1976; Britton, 1970, 1982; Bruner, 1986; Lemke, 1990; Martin, D'Arcy, Newton, & Parker, 1976; Pea, 1993; Prawat, 1993; Vygotsky, 1962; Wells, 1986). Based on their research findings, Durst and Newell (1989) have underlined "the importance of students using language—both oral and written—to explore, organize, and refine their ideas about themselves and their subject matter" (p. 375). Furthermore, Fellows (1994) had also observed that combining talk with writing was more effective for promoting conceptual change in students. Both the theoretical arguments and the research base supporting the use of talking as a learning strategy will be considered more extensively in the literature review which follows.

Establishing an Appropriate Context

Creating an environment in which the talk among students can flow freely without constraints, where students can explore ideas without fear of criticism or without being confronted prematurely with the authoritative or canonical explanation, is an essential condition for the effective deployment of these language-based learning strategies (McLaughlin & Talbert, 1993). Students must be given ample opportunities for formulating their own ideas about science concepts, for inferring relationships between and among these concepts, and for combining them into an increasingly more complex network of theoretical propositions (Brey, 1984). Moreover, studies suggest that talking within small peer groups, rather than talking with the teacher and other students in a whole-class situation, may be more effective for encouraging this kind of learning and conceptualization (Johnson & Johnson, 1990, 1991; Kagan, 1992; Slavin, 1983, 1990).

1.2 Purpose of the Study

The purpose of the study was: (1) to develop, pilot and refine research instruments for collecting data; (2) to investigate trends in the data collected during the pilot phase; (3) to develop questions and a proposal for future research work. The study investigated the role of writing alone, talking alone, and talking and writing in combination on science learning in the middle school. Moreover, since some types of instructional strategies have been found to be more effective with particular groups of students (Pressley, Goldchild, Fleet, Zajchowski, & Evans, 1989), both gender and ability were considered in the study to capture any possible aptitude-treatment interactions which might be addressed in a later study.

1.2.0 *Aptitude-Treatment Interactions*

The research designs which have been used in many studies of writing-to-learn in the classroom have not managed to isolate interactions that characterize the use of these strategies (Rivard, 1994). Some strategies may be more useful with particular groups of students. Race, ethnicity, academic ability and gender are all variables that can interact with the use of writing-to-learn strategies. Although ineffective with average or stronger students, a strategy may be quite effective with weaker students. If classrooms are to become multi-dimensional, accommodating different students with various methods and multiple tasks, then teachers should know what works and with whom.

Gender Differences

Gender differences in science achievement and in attitudes towards science or science learning have been noted in many studies (Department of Education and Science, 1985; Erickson & Farkas, 1991; Lapointe, Mead, & Phillips, 1989; Mullis & Jenkins, 1988; Science Council of Canada, 1984a; Tamir, 1989). Male students have generally outperformed females and have demonstrated more positive attitudes towards the study of science, whereas female students have tended to avoid secondary science courses altogether or to drop out after experiencing repeated failure or much frustration (Ferguson, 1984; Haggerty, 1991; Science Council of Canada, 1984a). Linn and Hyde (1989) have argued that "gender differences in cognitive and most social domains should be de-emphasized because they are small . . . [and are] so clearly a function of context or situation" (p. 26). However, other authors 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; Mason & Kahle, 1989; Rosser, 1990; Scantlebury & Kahle, 1993; Skolnick, Langbort & Day, 1982). For instance, Adams (1992) recommends the use of learning strategies like subjective knowing, sharing ideas, classroom dialogue, peer discussions, concrete experiences, journal activities and naturalistic inquiry. In a recent article, 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). The present study investigated the use of talking and writing as learning strategies with female students for learning science.

Ability Differences

Lehr and Harris (1988) have suggested that "teachers of low achievers need to present information in a variety of ways and build into the curriculum methods that students can use to process the information" (p. 36). Glatthorn (1991) has argued that an appropriate curriculum for low achievers would emphasize critical thinking, problem-solving, and conceptual understanding in a classroom environment which is rich in opportunities for communication, while relating school learning to the life experiences of these students whenever possible. Other classroom strategies which have been suggested for enhancing the learning of low-achievers have included peer-mediated instruction, small-group activities and cooperative learning (Cuban, 1989; Maheady, Harper & Mallette, 1991; Mastropieri & Scruggs, 1992; Thornburg, 1991). On the basis of his study of school leavers, Finn (1989) has suggested that "teaching practices that involve students in the learning process" might make school more appealing for some students (p. 122). A similar analysis led Wehlage, Smith and Lipman (1992) to recommend curricular reforms that focus on "problem-solving and the discovery of meaning, connections, and patterns as opposed to rote learning" (p. 60).

The literature suggests that classroom interventions which emphasize conceptual understanding, real-life applications, language use and small-group work may be particularly effective for enhancing the learning of marginal students: low-ability students, underachievers, and potential school leavers. Although the learning strategies which will serve as experimental treatments in this study would probably be effective with all students, ceiling effects with average- and high-ability students might limit improvements in learning compared to the results with low-ability students. For instance, Kirkpatrick and Pittendrigh (1984) had reported that explaining everday natural phenomena seemed to benefit average and poor students the most. Moreover, since one study has shown that low ability students have considerable difficulty applying science concepts (Department of Education and Science, 1985), analyzing the data for aptitude-treatment interactions related to student ability should show whether or not these learning strategies are particularly helpful for this group of students.

1.3 Research Questions and Hypotheses

Since the number of subjects that were made available for the classroom investigation was less than expected when the research was proposed, this study should be viewed as exploratory in nature. The ultimate purpose of the study was to serve as a template for new research. The analysis was oriented towards searching for trends in the data rather than seeking conclusions. Nonetheless, this pilot study was still guided by the following questions:

1. Does writing an explanation for a scientific phenomenon (the writing mode, or W) enhance learning more than a discussion with other students on the same topic, but without writing (the talking mode, or T)?
2. Does writing an explanation after a discussion with other students (the combined talking and writing mode, or TW) enhance learning more than individually writing an explanation of the same phenomenon, but without talking (W)?
3. Does writing an explanation after a discussion with other students (the TW mode), enhance learning more than a discussion of the same topic with other students (T)?
4. Is each of these treatments (T, W, TW) more effective than a control group which has been assigned descriptive and other simple content learning tasks (fill-in-the-blanks, true-or-false, etc.)?
5. Is there an interaction between these different groups using writing, and talking and gender?
6. Is there an interaction between these different groups using writing, and talking and student ability?

The following null hypotheses also guided this study:

1. There will be no significant difference ($p < .05$) in mean aggregate simple knowledge scores that can be attributed to:

a) treatment (T, W, TW, and C);

b) gender; and

c) interaction of treatment and gender.

2. 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.

3. There will be no significant difference ($p < .05$) in mean aggregate total knowledge scores that can be attributed to:

a) treatment;

b) gender; and

c) interaction of treatment and gender.

1.4 Definition of Key Terms and Variables

For the purpose of this study, key terms and variables were defined as follows:

- **Learning** - science knowledge as measured through various instruments and artifacts that were scored using two rubrics: (1) simple knowledge, or facts, terminology and basic concepts; and (2) integrated knowledge, or applications and explanations in which various concepts are interrelated or connected together.
- **Knowledge structure** - the concepts and relations among them as measured on a concept map which the student has constructed.
- **Phenomenon** - any natural event that is apparent to the senses and that can be described or explained using various science concepts.
- **Scientific Explanation** - an account of how or why a phenomenon occurs or occurred in the past that includes one or more factual or theoretical propositions.
- **Explanatory Task** - An elaborate problem based on a natural phenomenon that requires the student to write a detailed answer explaining how or why it occurs or

occurred in the past.

- **Discussion** - A learning strategy in which students work in small groups of three or four to talk about possible explanations for a problem task.
- **Writing** - a learning strategy in which students individually produce an expository text in response to an explanatory task.
- **Talking-and-Writing (TW) Mode** - A learning strategy in which students initially discuss an explanatory task in small peer groups before writing their individual responses to it.
- **Ability** - a variable combining the student's past achievement in science (grade in seventh-grade) with the teacher's prediction of the student's success in the grade eight science program.

1.5 Significance of the Proposed Study

1.5.0 *The Development of Knowledge*

Duschl (1994) has argued that many of the research studies now being published in science education journals may be theoretically or academically interesting, but without being helpful to classroom practitioners. He has stated that "what we need are more research studies that deal with Schwab's commonplaces of education—teachers, students, subject, and milieu" (p. 205). The present study answers this call for practical research: The research focusses on the use of learning strategies within the science classroom, the design is relatively unobtrusive, and the treatment is pedagogically sound.

Parker and Goodkin (1987) have suggested that the writing-across-the-curriculum movement in North America has "been concerned almost exclusively with writing, ignoring other uses of language, especially talking" (p. 12). This neglect has resulted in few studies investigating how talking and writing can influence classroom learning. This study will attempt to clarify the role of both of these modes on science learning. Moreover, Cazden (1986) had indicated that few interpretive research studies have focused on the links between classroom discourse and school content. One part of the present study will qualitatively examine the discourse, or talk, which occurs in small peer-groups in an attempt to relate it to content learning.

McCulley (1986) has argued that researchers "haven't done much to match types of writing (exposition, for instance) and their inherent logical and organizational structures with specific learning tasks" (p. 45). Explanation can be considered a product, a process or a social construction (Coleman, 1992). As a product, explanations can be assessed using traditional test measures. As a process, explaining involves the use of higher-order thinking by students. Finally, explanation as a social construction refers to the synergism which can occur when students interact during discussion while grappling with a problem task requiring an explanation. Schoenfeld (1989) has suggested that this synergism often gives rise to ideas beyond those which individual students contributed: "Ideas seem to emerge *de novo* in new contexts, but can also be seen to be *in the atmosphere* [italics added] if one looks at a broader, more enveloping context" (p. 79). On the one hand, an "explanation" can be considered a product or text characterized by a logical and organizational structure. On the other hand, "explaining" can be considered a specific learning task which invokes analytical thinking involving, not only conceptual, but relational knowledge of the content domain. As such, the task is well suited to the learning goal which is the production of an explanatory text. Moreover, the use of small discussion groups should enhance student understanding of the phenomena being considered in the problem.

1.5.1 *Implications for Educational Practices*

On the basis of an international assessment of science achievement in seventeen countries, the International Association for the Evaluation of Educational Achievement (1988) stated that "for a technologically advanced country, it would appear that a reexamination of how science is presented and studied is required" (p. 9). Although the report was specifically referring to the situation in the U.S., the same case could be made for science teaching in Canada. This study is important because it addresses the issue of "inert knowledge" in students by proposing an instructional strategy which can be easily implemented in science classrooms.

Survey studies have shown that teachers rarely use writing for learning content (Martin, D'Arcy, Newton, & Parker, 1976; Newell, 1986; Pearce, 1984; Sullenger, 1990 / 1991). When writing is used in the science classroom, it is generally for evaluating students rather than clarifying or refining their understanding of content (Tighe, 1991). Unless research can demonstrate how these writing strategies might be effectively implemented by classroom teachers, they likely never will adopt any of these strategies, no matter how promising these may be (Fowler, 1989/1990; Kamman, 1990). This study focuses on the use of one writing strategy, explaining, for enhancing student understanding in science.

Some educators have argued that using more extended writing can make science classes more appealing to females (Connolly & Vilardi, 1989; Smail, 1987). Since one national survey has shown that females generally like to write more than males, writing-to-learn strategies may be an effective teaching and learning strategy with this group (Coley, 1989). As the Science Council of Canada (1984a) has suggested: "In a world shaped by science and technology, Canada can no longer afford to shrug off the underrepresentation of girls in science classes in secondary schools" (p. 25). The study will attempt to isolate this variable to determine whether there exists an aptitude-treatment interaction between gender and the various treatments involving talking and writing.

Finally, the study is important because the instructional strategy may be particularly effective with low-achievers. Many of these low-achievers are potential school-leavers or drop-outs. According to 1991 data from Statistics Canada, 18% of twenty-year-olds (22% for males and 14% for females) still had not received a high school diploma (Gilbert & Orok, 1993). In a society which depends on an educated labour force for driving the economy and on a literate citizenship for participating in the political decision-making process, this is unacceptable. Since roughly one-fifth of these school leavers reported that they were bored with school, using alternative instructional strategies may be helpful in keeping them in school.

1.5.2 *Implications for Further Research*

The study was exploratory in nature. The fundamental objective was to develop and refine instruments and methods, and to formulate specific hypotheses and questions that could frame future research on the use of writing and talking as learning strategies in the science classroom. Moreover, analysis of the data should indicate whether or not a follow-up study specifically addressing aptitude-treatment interactions has any merit.

2. REVIEW OF THE LITERATURE

Since the purpose of the study was to investigate the role of talking and writing as strategies for learning science, the research literature in several unrelated fields of study, that are still relevant to the problem at hand, will be reviewed. First, the literature on talking and learning will be briefly discussed to establish the importance of informal language use in all human learning. Second, the literature underlying the notion of writing to learn, in which the act of writing itself can serve as a heuristic for "discovering meaning" or as a tool for enhancing learning will be described in considerable detail (Murray, 1982, p. 4). Third, certain writings in the philosophy of science will be explored to compare the prevailing conceptions of explanation, then some critical issues underlying the place of explanation in the classroom will be described, and the finally the link between explanation and understanding will be established. Fourth, organizational concerns about how to ensure that talk is meaningful and on-task while students work in small peer groups will be presented. Last, background information about concept mapping (definitions, origins, important aspects, role in the classroom and in research) will be described to justify its use as a measure in this study.

2.0 Talk and Learning

The relationship between talk, or oral language, and learning has been recognized by some educators since the early seventies (Barnes, 1976; Britton, 1970, 1982; Bruner, 1986; Martin, D'Arcy, Newton, & Parker, 1976). Barnes (1976) stated that talk is "a tool for making meaning" (p. 100). Rubin (1990) spoke of talk that can "transform knowledge" (p. 10), whereas Gere (1990) refered to "talk as the instrument by which meanings are negotiated and created" (p. 116). Fulwiler (1987) suggests that "we carry out conversations with others to explain things to ourselves" (p. 5). Britton (1982) characterized talk as an interpretive tool: "Talking . . . is the normal way in which we endeavour to make sense of our experiences, so that we store in memory not the raw data of events but the meaning we have come to attribute to them" (p. 115). Bruner and Haste (1987) have elaborated on Britton's views. They have identified three functions for human discourse: "The first is discourse as scaffolding, the second is discourse as the negotiation of meaning, and the third is discourse as the transfer of cultural representations" (p. 21). Talk as scaffolding, or for meaning-making or for knowledge acquisition, are all functional attributes that make talk particularly suitable for enhancing content understanding in the science classroom.

Britton (1982) has theorized about how understanding might be enhanced through talking:

> The two processes are interlocked: 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 also reflected on the underlying mechanisms that might explain the role of talk in learning. He 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). Finally, Lemke (1990) has specifically addressed the issue of language, not only talk, but other forms of language such as listening, reading, and writing, for supporting learning in the science classroom. He has stated that "the one single change in science teaching that should do more than any other to improve students' ability to use the language of science is to give them more practice actually using it. . . . Students should do more science writing during class, always following oral discussion of topics" (p. 168).

According to some educators, strategies involving the use of language have been seriously underutilized in the science classroom despite their theoretical potential (Driver, 1981 & 1989). In his review of research on student conceptions, Confrey

(1990) commented that "the role of language in the construction of understanding extends beyond labeling and communication of propositional knowledge into the social construction of knowledge" (p. 27). Pines (1985) also recognized the role of language in the construction of meanings: "A word is like a conceptual handle, enabling one to hold on to the concept and manipulate it" (p. 108). Barnes (1976) has written extensively about the role of language in exploring conceptual worlds. He suggests that "talk and writing provide the means by which children are able to reflect upon the bases upon which they are interpreting reality, and thereby change them" (p. 31). The Bullock Report also confirmed this view: "Language has a heuristic function; that is to say a child can learn by talking and writing as certainly as he can by listening and reading" (Department of Education and Science, 1975, p. 50).

Stenhouse (1986) has used Wittgenstein's concept of "language-game" to equate conceptual understanding with language use. He has argued that "we judge whether or not a student understands a concept by whether or not he or she can correctly use the words (or symbols) relating to that concept" (p. 417). He has stated that science teachers must encourage students to participate in this language-game by interpreting, or re-formulating, traditional scientific language found in textbooks and teacher explanations. He suggests that to do so "is likely both to be effective in the content/cognitive dimension and also to ameliorate the affective and social dimensions of the educational relationship [between teacher and students]" (p. 422). Pea (1993) has also argued that learning scientific concepts must include "inquiry and sense-making conversations involving authentic tasks in science practice such as making predictions, designing experiments to test them, careful observations, explanations, and revising conjectures in light of their observations" (p. 273).

Several empirical studies have confirmed the importance of language in classrooms. Driver (1989) reported on an ethnographic study which shows that conceptual-change classrooms spend more time on planning, discussion, reading and writing than traditional classrooms. She also observed that the participation of girls in class discussions, as well as their questioning behaviour, were far greater in these classrooms than that observed in traditional classrooms. Roth, Anderson, and Smith (1987) have conducted case studies investigating how different teachers approached a fifth grade unit on light and seeing. A year later, they studied how the use of specially prepared curriculum materials, which reflected a conceptual change perspective, had changed the patterns of classroom talk. The researchers observed major differences in classroom talk between conceptual change classrooms and traditional classrooms. In conceptual change classrooms, teachers used talk for: "(1) eliciting and responding to student misconceptions; . . . (2) focusing on explanations; . . . (3) probing after student responses; . . . (4) balancing open-ended and closed discussions; . . . and (5) providing practice and application" (pp. 544-546). Although talking in traditional classrooms has generally been viewed as unnecessary, it would appear to be indispensable in conceptual change classrooms where the students must negotiate meanings with peers while constructing personal understandings.

2.1 Writing and Learning

2.1.0 *Educator Concerns*

Some educators have suggested that writing might be used to enhance learning, whereas others have argued that writing is intimately related to thinking (Applebee, 1984a; Emig, 1977; Gere, 1985; J. Howard, 1983; Knoblauch & Brannon, 1983; Langer, 1986a; Langer & Applebee, 1987; Newell, 1984, 1986; Odell, 1980; Sensenbaugh, 1989). V. A. Howard (1988) suggested that too much emphasis has been placed on writing as communication and not enough on writing as articulation—what he has called "thinking on paper." He argued that the "act of writing is father to thought itself" so that "writing serves understanding first and communication second" (p. 88). The process of writing is important, not only for learning about something or acquiring knowledge but for generating a personal response to something, for clarifying ideas and for constructing knowledge.

Resnick (1987) recognized that writing has been neglected for too long and suggested that:

> [Writing's] potential role as a cultivator and an enabler of higher order thinking is very great, especially if we consider writing as an occasion to think through arguments and to master forms of reasoning and persuasion that are valued in various disciplines. (p. 38)

This current emphasis on a thinking curriculum is also undergirded by a recognition that discrete knowledge should not be learned for its own sake, as is too often the case in science classrooms (Hurd, 1991). Rather, students should be challenged to use this knowledge in solving meaningful problems (Resnick & Klopfer, 1989). In this learning environment, writing can serve as the medium for fleshing out responses to complex problems requiring higher-order reasoning.

Newell (1983) argued that writing can be a "powerful heuristic for learning new information when the writing is done for genuine communicative purposes and when the writer attempts to integrate the new information with prior knowledge" (p. 10). In V. A. Howard's view (1988), this writing as communication phase should follow an initial exploratory phase in which writers have articulated their ideas with few rhetorical constraints to inhibit their thinking. Applebee (1984a) theorized about the link between writing and thinking, arguing that the act of writing, by its very nature, may enhance thinking. He suggested that the permanence of the written word, the explicitness required for effective written communication, the richness of discursive tools for refining ideas and the active nature of writing all converge to make writing an extremely powerful process for shaping thought. Hayes (1987) utilized the organizing-demands hypothesis to explain how this might work: "Writing may achieve its effects by the demands it places on learners to organize language" (p. 334).

2.1.1 *Teacher Practices*

Britton, Burgess, Martin, McLeod and Rosen (1975) reported that British students do more writing in content courses than in language courses. Donlan (1974) confirmed that American students also do a lot of writing in content classrooms. However, Mullis and Jenkins (1988) wrote that nearly half of all grade eleven students participating in the National Assessment of Educational Progress' 1986 national assessment in science reported never writing reports of any kind (52%). Moreover, several studies suggest that this situation is no different in Canada. Fillion (1979) reported that Toronto secondary students wrote very little in science class without being guided by either teachers or textbooks. Another study conducted in Toronto by McTeague, Payne, Graham and Murray (1981) confirmed this limited use of student writing in Canadian classrooms.

Most classroom writing is informational and is intended solely for the teacher in the role of examiner with much of the writing involving simple mechanical tasks like filling in blanks (Applebee, 1984b). Rarely is expressive or persuasive writing used in the classroom. The use of writing which is extended or directed to oneself or to a wider audience of classmates or other readers also appears to be limited (Anson, 1988; Applebee, 1981, 1984b; Tighe, 1991). Students rarely use writing for performing meaningful authentic learning tasks in which they must write for a specific purpose with a real audience in mind. Nor do they appear to use writing for learning or for clarifying their own ideas about science topics (Lloyd, 1990; Pearce, 1984).

Langer and Applebee (1987) noted that there were clear differences in the use of writing between social studies and science teachers. Beliefs regarding teacher and student roles in the classroom seemed to determine whether writing would be used to provide information, foster inquiry, facilitate organization, develop understanding or mostly to evaluate students. Science teachers generally tended to use writing for evaluative purposes while social studies teachers reported more writing assignments which encouraged students to extend their learning. Based on case studies of writing in science and social studies classes, Marshall (1984) observed that, when the product is given more importance than the process, students tend to disregard writing, even those tasks which are designed to elicit personal responses or to enable them to make connections with prior knowledge. The way that writing is employed and evaluated in the classroom would appear to be critical in determining students' perceptions of its potential for learning content.

Sullenger (1990/1991) observed that the science teachers in her study generally expressed little concern for writing in the classroom. Pearce (1984) has suggested that using more writing in the science classroom might be achieved by showing teachers how to use writing-to-learn strategies.

2.1.2 *Writing, Learning, and Thinking*

A number of seminal studies have suggested that writing reflects the thought processes of the writer (Applebee, 1984a; Durst, 1987; Hayes, 1987; Langer, 1986b; Langer & Applebee, 1987; Newell, 1984). Although some of these studies have been conducted in disciplines other than science or have involved small numbers of subjects, they provide compelling evidence for the potential of writing to learn in the science classroom.

On the basis of case studies of 67 children from 8 to 14 years of age, Langer (1986b) concluded that children appear to be more aware of their use of strategies, rhetorical structures and background knowledge while writing than while reading. She suggested that this may be because strategic and text-structure usage are not as readily apparent to readers as they are to writers. She argued that this metacognitive awareness may make writing a particularly effective tool for writing to learn.

Newell (1984) studied the interactions between writing tasks and conceptual learning in eight eleventh-grade students, all above average in reading and writing abilities, as they worked with passages selected from science and social studies textbooks. The writing tasks included taking notes, answering study questions, and writing an analytical essay (applying concepts to a new situation), while learning was measured using a recall task, a measure of passage-specific knowledge and application questions. Although the results suggested that the type of writing task had no effect on either recall or application of the concepts, there was a significant effect for the essay-writing task on passage-specific knowledge. Analysis of the think-aloud protocols and the writing samples also suggested that essay writing resulted in more writing and learning operations (e.g., elaborating, interpreting and hypothesizing).

Newell (1986) also studied eight eleventh-grade students using research tools, materials and measures similar to those used in his earlier study. Newell concluded that the type of learning reflects the particular writing task engaged in by the learner. He suggested that with note-taking and question-answering the learner processes discrete bits of knowledge without making significant connections among them. He concluded that analytic essay writing encouraged the learners to integrate new information with relevant prior topical knowledge.

Hayes (1987) investigated the effect of combining different writing tasks with reading on the hierarchical organization of information recalled later. This recall measure allowed him to compare the proportion of top-level and low-level information included by students. The sample consisted of 176 tenth-grade students, average to above average readers, who read a target passage selected from a science textbook and completed a pretreatment recall measure on it. Students then read other passages on related topics, reacting to them using one of three assigned writing tasks which consisted of paraphrasing, formulating question or comparing and contrasting. Finally, students completed a posttreatment recall test on the initial target passage. Although the quantity of information recalled was not significantly different for any of the treatment groups, students who were either formulating questions or comparing and contrasting appeared to make more connections among the isolated bits of information. In addition, the group formulating questions tended to include more superordinate information in their recalls compared to all other treatment groups. Hayes concluded that the "way writing is combined with reading differentially impacts students' remembering a topic under study" (p. 346).

Durst (1987) investigated the cognitive and linguistic demands of writing summaries compared to writing analytical essays in ten high-ability and ten average-ability eleventh-grade history students. The results suggested that summary writers essentially engaged in low-level planning operations for restating content, while analytic writers engaged in more high-level planning and questioning, more evaluating and more thinking about the task while they construct meaning from the written text. Marshall (1987) observed that personal analytic writing in which students include personal observations in their responses to literature was as powerful as formal analytic writing in which students developed arguments solely from the text. Moreover, both of these forms of extended writing were significantly better than restricted writing (short-answer questions) or no writing at all on text recall and interpretation. Durst (1989) later observed that analytic writing encouraged students to devote more time to comprehending the demands of the writing task, to understanding the underlying content and to monitoring their use of writing strategies.

Langer and Applebee (1987) reported on a series of three related studies investigating the effects of writing on learning and thinking. Students in grades nine and eleven were assigned different writing tasks: Some involved limited writing (answering study questions or taking notes), others more extensive writing (writing summaries or analytic essays) to learn concepts from social studies passages. They observed that students focus on discrete bits of information when answering study questions. With note-taking, students focused on larger chunks but only superficially, whereas they appeared to integrate the information while engaging in more complex thought processes when writing essays. The use of a reading-only control group in the third study suggested that almost any kind of writing is better than simply reading for learning from texts and that more extended writing tasks, like summaries and analytical essays, result in better text recall. In summarizing the role of writing on learning, Langer and Applebee (1987) concluded that:

> First, the more that content is manipulated, the more likely it is to be remembered and understood . . . . Second, the effects of writing tasks are greatest for the particular information focused upon during the writing . . . . Third, writing tasks differ in the breadth of information drawn upon and in the depth of processing of that information that they invoke . . . . Finally, if content is familiar and relationships are well understood, writing may have no major effect at all. (pp. 130-131)

The findings of these studies have generally been interpreted using a "depth-of-processing hypothesis" which suggests that different writing tasks invoke different cognitive strategies for processing and encoding information (Langer & Applebee, 1987, p. 92). This encoded information can be recalled later in response to particular questions requiring simple comprehension, interpretation or application on the part of the learner. As Langer and Applebee have suggested:

> writing tasks differ in the breadth of information drawn upon and in the depth of processing of that information that they invoke. Thus note-taking, comprehension questions, and summarization tasks, which focus attention across a text as a whole, have relatively generalized effects, though they lead to relatively superficial manipulation of the material being reviewed . . . . Analytic-writing tasks, on the other hand, focus the writer more narrowly on a specific body of information . . . . This attention is also more directly focused on the relationships that give structure and coherence to that information. (p. 131)

Langer and Applebee (1987) have identified three functions for writing to learn in the content classroom:

(1) to draw on relevant knowledge and experience in preparation for new activities;

(2) to consolidate and review new information and experiences; [and]

(3) to reformulate and extend knowledge. (p. 41)

Langer and Applebee (1987) have given examples of the kinds of writing tasks that might be used for invoking these three functions of writing to learn. Freewriting is suggested as a classroom activity for stimulating students to draw on relevant knowledge and experience. Journal writing, summarizing, note-taking and answering study questions all are proposed as activities for consolidating and reviewing new information and experiences. Finally, they suggest an activity in which students "create-an-animal" as a writing task in which students reformulate and extend knowledge. However, other learning tasks in which students analyze, synthesize, evaluate or apply knowledge would likely meet the demands of this third function.

2.1.3 *Writing to Learn in Science*

Since the literature on the use of writing to learn in science has been recently reviewed by Rivard (1994), only those studies which are relevant to the dissertation will be discussed here. Moreover, studies that have been conducted in other content areas, but which still merit serious consideration by all science educators, were also reviewed. The focus in this section will be on the use of writing for reformulating and extending knowledge that might include tasks like analysing, interpreting, explaining, and evaluating. This is one of Langer and Applebee's (1987) three functions for writing to learn in the content areas. These writing and thinking tasks have been underutilized in the science classroom (Mullis & Jenkins, 1988).

Analytic Writing

Applebee, Langer, and Mullis (1986) indicated that "even at grade 11, relatively few students were able to provide adequate responses to analytic writing tasks" (p. 11). Despite their reported difficulties with analytic writing tasks, students appear to learn more while using them. The use of analytic writing in the science classroom is also supported in a study by Weiss and Walters (1980) who reported that "written concepts were learned with more clarity than non-written concepts" (p. 4).

Newell has examined the use of answering study questions, notetaking and writing analytic essays for learning course content in science and social studies. The writing task used in one study clearly involved explaining: "Explain in an essay how a spectroscope could be used to study the elements of the sun" (Newell, 1986, p. 298). In this study with eight eleventh-grade students, he observed that explaining enhanced content learning more than notetaking or answering study questions.

In an earlier study with eight eleventh-grade students, all strong in reading and writing abilities, Newell (1984) had employed an analytic essay task which "required [the] application of concepts to solve new problems" (p. 268). This writing task again involved the use of explaining. Newell (1986) has also stated that "with essay writing . . . we can focus on explanation and interpretation of concepts and principles that we want students to integrate into their prior learning" (p. 300). Langer and Applebee (1987) also used a writing task that required the application of new information. They indicated that this task required "the same type of writing as the analytic-writing study condition" (p. 106).

Newell (1984, 1986) has not differentiated between analytical writing in science and analytical writing in social studies in his studies of writing and learning. He did, however, recognize the need "to examine variations within function in different disciplines" (Durst & Newell, 1989). The role of explaining as a variant of analytic writing still merits further study.

Explaining

The use of short explanatory writing tasks for enhancing learning has been recommended by some educators (Ammon & Ammon, 1990; Comley, Hamilton, Klaus, Scholes, & Sommers, 1984). Van Nostrand (1979) has argued that writing can enhance learning because it involves "joining bits of information into relationships, many of which have never existed until the composer utters them" (p. 178). Jensen (1987) has argued that "writing about the concept and explaining it to others helps one understand that concept" (p. 330). Strenski (1984) has suggested using writing tasks in which students write a paragraph from a set of facts. He reported that this technique, which required making connections among concepts, helped the secondary students link the new topic being studied with their prior knowledge about it. Kirkpatrick and Pittendrigh (1984) used focused essay questions which required college students to explain everyday natural phenomena. Although this writing task benefitted average and poor students the most, almost 90% of the students participating in the project reported that the writing assignments had enhanced their learning of physical science concepts. Expository writing tasks in which students explain scientific phenomena were also found to be effective in college chemistry (Thall & Bays, 1989; VanOrden, 1987). Beyer (1982) has indicated how explaining can enhance the learning process:

> "by engaging them [students] in the basic steps of the conceptualizing processes—identifying examples of a given concept, explicating attributes common to all examples, and identifying the general interrelationships of these attributes—as this sequence of writing activities does, students can use writing as a tool for accomplishing this task" (p. 102).

Scardamalia and Bereiter (1986) have suggested that this epistemic function of writing "presuppose[s] that the writer is engaged in active reprocessing at the level of concepts and central ideas" (p. 790). Using writing to learn in the classroom also receives support from Shanahan (1988) who recommends having "students write explanations of complex text in order to enhance their understanding and to increase their awareness of their own level of understanding" (p. 645).

Schumacher and Nash (1991) have recommended the use of writing tasks which involve explaining for bringing about conceptual change. They have suggested the use of:

(1) writing tasks which increase the likelihood of the writer experiencing anomalies; (p. 77)

(2) writing tasks which necessitate dealing directly with analogical or metaphorical content; (p. 79)

(3) [writing tasks which favour] the construction of multiple representations; (p. 79) and

(4) writing [tasks which necessitate] the creation of a wide variety of exemplars. (p. 83)

In justifying the role of writing for learning from analogies, they argue that "the very process, for example, of having to explain, extend, or criticize an analogy for a content area may prove particularly effective in bringing about new understanding in that area" (p. 79). This notion is further supported by McGinley and Tierney (1989) who have argued that cases, or various problems applying particular scientific concepts, provide the "means or 'routes' for 'traversing a topical landscape'" (p. 249) and that "knowledge is best acquired by "traversing [this topical landscape] from a variety of perspectives" (p. 250).

Explaining, interpreting and applying scientific information can be considered science process skills as important as observing, classifying or planning investigations (Kempa, 1986). Kempa has described several kinds of tasks which involve explanatory writing:

(1) explaining familiar facts, observations and phenomena in terms of scientific laws, theories and models;

(2) [recognizing] patterns and relationships from experimental or other data;

(3) suggesting scientific explanations of unfamiliar facts, observations and phenomena; and

(4) applying scientific ideas and procedures to solve qualitative and quantitative problems. (p. 55)

Using explanatory writing tasks in the science classroom might be a useful strategy for encouraging students to link concepts thereby enhancing content understanding (Dyson & Freedman, 1991). The strategy also dovetails nicely with the current emphasis on constructivist approaches for the teaching of science. Yager (1991) has identified "proposing explanations and solutions" as an appropriate constructivist strategy for science teaching (p. 55). The use of this strategy in the science classroom certainly merits further study. Moreover, whether or not explanatory writing tasks benefit certain groups of students, for instance poor achievers, makes the research particularly relevant for elucidating effective classroom strategies tailored to at-risk students.

2.2 Explanation and Understanding

2.2.0 *The Role of Explanation*

The National Center for Improving Science Education has suggested that "searching for causes and explanations is the major activity of science" (1989, p.14). This explanatory view of science is echoed by the recent *Project 2061 Report* which will orient American science education through the next several decades:

> Scientists strive to make sense of observations of phenomena by inventing explanations for them that use, or are consistent with, currently accepted scientific principles. Such explanations—theories—may be either sweeping or restricted, but they must be logically sound and incorporate a significant body of scientifically valid observations. The credibility of scientific theories often comes from their ability to show relationships among phenomena that previously seemed unrelated. (Rutherford & Ahlgren, 1990, p.7)

However, the concept of explanation is itself problematic. Although early Greek philosophers, and more recently, contemporary philosophers of science have all argued about the different conceptions of explanation, a concensus has never really been achieved on the nature of scientific explanation (Thagard, 1989). Despite this uncertainty, people use scientific explanations on a daily basis. Since scientific literacy requires abilities in dealing with scientific information, it therefore merits serious consideration, particularly by teachers who work with scientific information and who have a responsibility for developing scientific literacy in students.

2.2.1 *The Nature of Scientific Explanations*

Kourany (1987) has described three different conceptions of scientific explanation reflected in modern writings in the philosophy of science. The *inferential conception* views explanation as a logical argument consisting of a series of statements, premises or propositions. The *erotetic conception* sees explanation as an answer to a why-question. Finally, the *causal conception* views the explanation of a phenomenon as "laying bare its inner workings, its underlying causal mechanisms" (Kourany, 1987, p. 24).

Inferential Conception

Hempel (1965, 1987) is undoubtedly the most important proponent of the inferential conception of scientific explanation. In his view, scientific explanations are comprised of laws and conditions, with the latter describing the context for the phenomenon to be explained. These statements, taken together, make up the *explanans* of the explanation. The rules of logic are used to judge the merit of the argument, or explanation, with deductive arguments generally being considered stronger than inductive arguments. In a deductive argument, the explanandum or description of the phenomenon being explained, is a logical inference of the explanans, or propositions, which comprise the explanation. The notion of empirical validation by observation or experimentation is important to the Hempelian notion of explanation.

Erotetic Conception

The erotetic conception of scientific explanation involves answers to why-questions (Bromberger, 1966). The erotetic conception can also be considered a pragmatic view in that context is all-important in determining whether or not an explanation given in response to a question is adequate (van Fraassen, 1987). According to this view, both the form and the content of a particular explanation will depend on the situational context. The importance of context in establishing explanatory relevance also receives support from other scholars. For instance, Braithwaite (1953) has argued that explanations, by their very nature, are hierarchical in that one can always search for a higher-level explanation to account for an observed events. He suggests that:

> At each stage of explanation a 'Why?' question can significantly be asked of the explanatory hypotheses; there is no ultimate end to the hierarchy of scientific explanation, and thus no completely final explanation. (p. 347)

Weaver (1964) has coined the term 'vertical explanation' to represent this situation because explanations can "move vertically downward to deeper levels of simplicity or abstraction" (p. 1299).

Causal Conception

Salmon (1987) has explicated the causal conception of scientific explanation. He has suggested that the "explanation of a phenomenon essentially involves locating and identifying its cause or causes" (p. 52). In this view, natural events are seen as nodes, or interactions, in a web of causal processes, with one or more antecedent events, the causes, accounting for the occurrence of another event, the effect. According to the causal conception, explanation involves describing the relevant causal connection or connections between the event to be explained and those events leading to it. The causal conception underscores the intimate relationship which exists between the two fundamental aims of science: explanation and prediction.

Although it is commonly assumed that good scientific explanations should also possess good predictive power, this may not be necessary. On the one hand, the physical sciences appear to be quite adept at predicting natural events. However, the underlying "explanations" in the physical sciences may actually be simply more adequate descriptions, or descriptions of a more general phenomena which can subsume the occurence of the event. The biological sciences, on the other hand, would appear to have minimal predictive power, but good explanations of the causal connections between events.

Biological Explanations

The nature of scientific explanations can differ depending on whether they are situated in the physical or biological sciences. Explanations in the physical sciences are generally based on universal laws, whereas those in the biological sciences are more likely to be based on probabilistic laws, if they are based on laws at all. Moreover, Von Wright (1971) has stated that "functional explanations in biology . . . are typically quasi-teleological" (p. 85).

Teleological explanations have been described by Hempel and Oppenheim (1987, p.36) as "causal explanations in which some of the antecedent conditions are motives of the agent whose actions are to be explained." They also suggest that teleological explanations are a "facile construction of *ex post facto* accounts without predictive force" (p.36). For instance, mimicry might be explained by suggesting that an animal evolved protective coloration to thwart its predators' attempts to detect it, thus ensuring the preservation of the species. Braithwaite (1953) has differentiated between causal explanations and teleological explanations:

> In a causal explanation the *explicandum* is explained in terms of a cause which either precedes or is simultaneous with it: in a teleological explanation the explicandum is explained as being causally related either to a particular goal in the future or to a biological end which is as much future as present or past. (p. 324)

Baier (1980) has argued that teleological explanations are really answers to why-questions, the erotetic conception of scientific explanation, whereas causal explanations are generally answers to how-questions. Nonetheless, Hempel and Oppenheim (1987) suggest that teleological explanations may still serve a useful function as a heuristic device for furthering biological research. Teleological explanations also may serve an important psychological function because they make the explanation more acceptable by assigning purposes which are readily comprehensible to human experience. Moreover, in many instances the teleological nature of the explanation is embedded in the language used to express it. Simple reformulation of the explanation can often remove this teleological veneer.

Classifying Explanations

Kourany (1987) has proposed a useful method for classifying scientific explanations according to four basic types: *Compositional, evolutionary, functional* and *transitional* explanations. First, compositional explanations rely on explaining "properties of objects . . . in terms of the properties of their parts—their composition" (p.26). For instance, the shape of a protein, a property, might be explained in terms of the polarities of its constituent parts, the amino acids. Second, evolutionary explanations explain "properties of objects . . . in terms of the temporal development of these objects—their evolution" (p.27). The following explanation, which is taken from a science magazine article, is a good example of the evolutionary type.

> The scenario goes like this. An avian flu virus passes from migratory birds and waterfowl, particularly ducks, to pigs. At the same time, human flu virus passes from the farmer to the pig, which becomes a "mixing vessel." (Pigs and humans, in fact, overlap considerably in the viruses each is susceptible to.) Once inside pigs the human and avian viruses exchange genes that code for key surface proteins on the virus called hemaglutinins, which trigger the immune response in humans. After what Morse calls " a trial run" in the local pigs, a new virus emerges in a form that can be passed back to humans. This variant might be more lethal than the original human virus, and because humans have never been exposed to it, they lack specific immunity - and a global pandemic catches fire. (Langone, 1990, p.66)

The author of this article has explained how human intervention can inadvertently enhance the emergence of new influenza strains, thus accelerating viral evolution.

Third, functional explanations rely on explaining capacities in terms of the organization of simpler parts. The following explanation, taken from the textbook *Biology* by Slesnick (1985), could be considered functional in nature.

> Leaves play an important role in the movement of water over the longer distance from the roots to the top of a plant. During daylight hours a leaf's stomata are open. Evaporation of water takes place rapidly through these openings in a process called transpiration. The xylem tubes leading to the stomata are filled with water, creating a continuous column of water from the roots to the leaf tips. As water evaporates through the stomata, more water moves upward to replace the water that is lost . . . . Through this process, enormous amounts of water move through large trees every day. (p. 358)

Thus, the transpiration-cohesion theory is used to explain the movement of water through trees. This capacity for moving water is explained by describing the structure and function of simpler parts like roots, xylem tubes and leaves.

Finally, in transitional explanations a "change of state in an object . . . is explained in terms of a disturbance in the object and the state of the object at the time of the disturbance" (Kourany, 1987, p. 28). An article from *Science*, which described the role of parasites in mediating the population characteristics of mosquitos, includes the following transitional explanation:

> In high-density microcosms without parasites, population densities and resource competition remained high, resulting in the production of smaller adults. Within low-density populations, adults from both treatment and control microcosms were of similar size because food was much less limiting. (Pechmann, Scott, Semlitsch, Caldwell, Vitt, & Gibbons, 1991, p. 51
>
> 893)

The presence of parasites disturbed the ecological balance thus affecting characteristics of the host mosquito population.

Although theoretically cogent, Kourany's typology is difficult to apply to authentic scientific writing. Compositional and functional explanations are not always easily differentiated, whereas transitional and evolutionary explanations appear to overlap extensively. However, these four types can be easily collapsed into two with compositional and evolutionary explanations subsuming the other two according to whether they are concerned, respectively, with stasis or with change.

2.2.2 *Explanation and Understanding*

Antaki (1988) suggests that: "If there is a feature that all explanations reliably possess, it is that they reveal something of the explainer's mind to the enquirer" (p. 6-7). In a similar vein, Klemke, Hollinger and Kline (1980, p.84) indicated that "the pragmatics of explanation concerns the important considerations in explaining something to someone." Martin (1970), who shares this last viewpoint, distinguishes between "explaining something," which is related to scientific research and inquiry, and "explaining something to someone," which is instead related to teaching and learning (p. 16). She argues that explaining has not been emphasized enough in the classroom for fostering student learning. She notes that a person who explains "might even find himself gaining understanding, for it is well known that there is something about having to organize and formulate one's ideas about some topic for another which leads to new sorts of enlightenment for oneself" (p. 208). Martin's notion of understanding seems to implicitly acknowledge the causal conception of explanation. She writes that "understanding is not simply a matter of seeing connections: One must . . . see the things to be connected as well as the connections between them if one is to understand" (p. 167). She recommends the use of "genuine explaining episodes" to enhance conceptual understanding in the classroom (p. 216).

The role of explanation for enhancing conceptual understanding has received strong support from many scholars, both from within and beyond the science education community. Brown and Campione (1990) have suggested that "the burden of explanation is often the push needed to make students evaluate, integrate and elaborate knowledge in new ways" (p. 114). This notion is further supported by studies in experimental psychology involving elaborative interrogation, a strategy in which learners use 'why' questions to draw upon prior knowledge, that seems to improve the learning of facts (Wood, Fler, & Willoughby, 1992).

More recently, Cohen, McLaughlin, and Talbert (1993) have argued that schools must "move away from transmitting knowledge and facts to promoting students' deeper understanding of academic subjects—understanding based in active engagement with subject matter concepts" (p. xi). Perkins and Blythe (1994) suggest that "understanding is a matter of being able to do a variety of thought-demanding things with a topic—like explaining, finding evidence and examples, generalizing, applying, analogizing, and representing the topic in a new way" (p. 5-6). Finally, Vitale, Romance, Parke and Widergren (1994) state that teaching for conceptual understanding should enhance the student's ability:

(1) to use science concepts as a guide to observe relevant aspects of their environment;

(2) to use concept relationships as a basis to generate predictions about future events;

(3) to use concept relationships as a basis to analyze and explain events that occur;

(4) to link concepts to other concepts in constructing knowledge;

(5) to demonstrate an awareness of using existing conceptual knowledge to learn new conceptual knowledge; and

(6) to communicate concepts and concept relationships to others. (pp. 20-21)

These standards suggest ways in which an explanatory task might be operationalized in the classroom.

A qualitative study which investigated how writing can be used to monitor conceptual change in the science classroom also provides further evidence for this link between explanation and understanding. Fellows (1994) used instructional strategies, such as oral and written student explanations, small-group and whole-class discussions, and problem-solving situations which applied some of the target concepts. She observed that "understanding . . . appeared in students' writing more often when students had opportunities to explain their ideas orally and in writing" (p. 985).

Brophy and Alleman (1991) described five attributes of teaching for understanding and application of knowledge:

(1) content coverage is intentionally limited so as to favour in-depth study;

(2) the content possesses a coherent knowledge structure centered on important ideas;

(3) the instruction focuses on the relationships and connections among these ideas;

(4) the activities emphasize processing information as students talk to negotiate meaning and to develop personal ownership over the ideas;

(5) opportunites for communicating, problem-solving, decision-making and creative or critical thinking are provided.

Although most of these attributes are in the cognitive realm, the last two acknowledge the social nature underlying all learning: Students individually construct their knowledge within a particular social context, the science classroom, while talking about content with the teacher and peers (Newman, Griffin, & Cole, 1989; O'Loughlin, 1991; Prawat, 1989a).

Understanding as a Social Act

Hatano and Inagaki (1991) have referred to learning within a social context as a "collective invention of knowledge" (pp. 332-333). This notion of student learning is far removed from the traditional science classroom which generally still subscribes to a transmission-absorption model of science teaching-learning (Prawat, 1989b). Moreover, the social-interactionist viewpoint is further supported, not only by work in education, but also by studies in linguistics, psychology, sociology and anthropology (Brown, Collins, & Duguid, 1989; Collins, Brown, & Newman, 1989; Edwards & Mercer, 1987; Halliday, 1975; Hatano & Inagaki, 1991; Newman, Griffin, & Cole, 1984; Palincsar, 1986; Resnick, 1991; Rogoff, 1984; Vygotsky, 1962; Wertsch, 1991). Glaser (1991) succinctly describes social-interactionist thinking: "Students are viewed as constructors of meaning and sense-makers who are greatly influenced by the context in which they work to learn" (p. 143).

Solomon (1987) has indicated how scientific explanations might be the object of social construction:

> The essential criterion is no longer the internal logic of the explanation but that it should be recognized and shared with others. We take it for granted that those who are close to us see the world as we do, but, through social exchanges, we seek always to have this reconfirmed. (p. 67)

The literature suggests that talk among students in the classroom is not only useful but essential for ensuring a good understanding of science concepts. Without this social interaction for sharing, refining and creating knowledge, students are left with inert ideas: adequate for responding to simple questions, but which cannot be applied to novel situations or problems.

2.3 Classroom Discussions

The literature appears to strongly support the use of talk, or discussion, in school classrooms (Costa, 1990; Dillon, 1988; Gall & Gall, 1990; Wilen, 1990). Barnes (1976) had advocated a teaching approach which is interactive and strongly discussion-based. Yet, the manner in which these discussions are actually implemented in the classroom would appear to be a critical determinant of their effectiveness. Studies suggest that classroom discourse is still firmly entrenched in the recitation tradition with the teacher initiating discussion through questioning, students responding to these questions, and finally the teacher evaluating the student responses (Cazden, 1986). This pattern of initiation-response-evaluation (I-R-E) utilizes questions as evaluative or control devices rather than as meaning-making adjuncts or cognitive scaffolding. O'Flahavan, Hartman, and Pearson (1988) have also observed that teachers tend to monopolize the flow of discussion in a whole-class situation.

2.3.0 *Whole-Class Discussions*

Bridges (1990) has outlined four characteristics of 'true' discussions:

(1) Students talk to each other (and not just in dialogue with the teacher);

(2) Students listen to each other;

(3) Students are responsive in thought and word to what others say; and . . .

(4) The talk is purposeful in relation to the development of understanding on the topic under discussion. (p. 106)

True discussions do not occur often in science classes. In a naturalistic study of whole-class discussions in content classrooms, Alvermann, O'Brien, and Dillon (1990) observed few "open forums, in which students present multiple perspectives, interact with one another as well as with the teacher, and use discourse longer than the one- or two-word responses associated with recitation" (p. 308). The balance of the discussions were classified as either recitation or lecture-recitation. Furthermore, on the basis of a large-scale study of U. S. classrooms, Goodlad (1984) reported that the frequency of discussions observed by the research team during visits to schools ranged from four to eight percent depending on the particular grade level (p. 107).

Nonetheless, learning strategies involving discussion have been suggested by some science educators. Champagne, Gunstone and Klopfer (1985) have proposed several instructional strategies that involve discussion, or what they have termed "interactive dialogue," for enhancing conceptual change (p. 71). One variant of interactive dialogue is "ideational confrontation" in which students interact with peers in reflecting on their own, as well as on other conceptions (p. 77). The emphasis in this strategy is on justifying one's thinking, resolving contradictions, and constructing ideas that will stand up to peer criticism (Champagne, Gunstone, & Klopfer, 1983). The effectiveness of such an approach has been demonstrated in a number of studies (Gunstone, Gray, & Searle, 1992; Hynd, Qian, Ridgeway, & Pickle, 1991). Gunstone, Gray, and Searle (1992) showed that a conceptual-change instructional model involving peer discussion based on Barnes' communicative approach and use of a Predict-Observe-Explain teaching strategy engendered long-term effects that were still measurable a year later in high school physics. Those students in conceptual change classrooms ranked higher in the study of mechanics than other students who had received traditional instruction.

The "discussion web" is an instructional strategy which has been developed by Alvermann (1991) that focuses the reader on a reflective question opposing the scientific concept with typical common misconceptions (p. 92). After reading a text, students work in pairs to identify arguments for or against their beliefs, then these dyads are combined into groups of four students with the task of defending their viewpoints using evidence from the text. Finally, a whole-class discussion ensues in which the teacher takes a more active role in clarifying views and refuting lingering misconceptions while using the text as the primary evidential source (Alvermann, Dillon, & O'Brien, 1987). More recently, Hynd, McWhorter, Phares and Suttles (1994) demonstrated that discussing ideas in groups mediated the effect of other variables such as reading a refutational text or viewing a demonstration, both of which challenged nonscientific intuitive notions about motion.

2.3.1 *Small-Group Discussions*

Whole-class discussions may actually be less effective in enhancing learning than small-group discussions with one's peers. Britton (1982) has suggested that:

> The normal procedure is to draw out connexions by *talking* about the new information; and talk with an expert who does not understand our ignorance may be less helpful than talking with others who share our experiences and share our ignorance. (p. 115)

On the basis of many studies of small-group discussion, Webb (1989) concluded that explaining, clarifying and elaborating one's thinking to peers enhances learning more than simpler content tasks which emphasize low-level knowledge.

Glaser (1991) has explained why small group discussions can enhance student understanding:

> First, it extends the available knowledge and thereby supports alternative problem approaches and inferences; second, it multiplies the loci of self-regulatory activity by providing numerous triggers for cognitive dissatisfaction. An audience monitors individual thinking, opinions, and beliefs and can elicit explanations that clarify points of difficulty. The learner's exposure to alternative points of view challenges his or her initial understanding". (p. 134)

Brown, Collins and Duguid (1989) have argued that groups are more than the simple addition of the individuals constituting them. Rather, the group interactions result in a certain synergism: "Groups are not just a convenient way to accumulate the individual knowledge of their members. They give rise synergistically to insights and solutions that would not come about without them" (p. 40).

A number of recent studies of learning in the science classroom confirm the view that small-group discussions may be a useful strategy for enhancing learning. Gaskins, Guthrie, Satlow, Ostertag, Six, Byrne, and Connor (1994) observed that peer discussion enhanced student understanding of science concepts. Finally, Fellows (1994) suggested that "writing ideas to themselves to explore and share informally with peers, reflecting on the ideas to reproduce new writing, and talking about the ideas with other students appeared to be important mechanisms for conceptual change" (p. 998-999).

A study investigating how talk or writing, or a combination of both language modes, affect learning certainly merits serious consideration. Bruffee (1984) has stated: "If thought is internalized conversation, then writing is internalized conversation re-externalized" (p. 641). Bruffee suggests that talk mediates between writing and thinking. According to this view, having peer groups discuss a topic enables students to shape the language which can be used later while writing. This view is further supported by Barnes (1990) who has observed that "exploratory talk often occurs when peers collaborate in a task" (p. 50). However, the critical issue here is how to organize these peer groups so that the interactions among students are focused on content and result in meaningful talk for clarifying, elaborating and shaping personal knowledge.

Although espousing the benefits of group learning, particularly for "accessing distributed knowledge" and for generating ideas, Linn and Burbules (1993) have cautioned educators about some of its pitfalls: An uncritical acceptance of ideas; a tendency to bring closure too soon while generating new ideas; and the role of student status in biaising group deliberations (p. 102).

Gall and Gall (1993) suggest that discussion may be used for cooperative learning and problem-solving, for simply mastering content or for analyzing complex issues. Cooperative learning through discussion establishes an interdependency among group members for accomplishing shared goals like learning content or solving problems (Johnson & Johnson, 1991; Slavin 1990). Content can still be learned through a discussion-based approach that gives students a forum for sharing, clarifying and elaborating knowledge without necessarily involving cooperative learning principles. However, the group structure is not as intricate as the cooperative model and not so dependent on collective action. Issues-oriented discussions generally focus on complex questions for which there are no simple answers and involve higher-order thinking for critical analysis and evaluation of alternate views (Cushman, 1992).

Using Cooperative Learning Principles

The cooperative learning model may be useful for structuring group work so that the discussion which ensues is meaningful and productive. Johnson and Johnson (1991) have described five components that characterize small-group cooperative learning situations: (1) social interdependence; (2) face-to-face interaction; (3) individual accountability; (4) interpersonal and small group skills; and (5) group processing. Face-to-face interaction refers to the importance of discussion, or verbal interaction, within the peer group. Interpersonal and small group skills refers to those skills which are essential if the group is to function effectively. Group processing refers to the importance of reflecting on individual and collective experiences in order to improve upon the functioning of the group.

Interdependency and Accountability. Advocates of cooperative learning have argued that two components, social interdependence and individual accountability, are fundamental in determining the effectiveness of this groupwork model (Johnson & Johnson, 1991; Slavin, 1990). Individual accountability occurs when "the best learning efforts of every member of the group must be necessary for the group to succeed" (Slavin, 1983, p. 441). Individual accountability can be established by either summing individual scores and assigning everyone in the group the average score or by assigning everyone a separate but well-defined task. Social interdependence "exists when each individual's outcomes are affected by the actions of others" in the group (Johnson & Johnson, 1991, p. 3). Slavin (1993), in more recent writings, has defined interdependency as establishing "group goals" within the cooperative learning group (p. 546). Social interdependence can involve depending on others for access to limited resources, for sharing information, for obtaining a reward or other incentive, or for performing the assigned task. For instance, giving group members different learning materials which are all required for task completion produces resource interdependency. A reward structure can be manipulated so that individual students will only benefit if everyone in the group achieves the stated goal. Finally, task interdependency can be created by assigning different roles to the group members: For instance, students might be assigned roles such as summarizer, facilitater, recorder, researcher, and observer, etc. (Cohen, 1986; Johnson & Johnson, 1991). This suggests that the discussions in this study may be more productive if students are assigned well-defined tasks or roles within the peer groups.

Giving Explanations. Webb (1982, 1985) has studied the nature and dynamics of groupwork in cooperative learning since the early eighties. In reviewing a series of studies which examined how peer interactions within small groups affect student learning in mathematics and computer science, Webb (1989) concluded that "giving elaborated explanations was positively related to achievement" (p. 29), whereas "receiving explanations is [only] sometimes helpful" (p. 26). Webb explains that whether or not receiving explanations is helpful depends on the knowledge state of the student receiving help. For instance, the explanation received must be relevant, at an appropriate level of elaboration, timely and understandable if it is to benefit the student receiving help. However, the most important finding here is the notion that the act of "explaining" enhances learning in the "explainer." This suggests that arranging a small-group task so that all students must give explanations has considerable merit for enhancing learning.

Studies in the Science Classroom. Lonning (1993) compared the use of discussion based on cooperative learning principles and unstructured discussion on achievement, defined as conceptual change, in a tenth-grade physical science unit for low-ability students. Students in the cooperative treatment group used significantly more verbal interactions which, according to Lonning, are theoretically correlated with conceptual change. Although this article lacks clarity in the reporting of the data and in their interpretation, the author still concluded that the experimental group showed greater conceptual change in their responses to test questions than the control group.

A recent study suggests that peer discussions within a cooperative learning framework can be an effective classroom strategy for learning science content at the elementary level. The study is particularly interesting because it contests cooperative learning dogma about the essential features of effective group work. Meloth and Deering (1994) compared the use of a reward condition, the traditional approach in cooperative learning activities, with a strategic condition which instead emphasized metacognitive knowledge. The reward condition used test scores as an incentive to reinforce the interdependency among the members of the groups. The strategic condition focused instead on developing metacognitive awareness of the task and the content among group members. Think-sheets, which were developed by the participating teachers using models provided by the researchers, were also used in both conditions to focus the discussion groups on the assigned tasks. Teachers reported, however, that the development of these think-sheets was an onerous task which they would not ordinarily perform on their own. Meloth and Deering concluded that "groups are more likely to focus their discussions toward task content and improve their awareness of learning goals and task dimensions when discussions are directly oriented toward the metacognitive features of their cooperative tasks" (p. 164).

Other strategies have been developed for focusing small-group discussions. Palincsar, Anderson and David (1993) developed a "metascript" to provide a structure for the discussion task (p. 646). They also worked closely with students to implement a set of social norms which provided additional guidance for small-group discussions. These norms included the need for everyone "(1) . . . to contribute to the group's efforts and help others contribute; (2) . . . to support one's ideas by giving reasons; (3) . . . to work to understand others' ideas; [and] (4) . . . to build on one another's ideas" (p. 647).

The Meloth and Deering (1994) study seems to suggest that a metascript, or a think-sheet of a more generic nature, focusing on the metacognitive knowledge underlying certain kinds of tasks, might be useful for guiding peer-group discussions: For instance, a generic think-sheet guiding explanatory tasks in science. Interestingly, Meloth and Deering (1994) also indicated that, despite the widespread use of cooperative learning methods over the last decade, little has been reported about the actual "content of peer-group discussions . . ." (p. 139). Video-recording group discussions and analyzing the transcripts might provide details about how students actually construct explanations of natural phenomena.

2.4 Concept Mapping

2.4.0 *Maps of Cognition*

Concept maps have been variously described as representations of "substantive structures" (Finley & Stewart, 1982, p. 593), as "knowledge representation tools" (Novak & Musonda, 1991, p. 130), as "representation[s] of meaning" (Novak, 1990b, p. 29), as "windows to the mind" (Malone & Dekkers, 1984, p. 231), and as "maps of cognition" (Wandersee, 1990, p. 923). A concept map can be defined simply as " a schematic device for representing a set of concept meanings embedded in a framework of propositions" (Novak & Gowin, 1984, p. 15). Finley and Stewart (1982) have explained that concept maps "are networks of meaning composed of concepts and systematic relationships among those concepts" (p. 595).

Novak (1991) has defined *concept* as "a perceived regularity in events or objects, or records of events or objects, designated by a label" (p. 45). The concept map gives a two-dimensional overview of the concepts related to a topic with these being hierarchically arranged from general, more-inclusive, and abstract concepts at the top of the page to the more specific, concrete concepts down the page. Words between each pair of concepts define the nature of the bond, or relationship, linking the pair into a logical proposition. Concept mapping essentially shows how individuals have structured their knowledge about a topic.

2.4.1 *Origins and Fundamental Aspects*

Concept mapping had its origins in Ausubel's cognitive assimilation theory which recognizes the importance of the learner's prior knowledge in meaningful learning. As Novak (1980) explained: "Meaningful learning occurs when new knowledge is consciously linked by the learner to existing concepts or propositions the learner already knows" (p. 282). Novak has frequently asserted that rote learning is the typical *modus operandi* in schools with students simply memorizing definitions, descriptions and explanations *verbatim*.

According to Ausubel, Novak and Hanesean (1978), meaningful learning can involve four different but complementary processes: Subsumption, integrative reconciliation, progressive differentiation, and superordinate learning. Subsumption occurs when learners simply elaborate and refine concepts already in their possession (e.g., wolves and dogs both belong to the family *canidae*). Integrative reconciliation occurs when the learner combines two or more previously separate concepts, unifying them into a single concept (e.g., density is mass divided by volume). Progressive differentiation occurs as learners discriminate between two or more subgroups derived from the same initial concept (e.g., protozoans can be classified into ciliates, flagellates, sarcodines, and sporozoans). Superordinate learning occurs when the learner moves from more specific concepts to a more inclusive, general concept (e.g., fish, amphibians, reptiles, birds and mammals are all vertebrates).

The four learning processes can be related to different aspects of a concept map (Novak, 1984). Progressive differentiation is shown in a learner's concept map by an increasing number of propositions. Integrative reconciliation is related to the addition of more valid cross-links between different branches of the map. Finally, subsumption and superordinate learning are reflected by an increasingly hierarchical structure and by more branching in the concept map.

2.4.2 *Improving Instruction*

Concept mapping had been proposed as a tool for improving instruction (Ausubel, Novak, & Hanesean, 1978; Barenholz & Tamir, 1992; Bernard & Naidu, 1992; Cliburn, 1987; Jegede, Alaiyemola, & Okebukola, 1990; Lehman, Carter, & Kahle; 1985; Malone & Dekkers, 1984; Mason, 1992; Novak, 1990a; Novak, Gowin, & Johansen, 1983; Okebukola, 1992; Okebukola & Jegede, 1988; Pankratius, 1990; Schmid & Telaro, 1990; Stewart, Van Kirk, & Rowell, 1979; Willerman & Mac Harg, 1991). Horton, McConney, Gallo, Woods, Senn, and Hamelin (1993) have conducted a meta-analysis evaluating the effectiveness of concept mapping as an instructional tool. They reported an effect size of 0.46 standard deviations for achievement when concept mapping was employed in the classroom. Horton and his colleagues concluded that the "instructional strategy of concept mapping . . . has had generally medium positive effects in students' achievement, and large positive effects on student's attitudes" (p. 107-108).

2.4.3 *Assessing Understanding*

Concept maps have also been used as an assessment tool in various studies of science learning. Hegarty-Hazel and Prosser (1991b) have argued that "concept maps are measuring things other than achievement . . . and can be used reliably to explore some aspects of students' propositional structure" (p. 428). The State of Connecticut has used concept mapping to assess student understandings in science in an innovative project entitled *ConnMap* (Lomask, Baron, Grieg, & Harrison, 1992). Concept maps were constructed *ex post facto* from students' written responses to traditional essay-type questions in a state-wide assessment, then the results were compared with an expert map based on the same topic. The results were finally collated on a three-dimensional concept map to show strengths and weaknesses on a statewide basis in the students' knowledge of science.

Here in Canada, Ross and Munby (1991) have used concept mapping to study secondary students' understandings of acids and bases. In this study, the researchers constructed concept maps, based on the results of interviews and multiple-choice tests, to represent each student's knowledge of the topic. These concept maps were then compared with a "model concept map" which was constructed on the basis of the chemistry curriculum actually being studied by the students (Ontario Ministry of Education Curriculum Guidelines). Ross and Munby concluded that:

> The concept maps themselves are especially revealing for how they portray gaps in students' understanding. Such gaps are not simply representative of

> missing knowledge, or of misconceptions. Concepts within acid and base chemistry are intricately connected, so the gaps evident in the concept maps represent areas in which chemical conceptual relationships do not appear to have been made by the students. (p. 22)

This suggests that concept mapping has great potential as a research tool for analyzing the nature of student understandings in science.

2.4.4 *Scoring Maps*

Student understandings in science as indicated by concept maps have generally been evaluated using two distinct approaches. The first approach has involved comparing students' concept maps to a model map reflecting expert opinion (Lomask, Baron, Greig, & Harrison, 1992; Ross & Munby, 1991; Wallace & Mintzes, 1990). The second approach has involved quantifying student concept maps according to a number of criteria.

Shore, Hakerem, and Hickman (1993) reported that the levels of hierarchy, as well as the number of cross-links and examples, did not differentiate among the various treatment groups learning about fractals. The only criteria to differentiate among these groups included the number of propositions and the number of critical concepts, examples and propositions. Cronin, Dekkers and Dunn (1982) had observed that the number of propositions was the most significant difference between concept maps constructed by teachers with those made by students.

In their study of study strategies, Hegarty-Hazel and Prosser (1991a & b) included three concept mapping criteria as variables: The number of valid propositions, the number of branches, and the number of cross-links. Students who volunteered for this study completed a concept map prior to instruction, then completed a second map and a study-strategy questionnaire after the instructional unit. Hegarty-Hazel and Prosser observed that the use of meaningful study strategies by students was positively correlated with better scores on the final concept maps. Of the three criteria, the number of valid propositions was the only one which changed significantly from initial to final concept maps. They concluded that "concept-mapping techniques are valid indicators of meaningful learning and can be used reliably to explore some aspects of students' propositional structure" (1991b, p. 428).

Stuart (1985) argued that the component scores using Novak's evaluation scheme are best kept separate rather than combining them into an aggregate score. The aggregate scores gave calculated reliabilities of less than 0.35, whereas reliabilities for the separate component scores ranged from 0.713 to 0.823. Kept separate, the component scores proved to be quite satisfactory for comparing the scores of different individuals at any one time, or for comparing the scores of the same individual at two different times. Stuart noted that one component, hierarchy, should be deleted from the evaluation scheme because it appeared to be strongly related to the relationships component.

Novak and Musonda (1991) proposed a scoring scheme for concept maps that considers the number of relevant concepts, the number of propositions and the number of misconceptions. Each of these criteria includes different levels; for instance, relevant concepts can range from broad, inclusive concepts to specific instances; propositional interlinkages can span three levels for reflecting hierarchical complexity; and misconceptions can be either major or minor.

Lomask, Baron, Greig, and Harrison (1992) evaluated students' concept maps to estimate the size and strength of their knowledge structures. Size was defined as the proportion of concepts included by the student that were also a part of the expert map. Strength was defined as the proportion of valid relationships depicted between concepts compared with those included in the expert map. Size and strength were then combined to evaluate the student's understanding of the topic.

Constructing a concept map from a transcript of a peer-group discussion should reveal the knowledge structure or network of propositions which the group has invented, collectively, for explaining a phenomenon or real-world application. As such, this concept map might be considered a measure of the effectiveness of the peer-group discussion for enhancing science learning.

In summary, the literature suggests that concept maps are adequate as representations of an individual's knowledge structure. Concept maps appear to be a valid research tool for assessing understanding, both conceptual and propositional, and can be scored reliably using criteria like the number of concepts and propositions.

2.5 Conclusions

The relationship between talking and learning has been the object of considerable theoretical work. Although the research base supporting the role of talk in learning is still inadequate, talking is viewed in this study as a powerful mediator of student learning. In contrast, the relationship between writing and learning is grounded in theoretical writings that are much more recent. However, in contrast to the role of talk, the research arguments supporting the role of writing for content learning are substantial.

The literature on writing to learn suggests that researchers have attempted to study writing in isolation rather than study it in context. Some researchers have studied writing in situations which are decontextualized from regular classrooms (Langer, 1984, 1986b; Langer & Applebee, 1987). Several studies have focused on the thinking processes of very small groups of students (Durst, 1987, 1989; Newell, 1984, 1986), whereas others have involved larger samples but for very brief periods (Hayes, 1987). As such, these studies lack ecological validity. Studies of writing to learn in science have neglected important contextual issues that may well determine the success or failure of these strategies in the classroom. More classroom-based research is required to show how these strategies might be used effectively by practitioners in everyday instructional situations.

Most of the studies of writing to learn have been conducted at the college level rather than at the elementary and secondary levels, or in other subject areas besides science. Although the college studies which have been reviewed here merit close scrutiny, there is no assurance that these findings about writing and learning can be generalized to elementary and secondary classrooms. Similarly, studies which have examined writing to learn in social studies may not be applicable to the science classroom because the conceptual demands across disciplines may be substantially different.

Writing can enhance science learning when teachers tailor tasks to attain meaningful curricular goals, when learners possess the necessary metacognitive knowledge, and when the instructional environment sustains a view of scientific literacy that embraces deep conceptual understandings rather than encyclopedic knowledge. As Langer and Applebee (1986) observed, “emphasis on writing as a way of learning may be impossible to implement when models of instruction . . . emphasize the importance of 'coverage' of content rather than mutual exploration of interpretations” (p. 185).

Philosophically, scientific explanations involve describing relationships between and among concepts. By its very nature, the act of explaining requires that individuals reveal their understanding or lack of understanding about these relationships. Working with peers to develop an explanation should allow individual students to share information, refine their ideas and together invent this propositional knowledge.

Although traditional measures of student knowledge will be used to compare learning in this study, concept mapping will provide a different view of students' knowledge structure with respect to the content domain being explored.

3. DESIGN AND METHODOLOGY

3.0 Subjects

3.0.0 *Population Characteristics*

The study was conducted at Collège Louis-Riel (CLR), a school situated in Saint-Boniface that includes students from grades seven to twelve. The sample for this study consisted of two intact grade eight classes composed of francophone students. However, despite the fact that the language of instruction is French, this sample could adequately represent any typical middle school science class in a surburban Canadian city. Although CLR receives students from all over metropolitan Winnipeg (Saint-Vital - 17.3% of the 1993-1994 student enrollment, Transcona - 5.4%, Winnipeg - 2.9%, St-James - 2.3%, Norwood - 1.9%, plus smaller numbers from Assiniboia, Fort Garry, and River East) as well as from some rural areas surrounding the city (Seine River School Division, Red River S. D., Whitehorse Plains S. D.), most of the students are residents of Saint-Boniface (66.9%). However, most of the students in eighth grade are from Saint-Boniface: Either established families living in 'old' Saint-Boniface or younger families who have moved into the newer housing development in the northern part of Saint-Boniface. Most of the students are from families with similar cultural, linguistic and religious backgrounds: they are Franco-Manitoban, francophone and Roman Catholic. However, some students are from families who have recently immigrated to Manitoba over the last decade. These students may have moved to Manitoba from Québec, France, Vietnam, Sénégal or some other Canadian province with an important francophone population, or other countries where French is an official language or was during colonial times. The socio-economic level of these students would range 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 post-secondary level (M. Fisette, personal communication, May 31, 1994).

3.0.1 *Sample Size and Sampling Procedure*

The teachers were self-selected in that they agreed to participate after receiving a general description of the study. Of the two teachers who expressed interest in participating in the study, only one received a subject assignment which was appropriate for the study. All of the 44 students in the volunteer teacher's two grade eight classes were included in the study. However, one student moved out of the province before the end of the study leaving only 43 students in the final sample. Students in each class were randomly assigned to four different groups, stratified for gender and ability. Since all peer discussion groups in both the talk-only (T) and the talk-and-writing (TW) treatments were to consist of exactly four students, preferably two males and two females, exactly twelve students were assigned to each of these treatment conditions. Students were initially randomly assigned to the TW group until all such groups were established for the class, then students were randomly assigned to the talk-only treatment condition until all such groups were established. Similarly, students were assigned to the writing-only treatment group (W). Finally, the remaining students constituted the control group (C group). The assignment to the different treatments and control groups has been summarized in Table 1.

Table 1

Assignment of Students to Treatment and Control Groups

	Group			
Class	Talk and Writing	Talk-Only	Writing-only	Control
One	2 X 4	1 X 4	6	4
Two	1 X 4	1 X 4 / 1 X 3 (4)	4	6
Combined	12	11 (12)	10	10

Note. The numbers within parentheses shows the initial assignment of students before mortality.

As much as possible given the actual composition of each classroom and the assignment to the different treatment groups, efforts were made to balance all groups for both gender and ability. The composition for each of the four treatment and control groups is shown in Table 2.

Table 2

Composition of Treatment and Control Groups for Gender and Ability

	Group			
Variable	Talk-Writing	Talk-Only	Writing-Only	Control
		Gender		
Males	7	8	6	6
Females	5	3 (4)	4	4
Totals	12	11 (12)	10	10
		Ability		
High	3	3	2	2
Average	6	7	5	6
Low	3	1 (2)	3	2
Totals	12	11 (12)	10	10

Note. The numbers within parentheses show the initial assignment of students before mortality.

3.1 Instrumentation

The instruments which were used to measure student learning included a multiple-choice test, a test with short essay questions, and concept maps. These measures gave "snapshots" of learning during the study period (Hewson & Stewart, 1994). Some of the peer discussions in the TW group were also videotaped to provide "movies" documenting the learning process.

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 concept and includes applications and explanations (Alexander, Schallert, & Hare, 1991; Prawat, 1989a).

This differentiation is consistent with the prevailing views in both the philosophical and the psychological traditions, the latter essentially representative of the cognitive school of thought. On the one hand, Hempel (1966) reflects the philosophical view when he stated that concepts are "the knots in a network of systematic interrelationships in which laws and theoretical principles form the threads" (p. 94). On the other hand, Reif and Larkin (1991) express the psychological view: "The basic building blocks of knowledge in any domain are concepts . . . and relations among these concepts" (p. 745). The approach taken in this study was to try to tease out the simple knowledge composed of concepts, or building blocks, from the integrated knowledge composed of the "relations" or "threads" which link up these concepts that together make up the knowledge structure of students in a particular content domain.

3.1.0 *Data Collection*

Multiple-Choice Tests

Both versions of the multiple-choice tests, versions A and B, included exactly 30 items, with each item measuring either simple or integrated knowledge. The proportion of questions measuring each of these knowledge categories on the test was equal.

Most of the questions were borrowed or adapted from various international, national and provincial science assessments, as well as from published item banks, that are now in the public domain. The origin of these question included the National Assessment of Educational Progress (NAEP), the Educational Testing Service (ETS), the International Association for the Evaluation of Educational Achievement (IEA), the New South Wales Department of Education, the Learning Assessment Branch of the Ministry of Education in the province of British Columbia, the Manitoba Science Assessment Program, and the Association des Institutions d'Enseignement secondaire in the province of Québec. The remaining fifteen of sixty question were constructed by the researcher.

The multiple-choice tests were administered three weeks before the intervention (pretest), immediately after the intervention (immediate posttest), and again six 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. All tests were scored by hand by the researcher after the final posttest.

Essay Questions

Four short essay questions, in which students had to provide written explanations for natural phenomena or real-world applications relevant to the content domain under study, were also administered three times during the experiment at about the same time as the multiple choice tests. Two of these questions had been used in previous assessments in the province of Manitoba and two of them were constructed by the researcher. Two scores were established for each of these essay questions: a score for simple knowledge that was based on the number of target concepts included in the response, and 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 was 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 simple knowledge score for the four essay questions was 23 points. Student responses to each question were also assessed using a holistic marking scale to establish a score for integrated knowledge. 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.

Criteria for scoring the integrated knowledge component 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; Freedman, 1994; Gorman, Purves, & Degenhart, 1988; Hart, 1994; Lamb & Purves, 1988; Purves, 1984). Table 3 shows the criteria used for scoring question C, a question about the winter adaptations of animals.

Table 3

Criteria used for Scoring Essay Question C

Scale	Conceptual Relationships	Clarity	Organization	Adequateness of Explanation
4	Lists five adaptations and correctly explains how each one helps the animal to survive.	Clear	Well-organized	Complete
3	Lists four or more adaptations and correctly explains how each one helps the animal to survive.	Generally clear	Satisfactorily organized	Adequate
2	Lists at least three adaptations and correctly explains how each one helps the animal to survive.	Generally clear	Satisfactorily organized	Partial
1	Lists one or more adaptations without necessarily explaining how each one helps the animal to survive.	Not clear	Poorly organized	Partial
0	No response ((NR)	NR or unreadable	NR	NR or irrelevant

Concept Maps

The teacher taught the students how to construct concept maps several weeks prior to beginning the unit of study (Novak & Gowin, 1984). Students also completed several 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 thirteen 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 five 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 estabished by the expert panel.

Students received two points for each of the key concepts which were not given as seed concepts that were included in the map. Students also received one point for each seed concept and for each additional concept added to the map. The number of key concepts, seed concepts, and other concepts thus established the simple knowledge score for the students. In addition, 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 had 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.

These two variables, concept score and proposition score, respectively, represent the size and strength of the student's knowledge structure in ecology (Lomask, Baron, Grieg, & Harrison, 1992). The size of the knowledge structure is reflected by the number of concepts included in the map, whereas the strength of the knowledge structure is reflected by the links or connections which the student included within this conceptual network.

Case Studies of Discussions

Two peer discussions were videotaped for each of the five problem-solving sessions planned for the duration of the unit. The same group of students was videotaped each time in both classes so that the recording process might become less obtrusive over time. The videotapes were transcribed for later analysis and then destroyed. Some of these transcriptions, which constitute case studies of peer-discussion groups in action, were analyzed to determine how student understanding evolved during these group interactions. This data was analyzed qualitatively in order to flesh out some of the quantitative measures like the essay questions and the concept maps.

3.1.1 *Reliability and Validity*

Multiple Choice Tests

Validity. Both versions of the multiple choice tests were composed of three pools of items: items which have been used in provincial, state, and national assessments; items which have been used in international assessments; and still others which have been constructed either by publishers or by the researcher. Since all of the items in the first two categories had been reviewed by a technical advisory committee prior to their use, the content validity of these items was assumed to be adequate. Nonetheless, both versions of the tests were reviewed by a panel of four experts in education. One expert works as a curriculum consultant with the provincial department of education and is responsible for science education in Franco-Manitoban and French immersion schools. Another worked as a curriculum consultant, but completed his career as a private consultant in education until his untimely death during this study. A third expert works as a consultant with the same department in the area of measurement and evaluation. The fourth expert has worked as a science curriculum consultant in the past, but is now an administrator with the same department. He is responsible for curriculum development and implementation in all subjects, Kindergarten to grade twelve in both French immersion and Franco-Manitoban schools. All four experts are fluently bilingual and three of them have Master's degrees in education.

Three grade eight teachers reviewed all of the tests to ensure that the readability was appropriate for the target students. Items which were deemed too difficult were either discarded or rewritten until the language proved satisfactory. A curriculum consultant in science reviewed all test items to ensure that they were appropriate given the grade eight science curriculum being used in the schools. Once again, items which were inappropriate were discarded or revised. The four experts were asked to categorize each item into two categories: (1) those which measure simple knowledge (facts, terminology and concepts), and (2) those which measure integrated knowledge (applications, interpretations, and explanations). Disagreements were resolved through discussion and concensus.

Reliability. The two versions of the multiple-choice test were administered to all of the students at the delayed posttest session. Some students wrote version A before version B, while others wrote B before A. A reliability coefficient of equivalence was calculated on the basis of this data. The Pearson correlation coefficients for simple knowledge was 0.67 and 0.65 for integrated knowledge. However, the coefficient for the total score on the multiple choice test was 0.80. McMillan and Schumacher (1989) indicate that a coefficient above 0.70 is acceptable for research purposes.

Essay Questions

Validity. The validity of the essay questions was established using an expert panel composed of three biology professors and one science curriculum consultant with extensive experience in environmental education. Simple knowledge was determined as the number of target concepts represented in the student's written response. The target concepts for each of the four essay questions were established using this panel of experts. For example, according to the experts, the first question (A) which asked students to explain cougar sightings in the Red River basin involved a cluster of eight concepts: ecosystem, abiotic factor, population, community, food chain, food web, habitat, and niche.

The score for integrated knowledge was established using a holistic marking scale. Criteria for scoring these written responses included clarity, organization and adequateness of the explanation. The conceptual relationships considered appropriate for an adequate explanation to each of the four questions were derived using feedback from the expert panel. For example, question B asked students to explain how someone might use their knowledge of the feeding and reproductive habits of the jackfish for fishing. The conceptual relationships which were judged to be appropriate by the expert panel included: (a) The use of bait or lures that ressembles its prey; (b) Knowing where to fish (depth and temperature of water); (c) Knowing when to fish (time of the day, seasonal differences); (d) Being able to identify optimal water temperature for jackfish; and (e) Respecting laws that forbid fishing, for instance during spawning. Each of the four questions was evaluated using a four-point marking scheme (0-4) with the total possible score being 16 points.

Reliability. Student responses to these questions were scored by the researcher and two research assistants who had been trained for the scoring task. The student received the average of the three scores for each question. Inter-rater reliability was established for the scoring of each question. Pearson correlation coefficients were calculated between all possible pairs of scorers. Table 4 summarizes the minimum, maximum, and mean correlation coefficients for the four essay questions, both simple knowledge and integrated knowledge.

Table 4

Inter-Rater Reliability for Scoring Essay Questions

	Pearson Correlation Coefficients		
Question	Minimum	Maximum	Mean
	Simple knowledge		
A	0.85	0.89	0.87
B	0.90	0.92	0.91
C	0.83	0.87	0.85
D	0.91	0.94	0.92
	Integrated knowledge		
A	0.77	0.80	0.79
B	0.84	0.87	0.85
C	0.81	0.85	0.84
D	0.85	0.91	0.80

For all four questions, the minimum and maximum correlation coefficients for scoring simple knowledge in student responses were 0.83 and 0.94, respectively. For integrated knowledge, the coefficients ranged from 0.77 to 0.91. 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). Since the objective of this study was not to compare individual students, but experimental treatments, the reliabilities were considered adequate for making comparisons between the different groups.

Concept Maps

Concept mapping has been used extensively in studies of science teaching and learning, and an approach using this technique for representing student knowledge has considerable merit.

Validity. Some studies have shown that concept maps are as valid as clinical interviews for revealing conceptual understanding (Edwards & Fraser, 1983; Wallace & Mintzes, 1990). This suggests that the concurrent validity of concept maps with interviews is more than adequate. Other studies have reported significant correlations between achievement measures and concept-map scores (Rice, Ryan, & Samson, 1993; Rogan, 1988; Wallace & Mintzes, 1990). Novak and Musonda's (1991) twelve-year study also supports the adequacy of concept maps as "a representational tool for the individual's cognitive structure" (p. 134). The validity of concept mapping as a knowledge representation technique therefore seems to be adequate.

Reliability. Lomask, Baron, Grieg and Harrison (1992) concluded that, with sufficient training, raters could reliably agree on concept map scores. However, no figures were reported in this preliminary study. Powers (1990) calculated an inter-rater reliability of 0.955 for the scoring of concept maps on the basis of propositions, hierarchy, cross links and branching. Using criteria based on the number of concepts, their linkages, and their organization within the map, Mason (1992) calculated an inter-rater reliability of 0.80 for scoring maps. Rice, Ryan, & Samson (1993) reported an inter-rater agreement of 0.98 for scoring maps. Finally, Novak and Musonda (1991) reported an inter-rater reliability of about r = 0.95 for scoring relevant concepts, propositions and misconceptions (p.130). Concept maps can be reliably scored if criteria are well-defined in scoring keys and if scorers are properly trained.

Concept mapping appears to be a valid tool for representing conceptual knowledge and a reliable instrument for research purposes. 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; Hegarty-Hazel & Prosser, 1991a & b; Lomask, Baron, Greig, & Harrison, 1992; Mason, 1992; Novak & Musonda, 1991; Powers, 1990; Shore, Hakerem, & Hickman, 1993; Stuart, 1985.

The concept maps were scored by the researcher and two research assistants who had received prior training. The inter-rater reliability was calculated separately for the simple knowledge and integrated knowledge components for all possible pairs of scorers. The minimum, maximum and mean Spearman correlation coefficients are reported in table 5.

Table 5

Inter-rater Reliability for Scoring Concept Maps

	Spearman Correlation Coefficients		
Knowledge	Minimum	Maximum	Mean
Simple	0.98	0.99	0.98
Integrated	0.97	0.99	0.98

Since the Spearman correlation coefficients were all above 0.90, the inter-rater reliability was considered adequate for the study.

3.2 Procedures

3.2.0 *Identification of the Groups*

Three treatment groups and a control group were used in this study:

1. A talking-and-writing group (TW) in which students discussed the problem tasks in small peer groups prior to individually writing their response to each question.
2. A talking-only group (T) in which students discussed the problem tasks in small peer

groups, similar to the first treatment group. However, students in this treatment were not asked to write out their response to the questions.

3. A writing-only group (W) in which students individually wrote their responses to the problem task. However, these students did not benefit from the peer discussions.
4. A control group (C) in which individual students were assigned simple content learning tasks involving fill-in-the-blanks, true-or-false exercises, matching exercises, definitions and descriptions).

Random assignment, stratified by gender and ability, was used to establish the groups in each of the two grade eight classes. Ability in science for each student was determined by the grades obtained the previous year in grade seven science. Students in each class were sorted into three groups: The low-ability group included those students who obtained grades in the lowest quartile relative to other students; the high-ability group included those students who obtained grades in the highest quartile; and the middle-ability group included those students who were between these two groups. Oliver and Simpson (1988) have argued that course grades, despite researcher concerns about reliability, are "typically the only indicator that students have of their progress in science" (p. 147). Over a period of several years, course grades could be considered a measure of a student's demonstrated ability in science. Oliver and Simpson reported that the reliability of science grades for students over a period of three or more years, using Cronbach's formula, yielded an alpha value of 0.81. The use of course grades might therefore be considered an appropriate and unobtrusive measure of a student's ability in science. However, Lock (1992) had established the ability levels of students on the basis of teacher perceptions. The teacher was thus asked to review the placement of the students in the groups when course grades from the previous year were used. He reported that the assignment to groups using this criteria was consistent with his perceptions of their abilities in science.

3.2.1 *Specification of the Experimental Treatment*

A mixed factorial design (4 X 2 X 2) was used with treatment (4) and gender as between-subjects variables and time-of-test (2 - pretest and immediate posttest, or pretest and delayed posttest) as a within-subjects variable. Although the students were not randomly selected from the population at large, assignment to the treatment and control groups was random, stratifying for gender and ability.

An ecology unit in the grade eight science program established the context for the study which lasted six weeks during the Fall term of the 1994-1995 school year. At the beginning of this unit, the teacher instructed students about the nature of scientific explanation (Salmon, 1987) and about the different types of explanation: compositional, evolutionary, functional, and transitional (Kourany, 1987). The instructional material used by the teacher was prepared by the researcher and has been included in Appendix A.

Problem-Solving Sessions

Problem-solving can be defined as "thinking [which] is functional, active, and grounded in goal directed action" (Rogoff, 1990, p. 8). According to Rogoff, problem-solving can involve a variety of different tasks, including writing an explanation for some everday natural phenomena or simply exploring new ideas which apply basic biological concepts. Problems which involved explaining natural phenomena were assigned five times during the teaching of this instructional unit. Each of these 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 first session asked students to explain whether or not a dead log could be considered an ecosystem. These explanatory tasks required students to relate various ecology concepts being studied during the unit. The explanatory tasks which were given to students during the problem-solving session on ecosystems, populations and communities have been included in Appendix B.

Students in the three treatment groups were separated for the duration of the problem-solving sessions, which lasted a full fifty-five minute period. The students in the writing-only and control groups worked individually in one classroom. Students in the discussion, or talk-only group, and the talk-and-writing group worked in separate classrooms which accommodated small-group work. The three treatment groups were supervised from one week to the next by the teacher, the researcher, or a number of research assistants who rotated among the groups each time. Apart from these problem-solving sessions, the students in each of the two class sections (A and B) were together for the remainder of the teaching unit. The students in each class read the same materials, did the same assignments and received essentially the same instruction throughout the entire unit. What differentiated the three groups was the nature of the experimental treatment: One treatment isolated writing as a learning strategy, a second treatment isolated talk or discussion, whereas a third treatment combined talk with writing for explaining real-world applications of the ecology concept being studied. Students in the control group were assigned simple learning tasks related to ecology, the content domain which established the context for the problem: Fill in the blanks, true-or-false exercises, matching exercises, definitions, and observational and descriptive tasks. The simple learning tasks which were given to students in the control group during the problem-solving session on biomes have been included in Appendix C.

Students in the writing-only group were instructed to individually respond to the explanatory tasks while working alone at their desks. Similarly, students in the control group received simple learning tasks to complete alone. Students in the talk-only treatment were assigned to peer-discussion groups. These students were instructed to discuss the explanatory tasks during the class period. They were not asked to write out a response to these questions. Similarly, students in the talk-and-writing group were also assigned to peer-discussion groups and instructed to discuss the explanatory tasks. However, they were instructed to individually write a response to the explanatory task once they had finished discussing each question. The instructions which were given to the persons supervising the problem-solving sessions have been included in Appendix D.

Peer Discussions

The peer discussion groups were instructed to follow a three-step procedure during the problem-solving sessions. The first step involved brain-storming possible explanations and emphasized the participation of all students (Linn & Burbules, 1993). The second step required students to elaborate, to clarify, and to ask or to answer questions about the proposed explanations (Webb, 1989). The third step invited students to evaluate, to criticize, to justify, and to revise their ideas. Students were given a written prompt during these discussion to scaffold metacognitive awareness during the explanatory session (Coleman, 1992; Meloth & Deering, 1994; Palincsar, Anderson & David, 1993). Students also received guidelines about the kinds of behaviours that promote constructive discussions (Bridges, 1990; Cohen, 1986; Costa, 1990; Johnson, Johnson, Holubec, & Roy, 1988; King, 1993; Lonning, 1993; Palincsar, Anderson, & David, 1993). The three-step procedure, the guidelines for discussing the explanatory tasks, and the written prompt which peer groups used during discussions have been included in Appendices E, F and G, respectively.

Variables and Measures

The independent variables in this study included: (1) the particular learning strategy employed during the explanatory task sessions (writing-alone, talk-alone, talk-and-writing, and a control); (2) gender (male or female); and (3) ability (low, average or high). Dependent variables included simple and integrated knowledge. The specific measures of simple and integrated knowledge that were used in the study included: (1) multiple choice tests; (2) four essay questions; and (3) concept maps. The aggregate score was defined as the total of the three measures. Table 6 summarizes information about the dependent variables in this study.

Table 6

Description of Dependent Variables

Knowledge	Measures: Multiple Choice	Essay Questions	Concept Maps	Aggregate
Simple	15 questions	4 questions	Number of concepts	Total simple
Integrated	15 questions	4 questions	Number of propositions	Total integrated
Total	30 questions	4 questions	Concepts and propositions	Total score

All three measures were administered at three different times: (1) A pretest given 2-3 weeks prior to beginning the unit; (2) An immediate posttest given upon completing the unit; (3) A delayed posttest given six weeks after completing the unit.

The multiple choice test was given in one class period and the essay questions and the concept mapping task were given in another class period. The two classes wrote the multiple choice pretests on consecutive days in September: Class B completed version B on Monday, while class A completed version A on Tuesday of the same week. Both classes completed the combined essay and concept map pretest on Thursday during consecutive periods to minimize leakage between classes.

The two classes wrote the immediate posttest in late November. The combined essay and concept map test was written first by both classes during the same period. The multiple choice tests were completed the following week on two different days: Class B completed version A on Monday, while class A completed version B on Tuesday.

For the delayed posttest, both classes completed the tests during the same periods in January: The essay questions and concept maps were completed during the first session, while the multiple choice tests were completed the following day. Although all students completed both versions of the multiple choice test during this final session to establish equivalence for an estimate of reliability, the results of only one version was used in the statistical analyses: Version A for class A and version B for class B.

Whenever students were absent during the testing sessions, arrangements were made for them to complete the tests whenever possible following their return to class.

3.3 Method of Data Treatment

Statistical treatment of the data involved analysis of covariance using the pretest scores as covariate each time. Hypothesis testing involved analyses using the aggregate scores only. However, separate analyses were performed on data from the multiple choice tests, the essay questions, and the concept maps to recommend changes in instrumentation in any future study. The first set of analyses examined effects immediately following the completion of the unit (posttest 1), whereas the second set examined longer term effects at the delayed posttest stage (posttest 2). Since the study was exploratory in nature, the objective during analysis was to explore the data for trends and relationships rather than to draw firm conclusions regarding experimental effects.

4. ANALYSIS OF DATA

The data included various measures of simple and integrated knowledge that were obtained at three different times during the study: (1) a pretest given two to three weeks before beginning the unit of study; (2) an immediate posttest given upon completing the unit; and (3) a delayed posttest given six weeks later. The measures included scores based on (1) multiple choice tests; (2) essay questions; and (3) concept maps. The scores on these measures were also combined into separate aggregate scores for simple and integrated knowledge with the sum of these two serving as a total knowledge aggregate score.

Separate analyses were conducted for the data obtained from the multiple choice tests, the essay questions, the concept maps, and the aggregate scores. First, descriptive statistics for each measure will be presented using tables and graphs. Second, the tests of hypotheses which guided the study will be presented and discussed. However, only the aggregate scores were used for testing the hypotheses. The other measures were analysed for both exploring trends in the data and for improving the instruments themselves. Third, the data will be reviewed to answer the research questions that were addressed in this study. Fourth, the limitations and assumptions underlying the study will be considered. Finally, excerpts of peer discussions which were audiotaped will be reviewed to flesh out the statistical analyses.

Unfortunately, the number of students that could be included in the study was less than expected when the study was first proposed. The data obtained in this small - *N* study would thus have to be interpreted very cautiously for definitive conclusions about language and learning in the science classroom. However, this study will also serve an exploratory function, allowing more latitude in interpreting trends in the data, so that a stronger follow-up study might result.

The data was analysed using SYSTAT, Version 5.2, on a Macintosh Plus computer which had been upgraded to four megabyte of RAM. The SYSTAT program was developed by SYSTAT, Inc. of Evanston, Illinois and was copyrighted in 1992.

4.0 Descriptive Statistics

4.0.0 Multiple Choice Tests

Table 7 gives the means and standard deviations for the pretest, the immediate posttest (posttest one), and the delayed posttest (posttest two) results obtained from the multiple choice tests. Simple knowledge and integrated knowledge were each measured using fifteen questions, whereas the total possible score which was based on both groups of questions was worth 30 points.

Table 7

Descriptive Statistics for the Multiple Choice Tests

Group	Pretest		Posttest 1		Posttest 2	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
			Simple Knowledge			
Control (*n* =10)	6.7	2.00	8.0	2.71	8.7	2.58
Talk (*n* =11)	7.5	2.46	10.8	2.60	10.5	2.88
Talk-writing (*n* =12)	6.3	2.14	9.9	2.31	9.6	2.28
Writing (*n* =10)	6.6	3.00	8.6	2.72	9.2	2.82
			Integrated Knowledge			
Control	7.6	3.20	10.8	2.20	10.5	2.76
Talk	8.6	3.38	12.7	1.74	11.7	3.10
Talk-writing	9.8	1.96	11.0	1.21	11.5	2.36
Writing	9.3	2.36	9.9	2.18	10.0	2.91
			Total Score			
Control	14.3	4.88	18.8	4.32	19.2	4.80
Talk	16.2	5.25	23.5	3.78	22.2	5.19
Talk-writing	16.0	3.59	20.9	2.84	21.1	4.32
Writing	15.9	4.82	18.5	4.01	19.2	5.14

Regardless of the relative order of groups on the pretest, the talk-only group and the talk-and-writing group ranked first and second, respectively, on the immediate and delayed posttests for simple, integrated and total knowledge mean scores. In comparison, the control group and the writing-only group alternated between third and fourth for these same measures. The talk-only group and the talk-and-writing group showed the biggest improvements in mean total scores on the multiple-choice tests. In comparison, the writing-only group and the control group posted the smallest increases from one test to the next.

Using group and gender as categorical variables, analysis of covariance (ANCOVA) was employed to explore the multiple choice data. Separate analyses were conducted for the immediate and delayed posttests. Table 8 shows the adjusted means and the standard errors for the multiple choice data, using the pretest scores as covariate each time, when both treatment and gender are used as categorical variables.

Table 8

Adjusted Means (M_{adj}) and Standard Errors (SE) for the Multiple Choice Data by Gender and Treatment

	Posttest 1				Posttest 2			
	Male		Female		Male		Female	
Group	M_{adj}	SE	M_{adj}	SE	M_{adj}	SE	M_{adj}	SE
				Simple Knowledge				
Control	8.2	0.96	7.7	1.17	8.6	0.89	9.0	1.09
Talk	10.1	0.84	11.2	1.35	9.2	0.78	12.2	1.25
Talk-writing	11.3	0.97	9.1	0.96	10.1	0.90	9.7	0.89
Writing	8.6	0.96	8.8	1.19	9.9	0.89	8.4	1.11
				Integrated Knowledge				
Control	11.2	0.72	10.9	0.85	11.7	0.70	11.2	0.83
Talk	12.7	0.60	13.0	0.98	12.0	0.59	11.6	0.96
Talk-writing	10.4	0.71	11.2	0.70	10.6	0.69	11.0	0.68
Writing	10.9	0.70	8.2	0.85	11.0	0.68	7.7	0.83
				Total Knowledge				
Control	19.7	1.36	18.8	1.64	20.3	1.23	20.1	1.49
Talk	23.0	1.16	24.2	1.89	21.0	1.05	23.7	1.72
Talk-writing	21.4	1.34	20.1	1.34	20.8	1.22	20.7	1.22
Writing	19.5	1.34	16.7	1.64	20.7	1.22	16.4	1.49

Male students in the talk-only and the talk-and-writing groups ranked first or second for all but one of the multiple choice measures on the immediate posttest, whereas female students in these same groups ranked first and second, respectively, on these same measures. The rank order of groups on the delayed posttest generally showed a similar pattern.

Table 9 shows the adjusted means and the standard errors, using the pretest scores as covariate each time, when treatment alone is examined.

Table 9

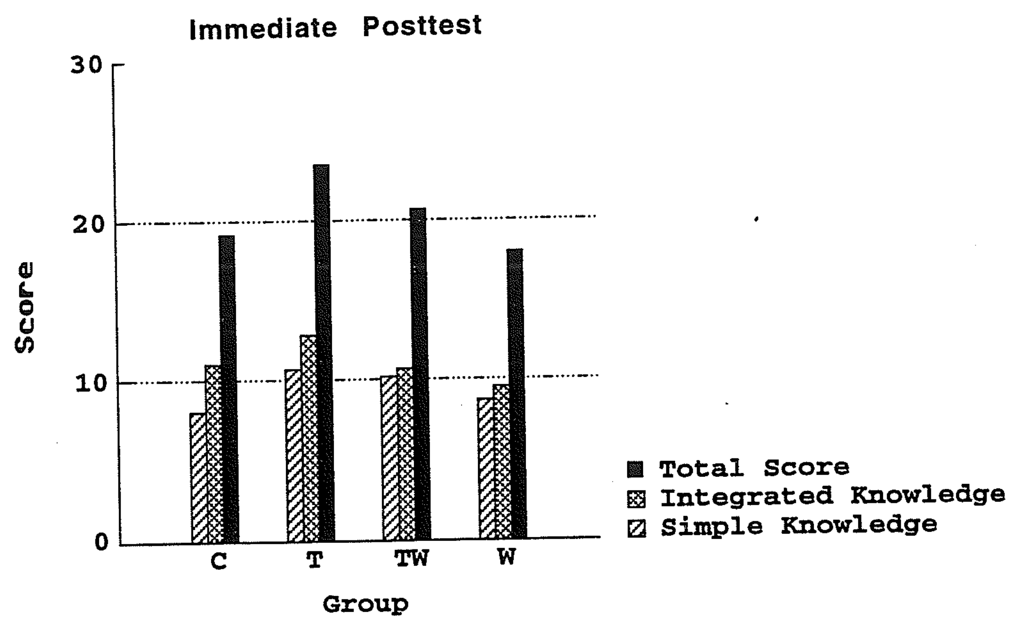
Adjusted Means and Standard Errors for the Multiple Choice Data by Treatment

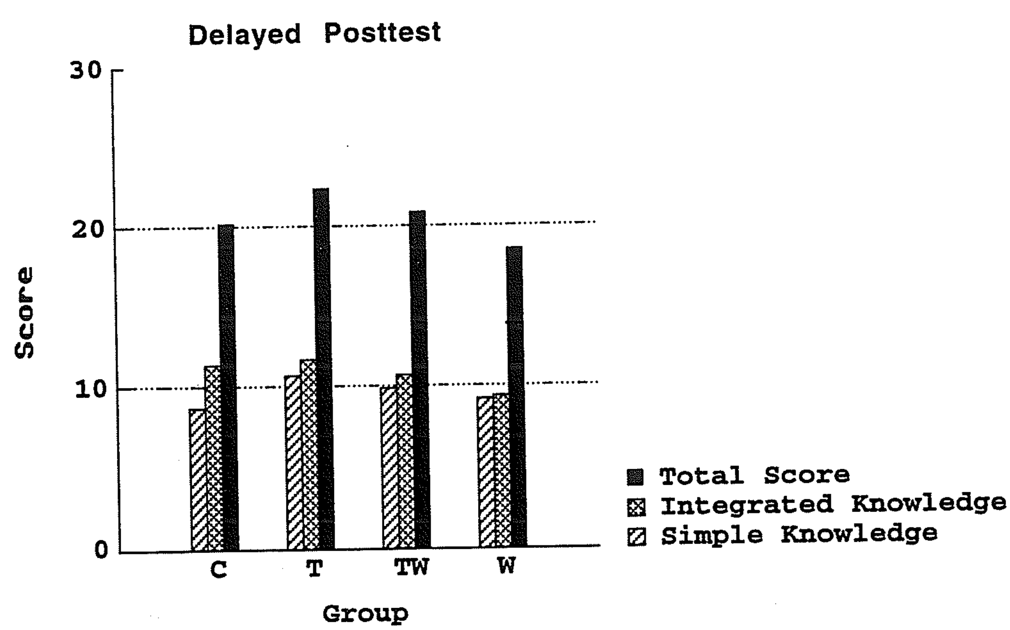
	Posttest 1		Posttest 2	
Group	Adjusted Mean	*SE*	Adjusted Mean	*SE*
		Simple Knowledge		
Control	8.0	0.75	8.8	0.70
Talk	10.7	0.80	10.7	0.74
Talk-writing	10.2	0.68	9.9	0.63
Writing	8.7	0.76	9.2	0.70
		Integrated Knowledge		
Control	11.0	0.56	11.4	0.55
Talk	12.8	0.58	11.8	0.56
Talk-writing	10.8	0.50	10.8	0.49
Writing	9.5	0.55	9.3	0.54
		Total Score		
Control	19.3	1.06	20.2	0.97
Talk	23.6	1.11	22.4	1.01
Talk-writing	20.8	0.95	20.8	0.86
Writing	18.1	1.06	18.6	0.96

Bar graphs of the adjusted means by treatment for the multiple choice tests from both immediate and delayed posttests have been included in Figure 1.

Figure 1

Bar Graphs of Adjusted Means for Multiple Choice Tests by Treatment





The talk-only and the talk-and-writing groups ranked first or second for all but the integrated knowledge adjusted mean score for which the control group did better than the talk-and-writing group, but not the talk-only group. For all other measures, the control group and the writing-only groups ranked third or fourth behind the other two groups using talk (TW and T).

4.0.1 ***Essay Questions***

Table 10 gives the means and standard deviations for the pretest, the immediate posttest, and the delayed posttest results obtained from the essay questions. Simple knowledge, integrated knowledge, and the total score were all measured using cumulative scores which students obtained on four different questions. The maximum scores for the simple knowledge and integrated knowledge components were 23 points and 16 points, respectively.

Table 10

Descriptive Statistics for the Essay Questions

	Pretest		Posttest 1		Posttest 2	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
			Simple Knowledge			
Control (*n* =10)	7.1	3.60	10.3	2.63	9.6	2.38
Talk (*n* =11)	9.6	2.57	12.4	2.11	11.2	1.82
Talk-writing (*n* =12)	7.5	4.18	11.0	2.62	9.9	2.15
Writing (*n* =10)	9.1	2.92	11.0	3.10	10.4	2.92
			Integrated knowledge			
Control	5.8	3.21	8.1	2.34	8.0	1.65
Talk	6.6	2.52	10.3	3.42	9.1	2.13
Talk-writing	5.4	3.17	9.0	3.13	8.5	3.47
Writing	6.3	2.86	9.4	3.11	8.3	3.22
			Total Score			
Control	12.1	6.77	18.4	4.72	17.6	3.42
Talk	16.3	4.74	22.7	4.73	20.3	3.57
Talk-writing	12.9	7.20	20.0	5.44	18.5	5.37
Writing	15.4	5.35	20.4	5.77	18.7	6.05

On the basis of mean scores for the essay questions, the biggest changes over time occured on the integrated knowledge measure. Although the talk-only group ranked first on all three tests (pretest, posttest 1, and posttest 2), the control group dropped from third place on the pretest to fourth place on both posttests. The writing-only group dropped from second place on both pretest and immediate posttest to third place on the delayed posttest. The talk-and-writing group improved in rank from last place on the pretest, to third place on the immediate posttest to second place on the delayed posttest.

Using group and gender as categorical variables, analysis of covariance (ANCOVA) was employed to explore the essay question data. Separate analyses were conducted for the immediate and delayed posttests. Table 11 shows the adjusted means and standard errors by treatment and gender for the essay questions data using the pretest scores as covariate each time.

Table 11

Adjusted Means and Standard Errors for Essay Questions by Gender and Treatment

	Posttest 1				Posttest 2			
	Male		Female		Male		Female	
Group	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*
				Simple Knowledge				
Control	11.7	0.78	10.0	0.92	10.7	0.72	9.6	0.85
Talk	11.6	0.65	11.6	1.11	10.5	0.60	10.5	1.03
Talk-writing	11.8	0.82	11.1	0.77	11.4	0.75	9.4	0.71
Writing	11.3	0.77	9.6	0.92	10.3	0.71	9.4	0.85
				Integrated Knowledge				
Control	9.1	0.94	7.1	1.14	8.7	0.83	7.3	1.01
Talk	10.2	0.80	9.0	1.36	9.2	0.71	7.2	1.20
Talk-writing	9.4	0.99	9.5	0.94	9.7	0.88	8.3	0.83
Writing	9.8	0.93	8.2	1.13	8.3	0.82	7.8	1.00
				Total Knowledge				
Control	20.8	1.49	17.2	1.78	19.5	1.39	17.0	1.66
Talk	21.8	1.25	20.6	2.14	19.5	1.17	17.6	2.00
Talk-writing	21.2	1.58	20.6	1.48	21.2	1.47	17.7	1.38
Writing	21.0	1.47	17.8	1.78	18.5	1.38	17.2	1.66

On the immediate posttest, male students in the talk-only and talk-and-writing groups ranked first and second, respectively, for total scores on the essay questions test.

On the delayed posttest, the rank-order of these two groups on the same measures was reversed. In comparison, female students in the talk-and-writing group tied the talk-only group on the immediate posttest and scored better than them on the delayed posttest.

Table 12 shows the adjusted means and the standard errors, using the pretest scores as covariate each time, when treatment alone is displayed.

Table 12

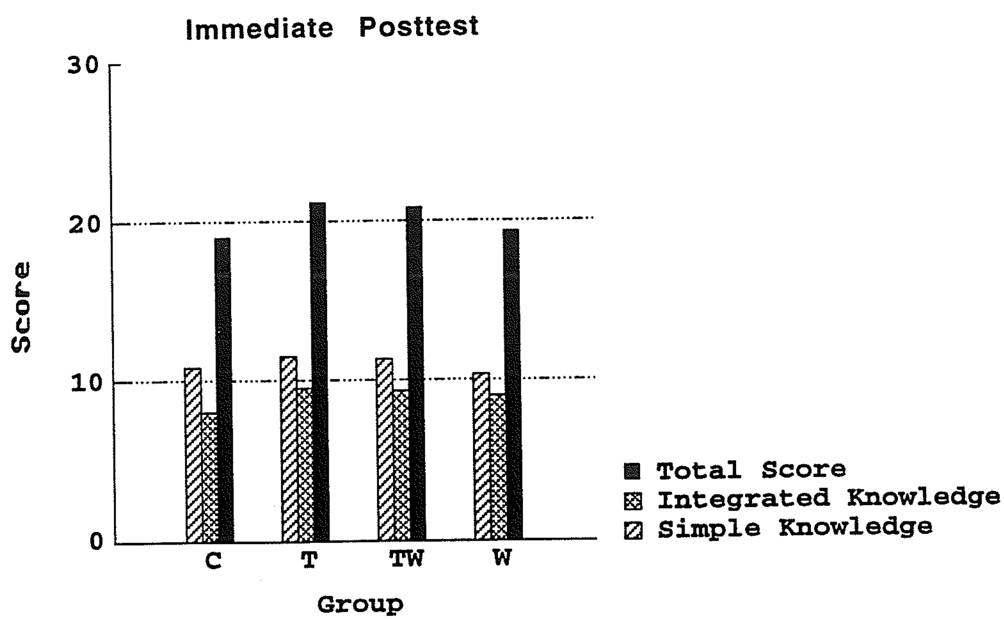
Adjusted Means and Standard Errors for Essay Questions by Treatment

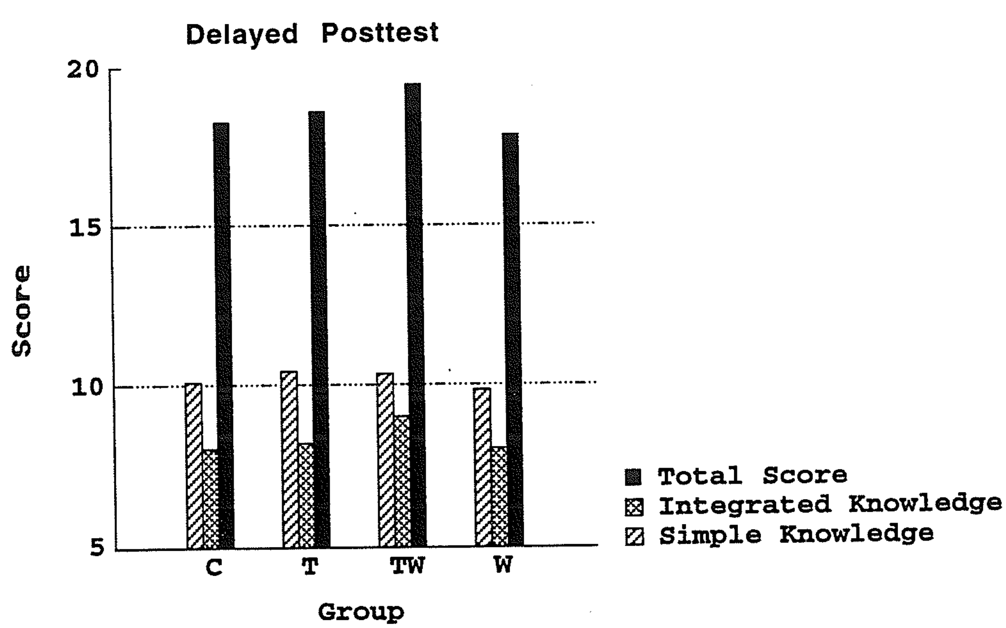
	Posttest 1		Posttest 2	
Group	Adjusted Mean	*SE*	Adjusted Mean	*SE*
		Simple Knowledge		
Control	10.9	0.60	10.1	0.56
Talk	11.6	0.65	10.5	0.60
Talk-writing	11.5	0.54	10.4	0.50
Writing	10.4	0.60	9.9	0.55
		Integrated Knowledge		
Control	8.1	0.73	8.0	0.64
Talk	9.6	0.79	8.2	0.69
Talk-writing	9.5	0.66	9.0	0.58
Writing	9.0	0.73	8.1	0.64
		Total Score		
Control	19.0	1.15	18.3	1.07
Talk	21.2	1.25	18.6	1.17
Talk-writing	20.9	1.04	19.4	0.97
Writing	19.4	1.15	17.9	1.07

Bar graphs of the adjusted means for the essay questions from both immediate and delayed posttest have been included in Figure 2.

Figure 2

Bar Graphs of Adjusted Means for Essay Questions by Treatment





The adjusted means for the different groups on the essay questions would suggest that talk enhances the learning of science. The talk-only group and the talk-and-writing group ranked first and second, respectively, on the immediate posttest for all three knowledge measures (simple, integrated, and total scores). In comparison, the control group and the writing-only group ranked third or fourth on every knowledge measure. An analysis of the adjusted means for the essay questions from the delayed posttest suggests that writing, combined with talk, may enhance the retention of science learning over time. Although the talk-only and the talk-and-writing group still ranked first and second on all three posttest two knowledge measures, the talk-and-writing group now scored better than the talk-only group on both integrated knowledge and total knowledge measures based on essay questions. Talk thus appears to be important for acquiring knowledge, whereas writing is important for the retention of this learning over time.

4.0.2 *Concept Maps*

Table 13 gives the means and standard deviations for the pretest, the immediate posttest, and the delayed posttest results from the concept maps.

Table 13

Descriptive Statistics for the Concept Maps

Group	Pretest Mean	Pretest *SD*	Posttest 1 Mean	Posttest 1 *SD*	Posttest 2 Mean	Posttest 2 *SD*
			Simple Knowledge			
Control (*n* =10)	6.5	4.30	9.5	4.16	6.1	2.83
Talk (*n* =11)	7.4	3.85	9.9	4.11	6.9	3.47
Talk-writing (*n* =12)	6.4	4.50	11.5	4.56	10.1	4.97
Writing (*n* =10)	4.8	3.81	10.0	7.45	6.7	5.76
			Integrated Knowledge			
Control	4.3	4.54	7.7	2.90	5.3	2.47
Talk	6.0	3.38	8.9	3.72	6.2	3.28
Talk-writing	4.9	3.46	10.3	3.44	10.0	6.52
Writing	3.7	3.20	7.9	5.40	6.0	5.17
			Total Score			
Control	10.8	8.60	17.3	6.84	11.4	5.16
Talk	12.8	7.57	18.9	7.75	13.0	6.63
Talk-writing	11.3	7.93	21.9	7.92	20.1	11.18
Writing	8.5	7.00	17.9	12.78	12.8	10.68

An analysis of the mean concept mapping scores suggests that the two groups which involved writing (TW and W) appeared to improve considerably over time, while the other two groups did not improve as much. The control group and the talk-only group dropped in rank over time, while the talk-and-writing and the writing-only groups improved their rank placement over the same period.

Using group and gender as categorical variables, analysis of covariance (ANCOVA) was employed to explore the concept mapping data. Separate analyses were conducted for the immediate and delayed posttests. Table 14 shows the adjusted mean scores and the standard errors for the concept-mapping data, using the pretest scores as covariate each time, when both treatment and gender are displayed.

Table 14

Adjusted Means and Standard Errors for Concept Maps by Gender and Treatment

	Posttest 1				Posttest 2			
	Male		Female		Male		Female	
Group	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*
				Simple Knowledge				
Control	9.7	1.80	9.0	2.23	7.2	1.50	4.3	1.86
Talk	10.0	1.56	7.3	2.51	6.1	1.30	7.0	2.09
Talk-writing	13.1	1.78	9.8	1.79	11.4	1.49	8.7	1.49
Writing	13.6	1.77	6.8	2.25	10.4	1.48	3.0	1.88
				Integrated Knowledge				
Control	8.5	1.42	7.1	1.71	6.4	1.71	4.4	2.07
Talk	9.3	1.20	5.6	1.94	5.2	1.45	6.1	2.35
Talk-writing	10.6	1.37	9.9	1.37	10.1	1.66	9.8	1.66
Writing	10.2	1.37	5.9	1.72	9.4	1.65	2.4	2.08
				Total Knowledge				
Control	18.3	3.13	15.9	3.85	13.4	3.07	8.6	3.78
Talk	19.2	2.70	13.9	4.34	11.3	2.65	14.1	4.26
Talk-writing	23.7	3.08	19.6	3.09	21.5	3.02	18.4	3.03
Writing	23.7	3.07	12.7	3.87	19.8	3.01	5.4	3.80

On the immediate posttest, male students in the group which discussed the problem tasks with peers before writing explanations scored better on the concept maps than males in all other groups for both integrated and total knowledge. Female students in this same group ranked first on all three concept mapping measures. On the delayed posttest, students in the talk-and-writing group, both males and females, ranked first for all three knowledge measures. Moreover, female students in the group which just discussed the problem tasks without writing explanations ranked second for all three delayed posttest knowledge measures.

Table 15 shows the adjusted means and the standard errors for the concept maps, using the pretest scores as covariate each time, when just treatment is displayed.

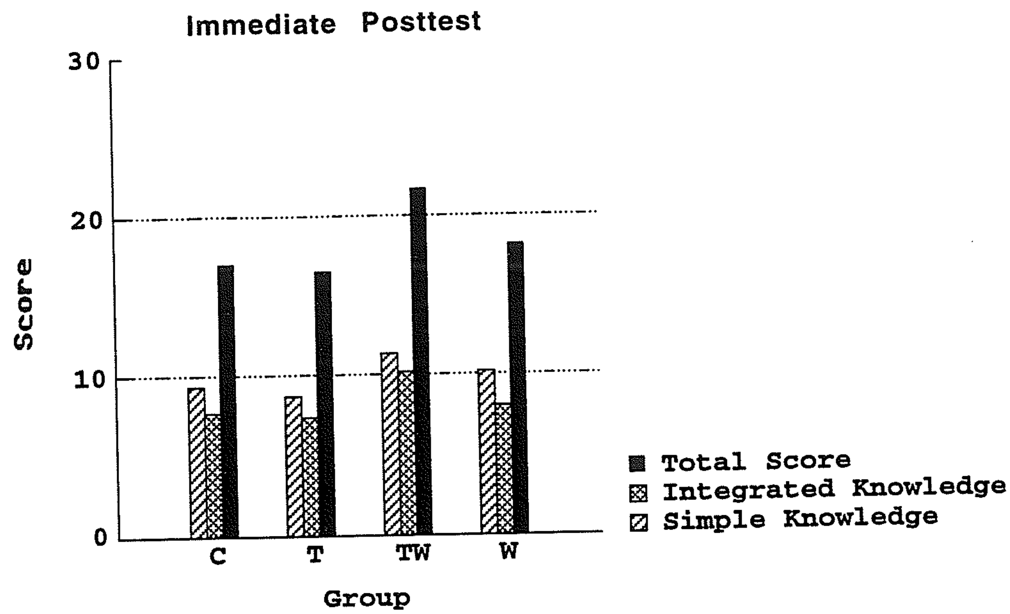
Table 15

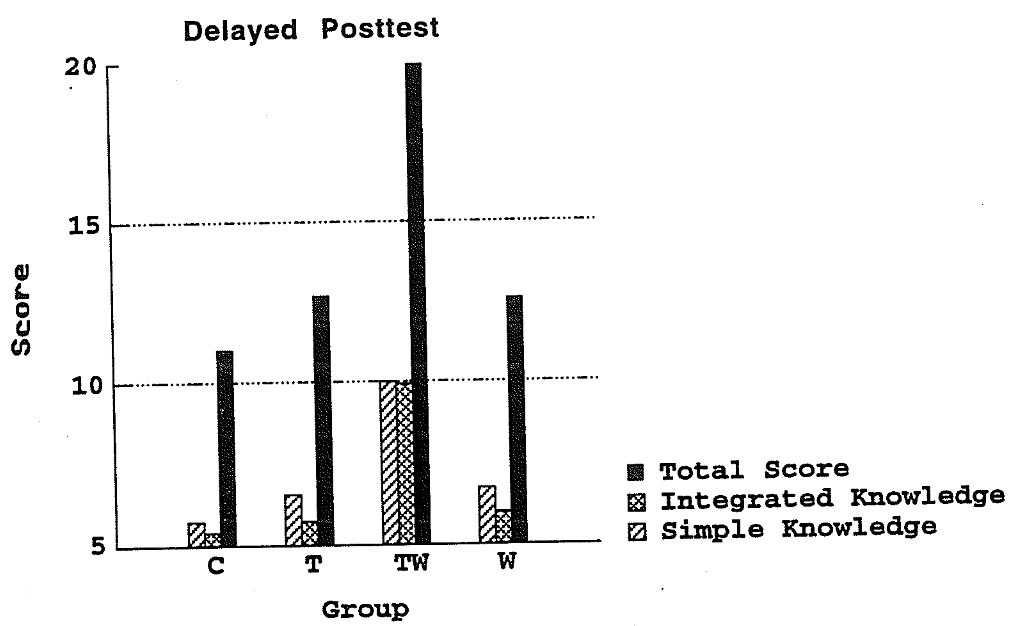
Adjusted Means and Standard Errors for Concept Maps by Treatment

	Posttest 1		Posttest 2	
Group	Adjusted Mean	*SE*	Adjusted Mean	*SE*
	Simple Knowledge			
Control	9.4	1.41	5.7	1.17
Talk	8.7	1.48	6.6	1.23
Talk-writing	11.5	1.25	10.0	1.05
Writing	10.2	1.44	6.7	1.20
	Integrated Knowledge			
Control	7.8	1.08	5.4	1.31
Talk	7.4	1.15	5.7	1.39
Talk-writing	10.2	0.97	10.0	1.17
Writing	8.0	1.10	5.9	1.33
	Total Score			
Control	17.1	2.43	11.0	2.38
Talk	16.5	2.55	12.7	2.50
Talk-writing	21.6	2.17	19.9	2.13
Writing	18.2	2.47	12.6	2.43

Figure 3

Bar Graphs of Adjusted Means for Concept Maps by Treatment





An analysis of the adjusted mean scores from the concept mapping suggests that writing may be important in learning science, particularly when combined with talk. The talk-and-writing group ranked first on all three knowledge measures for both immediate and delayed posttests. The writing group ranked second on all but the total knowledge score at the second posttest.

4.0.3 *Aggregate Scores*

Table 16 gives the means and standard deviations for the pretest, the immediate posttest, and the delayed posttest results obtained when the unweighted scores from the multiple choice tests, the essay questions and the concept maps are added together to give aggregate scores. These aggregate scores were calculated separately for simple knowledge and for integrated knowledge before being combined into aggregate total scores.

Table 16

Descriptive Statistics for the Aggregate Scores

Group	Pretest		Posttest 1		Posttest 2	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
			Simple Knowledge			
Control (*n* =10)	20.3	8.91	27.9	6.49	24.4	6.28
Talk (*n* =11)	24.6	6.35	33.1	6.62	28.5	6.48
Talk-writing (*n* =12)	20.2	8.15	32.4	5.38	29.6	6.54
Writing (*n* =10)	20.5	7.58	29.6	11.11	26.3	9.58
			Integrated Knowledge			
Control	17.7	9.35	26.6	5.99	23.9	5.94
Talk	21.3	6.86	32.0	7.22	27.0	6.05
Talk-writing	20.0	6.92	30.3	5.25	30.1	8.25
Writing	19.3	6.97	27.2	8.61	24.3	9.15
			Total Score			
Control	38.0	18.12	54.8	11.83	48.2	11.83
Talk	45.3	12.92	65.1	12.81	55.5	12.06
Talk-writing	40.2	14.91	62.7	9.99	59.7	14.15
Writing	39.8	14.29	56.8	19.23	50.6	18.33

An analysis of the mean aggregate scores over time suggests that both talk and writing can enhance learning in science. On the immediate posttest, the rank-order of the different groups for all three knowledge measures combined was as follows from first to last place: The talk-only group, the talk-and-writing group, the writing-only group, and the control group. On the delayed posttest, the rank-order of the groups was changed slightly with the talk-and-writing group now ranking first followed by the other three groups in the same order as before.

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 17 shows the adjusted means and the standard errors for aggregate scores by treatment and gender with the pretest scores being used as covariate each time.

Table 17

Adjusted Means and Standard Errors for Aggregate Scores by Gender and Treatment

	Posttest 1				Posttest 2			
	Male		Female		Male		Female	
Group	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*
				Simple Knowledge				
Control	30.1	2.25	26.4	2.71	27.0	1.97	22.3	2.38
Talk	31.4	1.93	29.7	3.12	25.3	1.69	29.2	2.74
Talk-writing	36.8	2.28	29.7	2.21	33.6	2.00	27.4	1.94
Writing	33.2	2.20	25.8	2.78	30.4	1.93	21.7	2.42
				Integrated Knowledge				
Control	29.4	2.10	25.3	2.50	27.1	1.90	22.5	2.26
Talk	32.3	1.76	27.9	2.90	26.2	1.59	24.8	2.62
Talk-writing	29.7	2.04	30.5	2.05	30.6	1.84	29.0	1.85
Writing	30.8	2.03	22.2	2.50	28.7	1.84	18.4	2.26
				Total Knowledge				
Control	59.5	3.90	51.5	4.67	54.0	3.56	44.6	4.27
Talk	63.7	3.29	58.7	5.37	51.3	3.01	55.3	4.91
Talk-writing	66.4	3.85	60.0	3.82	64.2	3.52	56.2	3.49
Writing	64.0	3.79	47.8	4.70	59.0	3.46	40.1	4.29

The trends across treatment groups were quite different for male and female students. For male students, 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 talking for male students. In comparison, female students 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 writing for female students.

Table 18 shows the adjusted means and the standard errors for the aggregate scores by treatment using the pretest scores as covariate each time.

Table 18

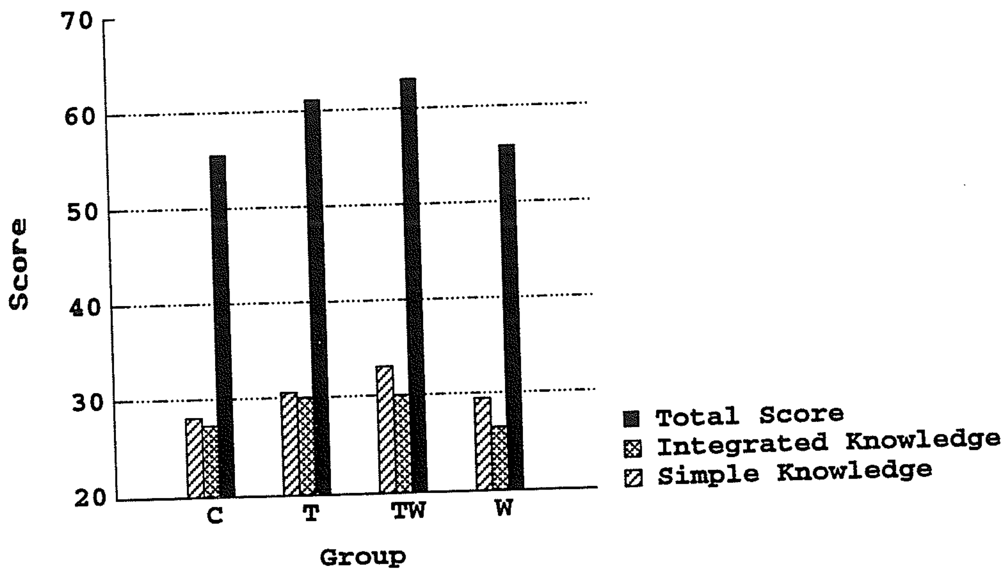
Adjusted Means and Standard Errors for Aggregate Scores by Treatment

	Posttest 1		Posttest 2	
Group	Adjusted Mean	*SE*	Adjusted Mean	*SE*
		Simple Knowledge		
Control	28.2	1.73	24.6	1.52
Talk	30.6	1.86	27.3	1.63
Talk-writing	33.3	1.55	30.5	1.36
Writing	29.5	1.74	26.1	1.53
		Integrated Knowledge		
Control	27.3	1.61	24.8	1.45
Talk	30.1	1.70	25.5	1.54
Talk-writing	30.1	1.43	29.8	1.30
Writing	26.5	1.60	23.6	1.45
		Total Score		
Control	55.5	2.99	49.3	2.73
Talk	61.2	3.17	53.3	2.90
Talk-writing	63.2	2.67	60.2	2.44
Writing	55.9	2.99	49.5	2.73

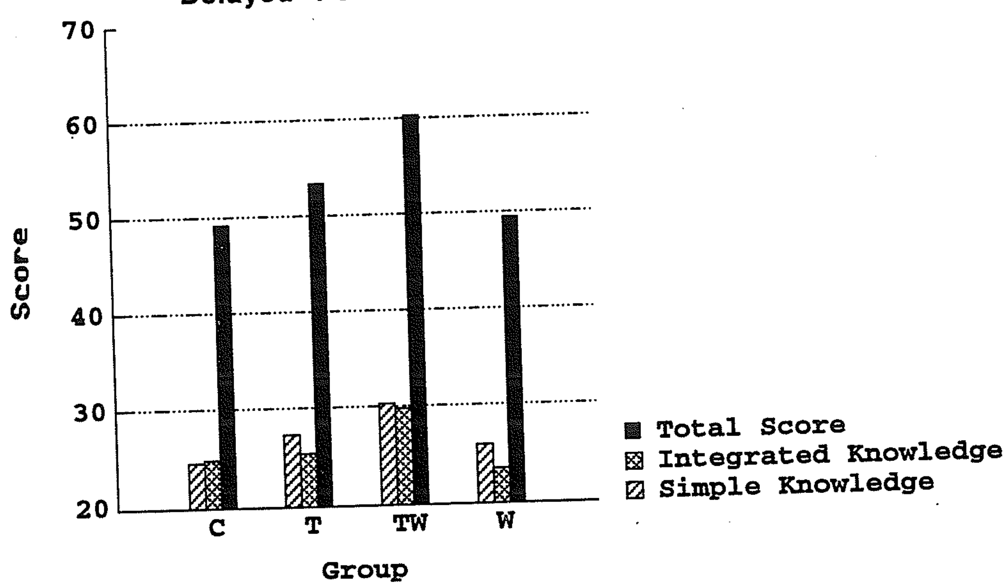
Figure 4

Bar Graphs of Adjusted Means for the Aggregate Scores by Treatment

Immediate Posttest



Delayed Posttest



An analysis of the adjusted means for the aggregate scores showed that the rank order of the talk-and-writing group and the talk-only group were again first and second, respectively. The rank order of the writing-only group and the control group were also unchanged except for the control group outperforming the writing-only group on the integrated knowledge aggregate scores at both the immediate and delayed posttests.

4.0.4 *Gender Differences*

Table 19 compares the adjusted means and standard errors by gender and treatment for the different measures and time-of-tests.

Table 19

Adjusted Means and Standard Errors for All Measures by Gender and Treatment

	Posttest 1				Posttest 2			
	Male		Female		Male		Female	
Knowledge	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*	*M adj*	*SE*
				Multiple Choice				
Simple	9.6	0.46	9.2	0.59	9.4	0.43	9.8	0.54
Integrated	11.3	0.34	10.8	0.43	11.3	0.33	10.4	0.42
Total	20.9	0.65	20.0	0.82	20.7	0.59	20.2	0.74
				Essay Questions				
Simple	11.6	0.37	10.6	0.47	10.7	0.34	9.7	0.44
Integrated	9.6	0.46	8.5	0.59	9.0	0.41	7.7	0.52
Total	21.2	0.72	19.0	0.92	19.7	0.67	17.4	0.86
				Concept Maps				
Simple	11.6	0.86	8.2	1.09	8.8	0.72	5.8	0.91
Integrated	9.7	0.67	7.1	0.84	7.8	0.80	5.7	1.02
Total	21.2	1.49	15.5	1.88	16.5	1.46	11.6	1.84
				Aggregate				
Simple	32.9	1.07	27.9	1.35	29.1	0.94	25.1	1.18
Integrated	30.5	0.99	26.5	1.26	28.1	0.90	23.7	1.14
Total	63.4	1.84	54.5	2.33	57.1	1.68	49.0	2.13

Female students scored higher than male students on the simple knowledge delayed posttest multiple choice measure. In comparison, the 23 other measures all favoured male students. Gender thus appears to be an important mediating variable. However, gender-based comparisons across groups were not quite as straightforward. Female students in the groups which involved talking with peers, either with or without writing, generally performed better on the various measures than those students in groups which did not include talk with peers. Nonetheless, talk with peers, either with or without writing, also appeared to benefit male students.

One of the questions which will be addressed later in this chapter is whether or not a gender-treatment interaction was present in the study. An examination of the rank-order placements for the various measures suggests that such an interaction may be present in the study. Table 20 compares the number of first-place rankings across all four measures and both posttests by gender and treatment.

Table 20

Comparison of the Total Number of First-Place Rankings across All Measures by Gender and Treatment

	Immediate Posttest			Delayed Posttest		
Group	Male (M)	Female (F)	M and F	Male	Female	M and F
Control	0	0	0	0	0	0
Talk	5	5	10	2	5	7
Talk-writing	5.5[a]	7	12.5	10	7	17
Writing	1.5	0	1.5	0	0	0

Note . [a] 0.5 signifies that the group was tied for first place with another group (Writing) for one measure.

The control group never ranked first on the basis of adjusted means for any of the measures, whereas students who just wrote in response to the problem task placed first for at least one measure and were tied for another, both at the immediate posttest. The balance of the highest rankings were shared by students who just discussed the problem tasks and those who wrote in response to the tasks after discussing them with peers. Differences between males and females were apparent on the delayed posttest. Male students in the groups which included both talk and writing generally performed better on the various measures than other male students (they ranked first for 10 out of 12 posttest-2 measures). In comparison, the female students in the two groups which used talk, either with or without writing, ranked first for all of the twelve measures, but the rankings were equally divided between these two groups. On the basis of the rank-order placements shown in table 20, the presence of gender-treatment interactions in the study is probable.

4.1 Hypothesis Testing

4.1.0 Aggregate Scores

The three hypotheses which guided the exploratory study were all based on aggregate scores. Aggregate scores were defined as the total of the unweighted scores from the multiple choice, essay question and concept mapping measures. Separate aggregate scores were calculated for both simple and integrated knowledge, then these were combined to also give an aggregate total score. The first hypothesis stated that there would be no significant differences in mean aggregate simple knowledge scores that could be attributed to:

a) treatment (C / T / TW / W);

b) gender; and

c) interaction of treatment and gender.

Simple Knowledge

For all of the analyses of covariance, both gender and treatment were used as categorical variables. Using the simple knowledge aggregate scores from posttest one as dependent variable and the simple knowledge aggregate scores from the pretest as covariate gave the results summarized in Table 21.

Table 21

Analysis of Covariance Summary for Simple Knowledge Aggregate Scores at the Immediate Posttest

Source	Sum-of-Squares	*df*	*MS*	*F* -ratio	*p*
Sex	238.274	1	238.274	8.309	.007
Group	149.470	3	49.823	1.737	.178
Sex*Group	55.186	3	18.395	0.641	.594
Pretest	885.586	1	885.586	30.881	<.0005
Error	975.022	34	28.677		

The analysis of covariance using the simple knowledge aggregate scores suggested that few differences existed at the immediate posttest with the exception of an effect for gender (p = .007). The scores for male students were significantly higher than those for female students.

Using the simple knowledge aggregate scores from posttest 2 as dependent variable and the simple knowledge aggregate scores from the pretest as covariate gave the results summarized in Table 22.

Table 22

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

Source	Sum-of-Squares	*df*	*MS*	*F* -ratio	*p*
Sex	150.490	1	150.490	6.808	.013
Group	201.961	3	67.320	3.046	.042
Sex*Group	204.372	3	68.124	3.082	.040
Pretest	835.243	1	835.243	37.787	<.0005
Error	751.538	34	22.104		

The analysis at the delayed posttest suggested that differences exist between the sexes (p = .013), between the groups (p = .042), and that an interaction exists between gender and treatment (p = .040). *Post hoc* analysis for pairwise differences using Fisher's Least-Significant-Difference Test suggested that girls who discussed the problems with peers showed better retention of facts and concepts over time than those who just wrote in response to the problem tasks (T>W, p = .05) and that boys who first discussed the problem tasks before writing individual responses showed better retention of simple knowledge than both groups of boys who simply completed descriptive tasks (TW>C, p = .022) or who just discussed the problems with other students (TW>T, p = .004). 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 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).

Integrated Knowledge

The second hypothesis stated that there would be no significant differences in mean aggregate integrated knowledge scores that could be attributed to:

a) treatment (C / T / TW / W);

b) gender; and

c) interaction of treatment and gender.

Analysis of covariance using both gender and treatment as categorical variables, with the integrated knowledge aggregate scores from posttest 1 as dependent variable and the integrated knowledge aggregate scores from the pretest as covariate, gave the results summarized in Table 23.

Table 23

Analysis of Covariance Summary for Integrated Knowledge Aggregate Scores at the Immediate Posttest

Source	Sum-of-Squares	*df*	*MS*	*F*-ratio	*p*
Sex	156.420	1	156.420	6.367	.016
Group	100.631	3	33.544	1.365	.270
Sex*Group	115.861	3	38.620	1.572	.214
Pretest	636.460	1	636.460	25.907	<.0005
Error	835.285	34	24.567		

The analysis suggested that an effect for gender was present at the immdiate posttest (p = .016). Once again, the aggregate integrated knowledge scores for males were superior to those for females.

Analysis of covariance using the integrated knowledge aggregate scores from posttest 2 as dependent variable and the integrated knowledge aggregate scores from the pretest as covariate gave the results summarized in Table 24.

Table 24

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

Source	Sum-of-Squares	*df*	*MS*	*F* -ratio	*p*
Sex	189.814	1	189.814	9.443	.004
Group	243.835	3	81.278	4.044	.015
Sex*Group	122.120	3	40.707	2.025	.129
Pretest	1054.790	1	1054.79	52.475	<.0005
Error	683.429	34	20.101		

The analysis at the delayed posttest suggested that differences exist between the sexes ($p = .004$) and between the groups ($p = .015$). The mean scores for male students was again superior to that of female students. *Post hoc* analysis for pairwise differences using Fisher's Least-Significant-Difference Test suggested that students who had talked with peers before writing showed better retention of integrated knowledge over time than all three groups of students: those who had simply completed descriptive tasks (TW>C, $p = .015$); those who had just discussed the problems with peers (TW>T, $p = .039$); and those who had simply written an explanation without prior discussion with peers (TW>W, $p = .003$).

Total Knowledge

The third hypothesis stated that there would be no significant differences in mean aggregate total knowledge scores that could be attributed to:

a) treatment (C / T / TW / W);

b) gender; and

c) interaction of treatment and gender.

Analysis of covariance using both gender and treatment as categorical variables, the posttest 1 total knowledge aggregate scores as dependent variable, and the pretest total knowledge aggregate score as covariate gave the results summarized in Table 25.

Table 25

Analysis of Covariance Summary for Total Knowledge Aggregate Scores at the Immediate Posttest

Source	Sum-of-Squares	*df*	*MS*	*F* -ratio	*p*
Sex	749.721	1	749.721	8.780	.006
Group	454.372	3	151.457	1.774	.171
Sex*Group	166.248	3	55.416	0.649	.589
Pretest	3057.439	1	3057.439	35.804	<.0005
Error	2903.378	34	85.393		

Similar to the two other analyses of posttest one aggregate scores, an effect for gender was observed with male students again scoring higher than female students (p = .006).

Analysis of covariance using both gender and treatment as categorical variables, the posttest 2 total knowledge aggregate scores as dependent variable, and the pretest total knowledge aggregate score as covariate gave the results summarized in Table 26.

Table 26

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

Source	Sum-of-Squares	*df*	*MS*	*F* -ratio	*p*
Sex	625.440	1	625.440	8.774	.006
Group	857.144	3	285.715	4.008	.015
Sex*Group	595.280	3	198.427	2.784	.056
Pretest	3707.095	1	3707.095	52.005	<.0005
Error	2423.635	34	71.283		

The analysis at the delayed posttest suggested that differences exist between the sexes (p = .006) and between the groups (p = .015). *Post hoc* analysis of pairwise differences using Fisher's Least-Significant-Difference Test suggested that students who discussed with peers before writing showed better retention of knowledge, generally, than both students who completed simple descriptive tasks (TW>C, p = .005) and students who just wrote without any peer support (TW>W, p = .006). Although the mean total knowledge aggregate score of the talk-and-writing group was also greater than that of the talk-only group, the differences were not statistically significant.

4.1.1 *Exploring the Data*

Since the purpose of this exploratory study was to investigate trends and relationships in the data that could serve as a template for new research, each of the measures was considered separately in the following analysis.

Multiple Choice Tests

Analyses of covariance using group and gender as categorical variables gave the results for the multiple choice tests that are summarized in Table 27.

Table 27

Analyses of Covariance Using Group and Gender as Categorical Variables: Summary Table of *p* - Values for Multiple Choice Tests

	Posttest 1			Posttest 2		
Score	Sex	Group	Sex*Group	Sex	Group	Sex*Group
Simple Knowledge	.628	.058	.421	.572	.266	.170
Integrated Knowledge	.362	.003	.120	.084	.016	.092
Total	.367	.007	.639	.607	.071	.110

The analyses of covariance for the multiple choice data from the immediate posttest suggests that differences do exist among the groups for both integrated knowledge ($p = .003$) and total score ($p = .007$). *Post hoc* analysis of pairwise differences using Fisher's Least-Significant-Difference Test suggests that students who just discussed the problem tasks with peers had recalled more integrated knowledge than all other groups when tested immediately after the completion of instruction. These students (T) were able to make connections among the concepts which had been taught better than students who had received simple descriptive tasks (T>C, $p = .032$), students who had just written in response to the explanatory tasks (T>W, $p < .0005$), and students who had discussed the problems with peers before writing (T>TW, $p = .012$). On the basis of total scores, differences appeared at the immediate posttest ($p = .007$) with students who had simply discussed the problem tasks learning more, generally, then students who had completed simple descriptive tasks (T>C, p = .008) and those who had just written in response to the explanatory tasks without any peer support (T>W, p = .001).

However, other differences appeared after six weeks on the delayed posttest. The results based on integrated knowledge again differentiated among the groups (p = .016). Analysis of pairwise contrasts suggested that students who showed the worst retention of integrated knowledge about ecology over time were those students who had just written explanations without prior discussions of the tasks with peers. Even students in the control group who had individually completed descriptive tasks had better integrated knowledge than this group (C>W, p ≐ .01). Moreover, the groups who had benefited from discussions with peers also showed better retention over time of integrated knowledge compared with the group in which students just wrote explanations alone (T>W, p = .003; TW>W, p = .049).

Essay Questions

Analyses of covariance using group and gender as categorical variables gave the results for the essay questions that are summarized in Table 28.

Table 28

Analyses of Covariance Using Group and Gender as Categorical Variables: Summary Table of *p* - Values for the Essay Questions

	Posttest 1			Posttest 2		
Score	Sex	Group	Sex*Group	Sex	Group	Sex*Group
Simple Knowledge	.098	.475	.754	.089	.870	.614
Integrated Knowledge	.137	.444	.749	.063	.623	.896
Total	.079	.446	.747	.047	.732	.899

The analyses of covariance for the essay questions suggests that the groups do not differ significantly from one another regardless of the time of test. Overall, there was only one difference, one based on gender, that was significant: Males outperformed the females for total knowledge at the delayed posttest (p = .047).

Concept Maps

Analyses of covariance using group and gender as categorical variables gave the results for the concept maps that are summarized in Table 29.

Table 29

Analyses of Covariance Using Group and Gender as Categorical Variables: Summary Table of *p* - Values for the Concept Maps

	Posttest 1			Posttest 2		
Score	Sex	Group	Sex*Group	Sex	Group	Sex*Group
Simple Knowledge	.021	.506	.541	.014	.039	.138
Integrated Knowledge	.022	.214	.601	.110	.036	.181
Total	.023	.399	.660	.046	.032	.110

The analyses of covariance for the concept maps suggests that differences do exist between males and females. At the immediate posttest, differences appeared for simple knowledge (p = .021), integrated knowledge (p = .022), as well as for the total score (p = .023). At the delayed posttest, gender effects were observed for simple knowledge (p = .014) and for total score (p = .046). All of these differences favoured males.

Differences among the treatment and control groups were observed for all three knowledge measures at the delayed posttest: Simple knowledge (p = .039), integrated knowledge (p = .036), and total score (p = .032). In every case, differences favoured the talk-and-writing group over the other three groups. For simple knowledge, students who discussed the problem tasks with peers before writing out their explanations showed better retention of simple facts and concepts over time than students who had received simple descriptive tasks to complete alone (TW>C, $p = .009$), or students who had just discussed the explanatory tasks with peers (TW>T, $p = .038$), or students who had just written their explanations without talking with peers (TW>W, $p = .044$).

Students who talked and wrote about the problem tasks also showed better retention of integrated knowledge over time than all three other groups of students. Pairwise comparisons suggested that the talk-and-writing group retained more integrated knowledge than the control group (TW>C, $p = .013$), the group in which students just discussed possible explanations for the problems (TW>T, $p = .024$), and the group in which students just wrote in response to the problem tasks (TW>W, $p = .029$). Moreover, students who talked with peers before writing their explanations also showed better retention of knowledge, generally, than all other groups: TW>C, $p = .008$; TW>T, $p = .034$; TW>W, $p = .029$.

Table 30 presents a summary of the analysis of covariance using group and sex as categorical variables for all measures at both the immediate and delayed posttests.

Table 30

Summary Table of *p* - Values for Factorial ANCOVA Using Group and Sex as Categorical Variables

	Immediate Posttest			Delayed Posttest		
Knowledge	Sex	Group	Sex*Group	Sex	Group	Sex*Group
			Multiple Choice			
Simple	.628	.058	.421	.572	.266	.170
Integrated	.362	.003	.120	.084	.016	.092
Total	.367	.007	.639	.607	.071	.110
			Essay Questions			
Simple	.098	.475	.754	.089	.870	.614
Integrated	.137	.444	.749	.063	.623	.896
Total	.079	.446	.747	.047	.732	.899
			Concept Maps			
Simple	.021	.506	.541	.014	.039	.138
Integrated	.022	.214	.601	.110	.036	.181
Total	.023	.399	.660	.046	.032	.110
			Aggregate Scores			
Simple	.007	.178	.594	.013	.042	.040
Integrated	.016	.270	.214	.004	.015	.129
Total	.006	.171	.589	.006	.015	.056

4.2 Addressing the Research Questions

The exploratory study was guided by six questions. In this section, each of these research questions will be addressed in the light of the data collected in the study.

4.2.0 *Writing or Talking?*

Question one asked: Does writing an explanation for a scientific phenomenon (the writing mode, or W) enhance learning more than a discussion with other students on the same topic, but without writing (the talking mode, or T)? The answer to this first question appears to be negative. None of the pairwise contrasts after the analysis of covariance for each of the measures and time-of-test favoured the writing-only group over the talk-only group. In fact, the opposite situation seemed to prevail with students who simply discussed the explanatory tasks acquiring more integrated knowledge and more knowledge, generally, as measured by the multiple choice tests, and showing better retention of this knowledge over time than those students who just wrote in response to the explanatory tasks, even when students were asked to write as a measure of learning.

One other pairwise contrast favoured the talk-only group over the writing-only group. The interaction between gender and treatment was significant for the simple knowledge aggregate measure at posttest two ($p = .04$). Pairwise contrasts suggest that girls who just discussed the explanatory tasks with peers showed better retention of facts and concepts over time than those girls who simply wrote in response to the same tasks. However, since these contrasts all employed the Least Significant Difference Test (LSD), they should be interpreted cautiously. Nonetheless, the data appear to suggest that writing alone is not better than talking with other students in small groups to learn science concepts, either measured immediately after completing the ecology unit or measured later for determining knowledge retention over time.

4.2.1 *Talk-and-Writing or Just Writing?*

Question two asked: Does writing an explanation after a discussion with other students (the combined talking and writing mode, or TW) enhance learning more than individually writing an explanation of the same phenomenon, but without talking (W)? For this question, the answer seems to depend on the time-of-test. Although none of the analyses at the immediate posttest showed any differences between these two treatment groups, many of the delayed posttest measures suggest that talking-and-writing does indeed enhance learning more over time than writing alone. For the multiple choice data, only the measure of integrated knowledge showed any significant differences between these two groups. Further, none of the essay question measures showed any differences between these two groups. In comparison, pairwise contrasts following the analyses of covariance on all three knowledge measures for both the concept maps and the aggregate scores all favoured the talking-and-writing group. In summary, seven of twelve delayed posttest measures suggest that students who wrote explanations after initially discussing the problem tasks with peers showed better retention of knowledge about ecology, generally, than students who just wrote explanations alone.

4.2.2 *Talk-and-Writing or Just Talking?*

Question three asked: Does writing an explanation after a discussion with other students (the TW mode), enhance learning more than a discussion of the same topic with other students (T)? The integrated knowledge measure based on the multiple choice test at posttest one suggests that students who just discussed the problem tasks acquired more integrated knowledge about ecology initially than those students who both talked and wrote about the same explanatory tasks. However, all three knowledge measures using concept maps at the delayed posttest favoured the use of both talk and writing combined over just talking. For the aggregate scores, the integrated knowledge measure at the delayed posttest suggests that students who talked and wrote in response to the explanatory tasks showed better retention of integrated knowledge than students who just talked about the same problems. One other pairwise contrast, which showed an interaction between gender and treatment, also favoured the talk-and-writing group. For the aggregate simple knowledge measure at the delayed posttest, boys who used both talk and writing showed better retention of facts and concepts over time than other boys who just talked about the problem tasks. Moreover, although the omnibus F - test for total aggregate knowledge was only marginally significant (p = .056), pairwise comparisons suggest that boys who discussed the problem tasks with peers before writing their explanations showed better retention of knowledge, generally, than boys who just discussed their explanations with peers. Writing appears to exert an effect, but only over time. The measures of long-term learning appear to be enhanced by writing, but only after initially talking about the problem tasks with peers.

4.2.3 *Comparisons with the Control Group*

Question four asked: Is each of these treatments (T, W, TW) more effective than a control group which has been assigned descriptive and other simple content learning tasks (fill-in-the-blanks, true-or-false, etc.)? Table 31 summarizes the results of pairwise contrasts for all twelve measures at both immediate and delayed posttests. For instance, significant differences were observed in the omnibus F - test for the integrated knowledge multiple-choice measure at the delayed posttest. *Post hoc* pairwise contrasts suggested that students in the talk-and-writing group showed better retention of integrated knowledge over time than those students who just wrote in response to the same explanatory tasks (TW>W, p = .049).

Table 31

Summary Table of *p* - values for Pairwise Contrasts Based on Immediate and Delayed Knowledge Measures for Both Groups Using Talk

		Talk-and-Writing			Talk-only		
Knowledge	Posttest	>C[a]	>T	>W	>C	>TW	>W
				Multiple Choice			
Simple	1						
	2						
Integrated	1				.032	.012	<.0005
	2			.049			.003
Total	1				.008		.001
	2						
				Essay Questions			
Simple	1 / 2						
Integrated	1 / 2			Note: No differences			
Total	1 / 2						
				Concept Maps			
Simple	1						
	2	.009	.038	.044			
Integrated	1						
	2	.013	.024	.029			
Total	1						
	2	.008	.034	.029			
				Aggregate			
Simple	1						
	2	.007		.038			
Integrated	1						
	2	.015	.039	.003			
Total	1						
	2	.005		.006			

Notes. [a]">C" signifies that the talk-and-writing group outperformed the control group.

Only one pairwise contrast favoured the control group over another treatment, specifically the writing-only group on integrated knowledge at the delayed posttest (C>W, p = .01). All other pairwise contrasts favoured one or more of the treatment groups over other groups. The talk-and-writing group outperformed the control group on all six concept map and aggregate knowledge measures at the delayed posttest. This suggests that students who discussed the problems with peers before writing their explanations showed better retention of knowledge, both simple and integrated, over time than the control group in which students just completed descriptive tasks alone. In contrast, pairwise comparisons showed that the talk-only group outperformed the control group on only two knowledge measures, both of them based on the multiple choice tests at the immediate posttest. This suggests that students who just discussed the problems with peers acquired more knowledge about ecology than students completing simple descriptive tasks. However, merely discussing the explanatory tasks did not enhance student retention of this learning over time. None of the pairwise contrasts favoured the writing-only group over the control group. In many respects, this treatment group was no different than the control group. It should be mentioned that the control group did receive meaningful and relevant learning activities to complete during the problem sessions. However, activities were designed so that students would not be encouraged to make connections among the ecological concepts. Nonetheless, the students did learn about ecology during these sessions. Some of these students may have been able to make connections in the absence of explicit tasks designed to foster the integration of ecology concepts. Several students in the control group reported that they had enjoyed the activities which had been assigned during these classes.

One other approach to answering this question might be to rank-order the different groups on the basis of the adjusted mean total knowledge scores for each of the measures. Table 32 summarizes the rank-order of each group for every measure at both the immediate and the delayed posttests.

Table 32

Rank Order of Treatment and Control Groups for each Measure and Time-of-Test on Total Knowledge Scores

	Group							
	Control		Talk		TW		Writing	
Measure	PO1	PO2	PO1	PO2	PO1	PO2	PO1	PO2
Multiple Choice	3	3	1	1	2	2	4	4
Essay Questions	4	3	1	2	2	1	3	4
Concept Maps	3	4	4	2	1	1	2	3
Aggregate	4	4	2	2	1	1	3	3

Note. Number 1 represents the highest ranking score and 4 the lowest ranking score. "PO1" and "PO2" represent immediate and delayed posttests, respectively.

The rank-ordering shows that the control group was in last place for half of the measures. In comparison, the writing group was in third place for half of the measures. With the exception of the concept maps at the immediate posttest (for which the writing-only group ranked second), these two groups ranked either third or fourth for all of the other measures. In contrast, the talk-only group and the talking-and-writing group ranked either first or second for all but one measure for which the talk-only group ranked fourth on the concept maps at the immediate posttest.

The collapsing of the four treatment and control groups into two clusters was also supported by user-defined contrasts following factorial ANCOVA. Table 33 summarizes the *p* - values for contrasts between the talk-and-writing and talk-only groups taken together in one cluster, and the writing-only and the control groups in a second cluster. Using abbreviations for the treatment and control groups, this contrast can be represented as TW+T≠W+C.

Table 33

Summary Table of *p* - values for

TW+T≠W+C Contrast after Factorial ANCOVA

Measure	Posttest 1	Posttest 2
	Multiple Choice Test	
Simple Knowledge	.009	.065
Integrated Knowledge	.008	.097
Total Score	.002	.028
	Questions	
Simple Knowledge	.143	.446
Integrated Knowledge	.175	.378
Total Score	.111	.381
	Concept Map	
Simple Knowledge	.841	.082
Integrated Knowledge	.396	.109
Total Score	.550	.064
	Aggregate	
Simple Knowledge	.088	.026
Integrated Knowledge	.057	.023
Total score	.037	.011

At the immediate posttest, differences were apparent between the two groups using talk (TW+T), on the one hand, and the two groups in which students worked alone (W+C), on the other hand. All three multiple choice measures were significantly different at an alpha level of .05 and all favoured the students using talk: The *p* - values were .009,

.008, and .002 for simple, integrated and total knowledge scores, respectively. One of the aggregate measures, specifically the total score, was also significant with a *p* - value of .037. Two other aggregate measures were marginally significant with alpha set at .10. At the delayed posttest, four measures were significantly different for treatment effects. The contrast using the total score for the multiple choice gave a *p* - value of .028. Furthermore, all three aggregate measures also were significantly different: The *p* - values were .026, .023, and .011, respectively, for simple, integrated, and total knowledge scores. Moreover, except for the essay questions, all remaining contrasts were marginally significant.

Since the aggregate total scores, which were significant at both the immediate and delayed posttest, reflect the domain knowledge of students in ecology as measured by an overall score for multiple choice tests, essay questions, and concept maps, one might argue that talking with peers enhances not only the immediate acquisition of ecology knowledge after instruction, but also the retention of this learning over time.

Interestingly, if the talk-and-writing group alone is contrasted with the other three treatment and control conditions taken together as one group, then other differences are significant, particularly at the delayed posttest. Table 34 compares the *p* - values for contrasts between the talk-and-writing group in one cluster and the the other three groups in a second cluster (TW≠T+W+C).

Table 34

Summary Table of *p* - Values for

TW≠T+W+C Contrast after Factorial ANCOVA

Knowledge	Immediate Posttest	Delayed Posttest
	Multiple Choice	
Simple	.194	.638
Integrated	.583	.931
Total	.696	.692
	Essay Questions	
Simple	.444	.705
Integrated	.487	.200
Total	.405	.319
	Concept Maps	
Simple	.182	.005
Integrated	.039	.004
Total	.102	.004
	Aggregate Scores	
Simple	.048	.010
Integrated	.226	.002
Total	.086	.003

Using a TW≠T+W+C contrast, only two measures were significantly different at the immediate posttest: the integrated knowledge concept mapping measure (p = .039) and the simple knowledge aggregate measure (p = .048). In comparison, six contrasts were signficantly different at the delayed posttest. All three concept mapping measures were signficantly different: simple knowledge (p = .005), integrated knowledge (p = .004), and total score (p = .004). As well, the contrasts for all three aggregate measures also differentiated the talk-and-writing group from the other three groups: simple knowledge (p = .01), integrated knowledge (p = .002), and total score (p = .003).

The data suggest that the three treatment groups (T, W, and TW) are not all better than the control group which had been assigned descriptive tasks. The control group and the writing-only group appeared to be quite similar in terms of science learning. Similarly, the two groups using talk also appeared to be quite similar in their learning patterns, particularly on the initial posttest which might be considered a measure of short-term retention. These two groups appeared to do better on four of the twelve measures of science learning when compared with the writing-only and control groups. Moreover, two other measures approached significance. On the delayed posttest, which might be considered a measure of long-term retention, the talk-and-writing group did better than the other three groups on six of the twelve science learning measures. This suggests that writing may be an important tool for the longer-term retention of science knowledge, but only if students are given the opportunity to first construct this knowledge while interacting with other students. Moreover, since the two contrasts (TW+T≠W+C and TW=T+W+C) highlighted three measures which were significantly different in one or both contrasts (the aggregate measures) with six measures which were significantly different in one or the other (the three multiple choice measures in the TW+T≠W+C contrast at posttest one, and the three concept mapping measures in the TW≠T+W+C contrast at posttest two), one can argue that this supports the idea of talk being important for personally constructing knowledge, whereas writing is important for the retention of science learning over time.

4.2.4 *Interactions with Gender*

Question five asked: Is there an interaction between these different groups using writing, and talking and gender? Gender differences were apparent for half of the 24 measures when analyses of covariance using group and sex as categorical variables were employed to explore the data. None of the multiple choice measures and only one of the essay question measures showed differences between the sexes. However, all but one concept mapping measure, integrated knowledge at the delayed posttest, showed significant differences between females and males with alpha set at .05. Moreover, all six aggregate measures showed gender differences. Every difference favoured male students over female students. Despite these differences, the statistical evidence for interactions between treatment and gender was not very strong. Only one measure, the aggregate simple knowledge score, suggested that differences due to a gender-treatment interaction were present ($p = .04$). The pairwise comparisons suggested that females in the talk-only group had learned more than those in the writing-only group ($p = .05$). *Post hoc* comparisons also suggested that males in the talk-and-writing group learned more than those in the talk-only group ($p = .004$) and those in the control group ($p = .022$). Nonetheless, the F -values exploring gender-treatment interactions for two other measures at the delayed posttest were marginally significant with alpha set at .10: The multiple choice measure for integrated knowledge ($p = .092$); and the aggregate total knowledge score ($p = .056$). In addition, the contrasts between females in both groups using talk with those in the combined writing-only and control groups (TW+T≠W+C) suggested that differences exist at the delayed posttest for the multiple choice integrated knowledge measure ($p = .036$), as well as for two aggregate measures, simple knowledge ($p = .012$) and total score ($p = .003$). Two of these same delayed posttest measures differentiated females in the talk-and-writing group from females in the three other groups taken together (TW≠T+W+C): Simple knowledge ($p = .01$) and total score ($p = .005$). These contrasts suggest that female students may learn more science when they talk with peers about meaningful problems which apply the target concepts. They may also show better retention of this learning over time when they are given the opportunity to write explanations after these peer discussions.

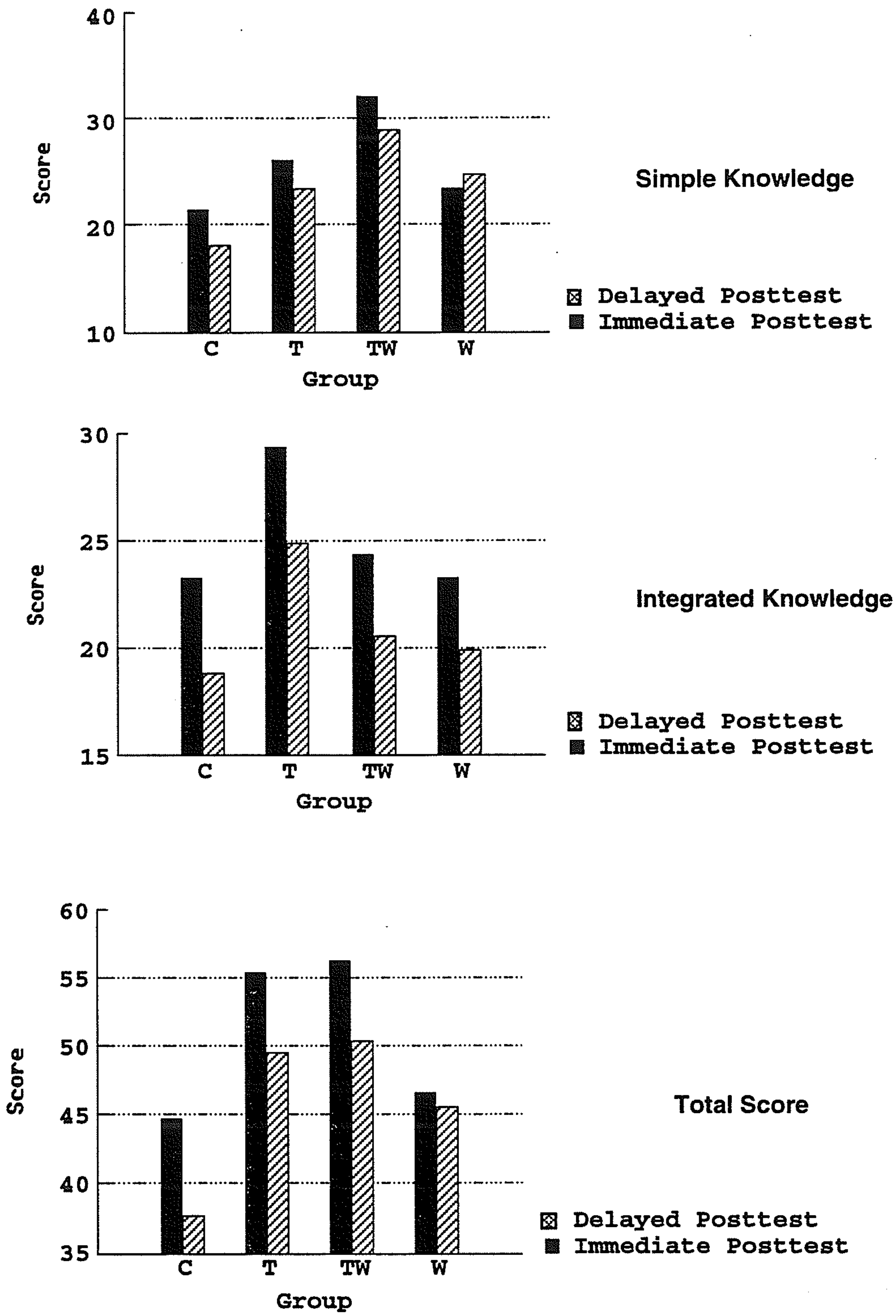
In summary, no conclusions are warranted regarding whether or not gender-treatment interactions were operating during the study. However, since differences are present between the scores of males and females on half of all knowledge measures, and since several analyses confirm the presence of possible interactions, the question merits further exploration.

4.2.5 *Interactions with Ability*

Question six asked: Is there an interaction between these different groups using writing, and talking and student ability? The sample size of 43 students was not adequate to objectively answer this question. Nonetheless, the data were examined to determine whether or not trends were discernible. The objective of this analysis was to ascertain whether or not a follow-up study which would specifically address this variable had any merit. Figure 5 compares the adjusted mean aggregate scores on both immediate and delayed posttests of low ability students in the four treatment and control groups. The adjusted mean scores for simple, integrated, and total knowledge have all been represented. Ability was defined as the achievement of students in the grade seven science course. Students scoring in the upper or lower quartile were considered to be of high or low ability, respectively. Those students who were in between these two ability groups were judged to be of average ability. The placement of students using this operational definition was corroborated by the teacher's perception of student abilities for achievement in the grade eight science course.

Figure 5

Bar Graphs of Adjusted Mean Aggregate Scores for Low Ability Students



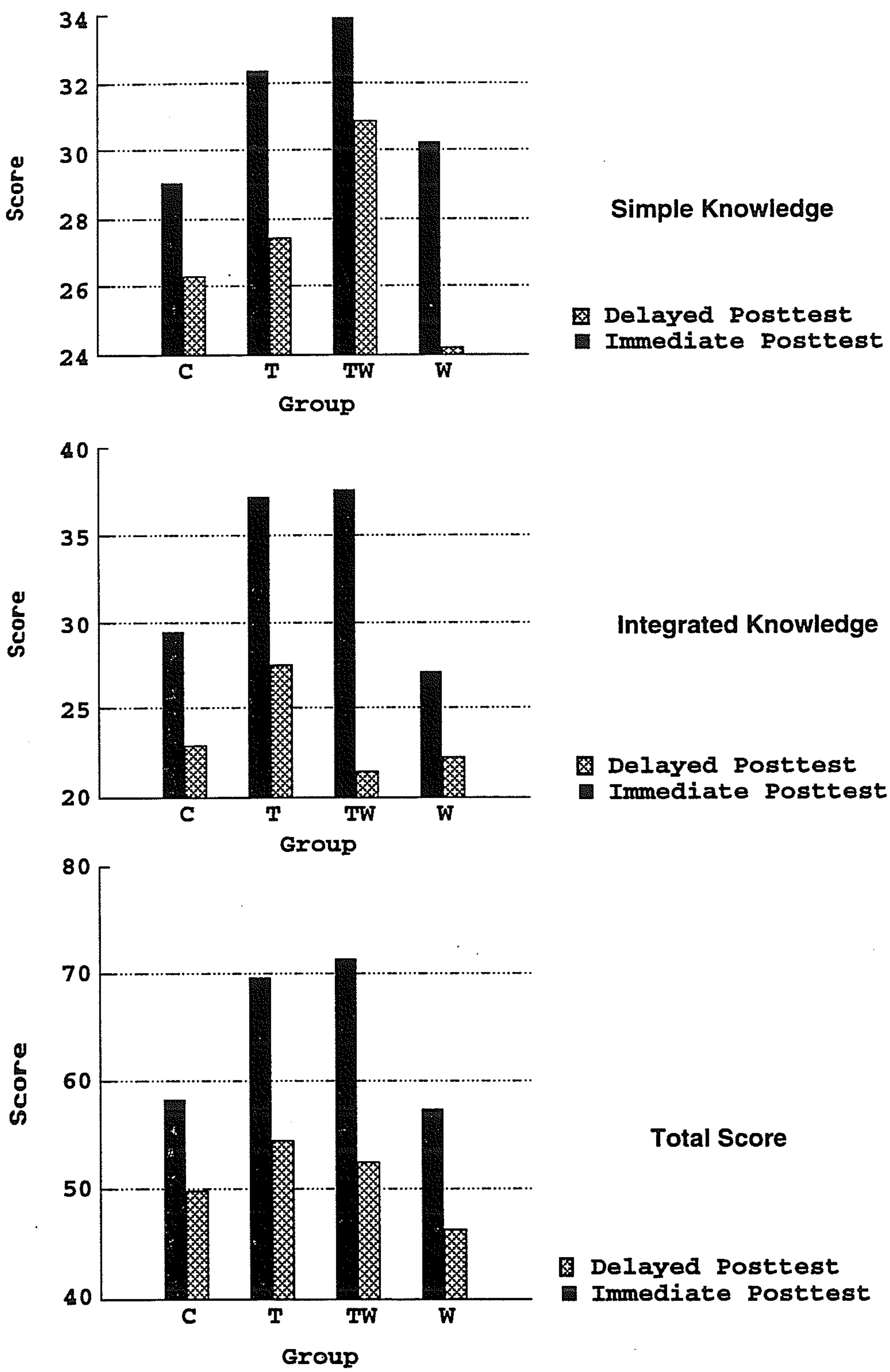
Note . Number of students in each group: C=2; T=1; TW=3; and W=3.

The low-ability students in the groups using talk appeared to learn more than low-ability students in the control and writing-only groups. Students in the talk-only and the talk-and-writing groups ranked either first or second for the mean adjusted integrated and total knowledge aggregate scores on both immediate and delayed posttests. For the mean adjusted simple knowledge aggregate scores, students in the talk-and-writing group ranked first on both posttests. For this same measure, students in the talk-only group ranked second on the immediate posttest, whereas those students in the writing-only group ranked second on the delayed posttest. In comparison, the control group ranked last for all three knowledge measures on both posttests.

Figure 6 compares the adjusted mean aggregate scores on both immediate and delayed posttests of average ability students in the four treatment and control groups.

Figure 6

Bar Graphs of Adjusted Mean Aggregate Scores for Average Ability Students



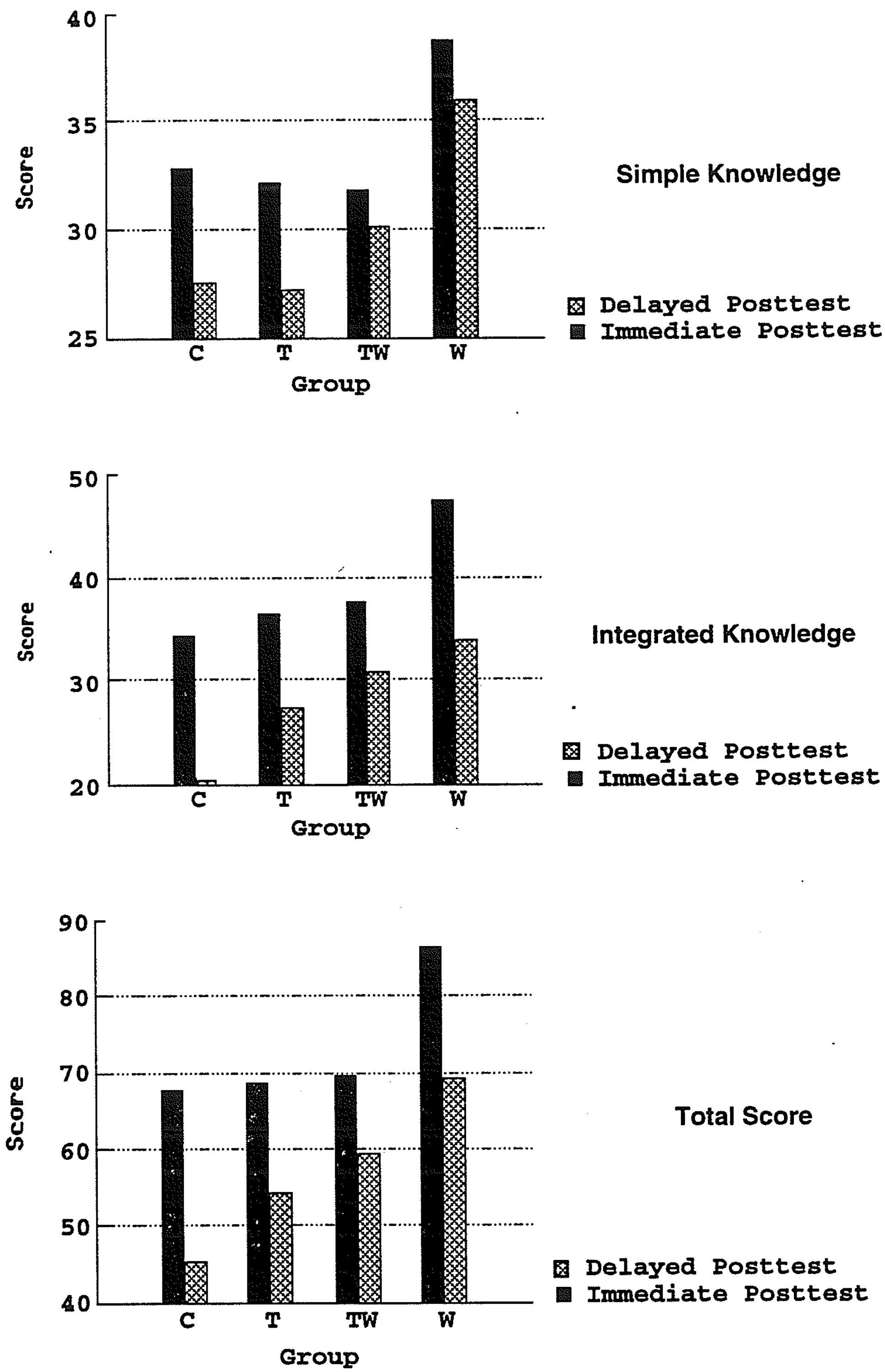
Note . Number of students in each group: C=6; T=7; TW=6; and W=5.

Students of average ability in the two treatment groups which involved talk or peer discussion ranked either first or second for all but the delayed posttest measure of integrated knowledge. For this measure, the control group ranked second. A comparison of adjusted mean aggregate scores among the four treatment groups for both the low and the average ability students suggests that talking with peers is important for learning science. By sharing ideas with others in the group, students appear to redistribute knowledge among them. In contrast, students in the writing-only and control groups, who worked individually on the same or related problem tasks, were not able to benefit from interactions with peers. Knowledge about ecology had to be constructed alone by these students limited by their own experiences and prior knowledge. For students of low or average ability, interacting with others enhances the knowledge pool shared by the group. This pool of ideas might be used by individuals in constructing personal knowledge about ecology.

Figure 7 compares the adjusted mean aggregate scores on both immediate and delayed posttests for high ability students in the four treatment and control groups.

Figure 7

Bar Graphs of Adjusted Mean Aggregate Scores for High Ability Students



Note . Number of students in each group: C=2; T=3; TW=3; and W=2.

In contrast to the trend which was observed with low- and average-ability students that suggested an important role for talk and peer discussion in science learning, high ability students appear to learn more when they use writing. High ability students in the writing-only group ranked first for all three aggregate knowledge measures (simple, integrated, and total score) at both the immediate and delayed posttest. For two of the measures, integrated and total knowledge scores, the students in the talk-and-writing group ranked second for both posttests. Although the talk-and-writing group ranked fourth on the simple knowledge measure at the immediate posttest, they later ranked second behind the writing-only group on the delayed posttest. In comparison, the control group ranked last for both integrated and total knowledge scores on both the immediate and delayed posttests. Writing appears to be a tool that high ability students can effectively use for learning science. Since high ability students in the talk-and-writing group did reasonably well when compared to both the control and talk-only groups, one might argue that peer discussion still plays an important role for these students by allowing them to exchange, clarify, elaborate and consolide their ideas about ecology. However, individually writing their responses to the ecology problems, either with or without peer interaction, appears to be a powerful tool for this particular group of students. More so than students of low or average ability, these students likely already possess adequate knowledge for responding to the different tasks.

In summary, the data suggest that low and average ability students learn more science when they talk with other students. The talk-only and talk-and-writing groups were generally ranked first or second for almost all measures at the immediate and delayed posttest. In contrast, the high ability students generally learned more when writing was involved, either with or without prior discussions with peers. However, since the factorial cell size ranged from a low of one student to a high of seven students, any interpretation must be considered highly tentative. Nonetheless, the trends are particularly interesting and merit further study.

4.3 Limitations and Assumptions

The analysis of variance (ANOVA) makes three assumptions about the statistical treatment of data (Glass & Hopkins, 1984). First, ANOVA assumes that the distribution of scores is normal (the assumption of normality). Second, ANOVA assumes that the scores show the same degree of variability from one treatment group to the next (the assumption of homogeneity of variance). The third assumption requires that the scores be independent from one another both within each treatment group and across groups (the assumption of independence or absence of systematic bias).

The analysis of covariance (ANCOVA) demands that the three assumptions underlying the use of ANOVA also be observed, but imposes three additional assumptions. The first assumption requires that the regression lines for each group in the statistical model be parallel. For instance, this parallelism assumption would be violated if there was an interaction between treatment and covariate. The regression equation in the covariance procedure also assumes that a linear relationship exists between the covariate and the dependent variable. This assumption would be violated if differences existed between groups on the covariate. These two requirements are generally combined and are referred to as the homogeneity of slopes assumption. Finally, a third assumption posits that the covariate contains no measurement error.

4.3.0 *Normality*

For each of the measures at both immediate and delayed posttests, normality was checked using stem-and-leaf plots and normal probability plots which compared residuals with expected values. Although the small sample size of this study ($N = 43$) made it difficult to ensure a normal distribution of scores, a visual inspection of these plots suggested that the assumption of normality did not appear to be violated by most of the measures.

An analysis of kurtosis and skewness suggested that the distributions of multiple choice scores for both integrated and total knowledge measures on the delayed posttest were slightly platykurtic (values of -1.166 and -1.028, respectively). The distribution of the integrated knowledge scores for concept maps on the delayed posttest was somewhat positively skewed (1.088) and leptokurtic (1.340). In comparison, the values for kurtosis and skewness of all other measures were within plus or minus 1.0 for these curve caracteristics. The distributions for all of the measures was thus considered to be within acceptable limits for normality.

4.3.1 *Independence*

The assumption of independence was verified by visually checking studentized residuals plots (S-R plots) and autocorrelation function plots (ACF plots). The S-R plot shows whether or not the residuals are randomly scattered above and below the zero horizontal, whereas the ACF plot measures whether the residuals are serially correlated. Analysis of the studentized residuals plots suggested that the errors were independent. Analysis of the ACF plots indicated that four values in different multiple choice immediate posttest measures might not be independent. However, since the ACF plots for the delayed multiple choice tests were not serially correlated, the data were considered to be satisfactory in terms of independence.

4.3.2 *Homogeneity of Variance*

The assumption of homogeneous variance is critical in a small - *n* study, particularly when the treatment and control groups have unequal *n* 's. This assumption was checked using box plots, studentized residuals plots, as well as the Levene test. Box plots allows a visual inspection of the variance for each of the treatment and control groups, whereas the studentized residuals plots show whether or not the errors have constant variance across groups. Visual inspection of the box plots and the studentized residuals plots suggested that the variance was different across groups on many measures. Moreover, outliers were observed in about half of the box plots for the twenty-four measures. In comparison, output accompanying the S-R plots identified outliers in 19 out of the 24 measures.

Separate Variance Error Term Estimates

Various tests, such as Hartley's *F* - test, Bartlett's test, and Levene's test, have been developed to statistically check the homogeneity of variance assumption (Glass & Hopkins, 1984). The manual accompanying the computer program SYSTAT describes a test for unequal variances based on work by Levene (Levene, 1960; Wilkinson, Hill, & Vang, 1992). The results of the Levene test suggested that three measures which were significant violated this assumption and should be statistically analyzed using an unequal variance *F* - test. These measures were all from the delayed posttest and included the analyses involving both integrated and total knowledge scores based on concept maps, and the integrated knowledge aggregate scores. The contrasts which violated the homogeneous variance assumption were recalculated using separate variance error term estimates.

The analysis suggested that one of the TW+T≠W+C contrasts, the aggregate integrated knowledge measure, was no longer significant. However, the TW≠T+W+C contrast based on the same measure still approached significance (p = .053). Moreover, two other similar contrasts using both the integrated and total knowledge concept mapping scores were still significant when separate variance error term estimates were employed (p = .011 and p = .014, respectively) These analyses using separate variance error term estimates thus confirm the role of writing as a tool for enhancing the retention of science learning over time. Since the immediate posttest measures did not violate the homogeneity of variance assumption, the role of talk as a tool for constructing science knowledge is supported by the data.

4.3.3 *Homogeneity of Slopes*

The homogeneity of slopes assumption was checked by estimating the probability value of a treatment by covariate interaction (Wilkinson, Hill & Vang, 1992). The probability of a treatment-covariate interaction was plausible for six of the twenty-four measures: essay question total knowledge scores at posttest 2; all three posttest one knowledge measures using concept maps; and the simple and total knowledge aggregate measures on the immediate posttest. Five of these measures were not significant in any of the contrasts. However, the TW+T≠W+C contrast based on total knowledge aggregate scores at the immediate posttest had been significantly different (p = .037). The results of this contrast must therefore be rejected.

4.3.4 *Equality Across Groups*

Factorial ANOVA was used to check whether or not there had been differences across the treatment and control groups on the various pretests or covariates. The results suggested that the differences across groups on all twelve covariates were not significant.

4.3.5 *Measurement Error*

Finally, the last assumption underlying ANCOVA states that the covariate contains no measurement error. This was not considered to be a problem because the reliability of all instruments, as reported in the instrumentation section in chapter three, was satisfactory. Although all instruments and measures contain some degree of measurement error, the issue is rather how much error is acceptable in a given study. As such, the covariates likely did not violate this assumption of ANCOVA in the study.

4.4 Peer Discussions

Five problems sessions were included during the experimental unit on ecology. The topics that were explored in these sessions included: (a) biomes; (b) plant and animal adaptations; (c) ecosystems, populations, and communities; (d) habitats and niches; and (e) food webs. One peer discussion group from the talk-and-writing treatment group in each class was videotaped during each of these sessions. Unfortunately, one of the videotapes had to be rejected because of technical difficulties. The remaining nine transcripts of the problem sessions were analysed to illuminate the statistical analyses and to recommend changes in a follow-up study. Most of the peer discussions confirmed the view that this strategy is effective for extending and distributing knowledge among student participants. The role of peer discussion in enhancing concept learning will also be shown using excerpts from selected transcripts. In addition, the roles of hypothesising, of asking questions, of formulating ideas together, and of explaining as engines for driving these peer discussions will be demonstrated. Occasionally, the written responses which students in the talk-and-writing group completed after these peer discussions will be examined to establish the link between talking and writing. This analytical section is organized in three parts. First, the role of talk for learning science will be reviewed. Second, the role of writing in consolidating this learning will be considered. Finally, the limits of talk as a learning strategy will be discussed.

Since the study was conducted in French, all questions, transcripts of discussions, and written responses which are reported have been translated by the researcher. A description of the codes that were used for transcribing the discussions has been included in Appendix H. Pseudonyms have been used to conceal the identity of students in the transcriptions and written responses.

4.4.0 *The Role of Talk*

Glaser (1991) has 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. Since one of the students in the group was absent from school when this problem session was held, only three students participated in the peer discussion.

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

Derek: And, like, there are plants that live on top of the moss.

Kelly: Like fungi.

The transcript demonstrates that peer interaction 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 re-examine their basic knowledge about these concepts and to determine whether or not the definitions which 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 many of the group discussions. 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 which involved examining various plant specimens. The question stated:

> Deciduous trees are well adapted for life in temperate regions where the climate is mild with moderate temperatures and sufficient precipitation. Conifers are well adapted for life in boreal regions where there is less precipitation and colder

temperatures. Cacti are well adapted for life in desert regions where the temperature is warmer with very little precipitation. Study the "leaves" of the following plants and explain how each group is well adapted for life in its particular environment:

a) deciduous trees, such as aspen poplar, willow, elm and maple;

b) conifers, such as spruce, pine, cedar and juniper (the leaves are actually scales or needles);

c) and three different kinds of cacti (the leaves are needles).

In this first excerpt, the students had been examining the coniferous specimens.

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

David: O.K. 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: O.K.

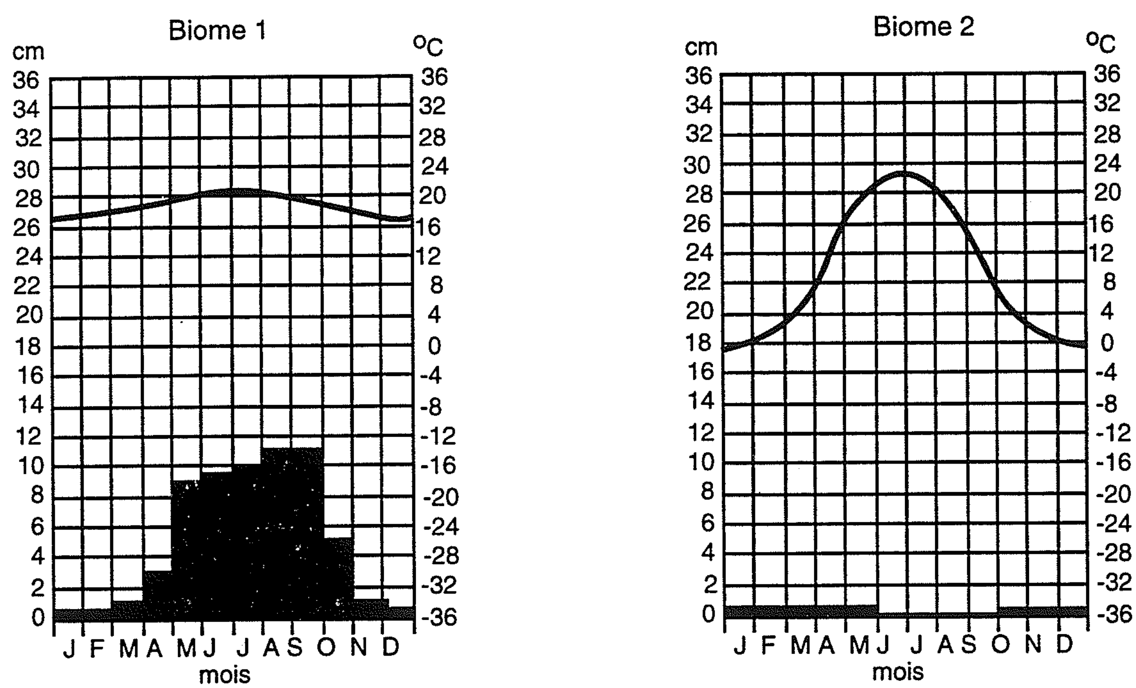
Lisa's questions and her requests for clarification forced the group to elaborate their explanations until they made sense to her and to the others.

Although all of the students contributed ideas during these discussions, high-ability students served a pivotal function in these groups as primary explainers or providers of information. One of the problem tasks asked students to compare and contrast two biomes, a tropical deciduous forest in Oaxaca, Mexico with a desert in Fallon, Nevada. The question stated:

> The following climatograms shows the mean monthly temperature and precipitation for two different biomes. Explain how these biomes are similar and how they are different.

Figure 8

Two Climatograms which Students were Asked to Interpret



In the following exchange, David, a high-ability student, clarifies a concept which is essential to interpreting these climatograms.

David:	Biome 1 // on the average, I think that biome is warmer than biome 2.
Melissa:	*Yeah.*
Elliot:	Me too.
David:	Hmm // hmm / there's a lot less precipitation in biome 2 than in biome 1. / Hmm // So it's probable that biome 2 is in a region /
Elliot:	that is warmer than the other one.
David:	No, it's probably in a desert. // It's probably a desert.
Melissa:	Why is the temperature? /
David:	Because it's night.
Elliot:	It doesn't go to zero?
David:	Well, that's the mean [temperature]. You remember? Yes, say you have minus four [° Celsius] at night / you take the average of the two to calculate the mean.
Melissa:	During the day you have 28 [° Celsius] and at night you have four [° Celsius].
David:	Minus four.
Melissa:	*Yeah, I guess.*
David:	Plus minus four minus 24 gives you 20 [+ (-4) - 24 = 20].

Students had difficulty interpreting the measures which were represented by these climatograms. David's explanations allowed his peers to correctly interpret these climatograms.

Hypothesising also served as an engine for discussions, seemingly keeping students very involved during these talk sessions. In the excerpt just discussed, students had hypothesized that biome 2 likely represented a desert. The following excerpt continues where the previous excerpt left off.

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 / 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.

Melissa: That could be the desert because it's quite warm.

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 millimetres 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. In their written responses after the discussion, all four students correctly indicated the types of biomes represented by these two climatograms.

Shoenfeld has 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 a slender muzzle and long pointed ears, 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 a reasonable 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 foxes' small ears were an adaptation for cold climates because they minimized heat loss and the possibility of freezing these extremities.

4.4.1 *The Role of Writing*

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 interpreted the ideas discussed with peers, and how these were translated into written text using the discursive tools available to each student. Analysis of these written responses framed by the initial peer discussion suggests that talk was used for interpreting the problem task, and for generating, clarifying, and evaluating ideas. Writing, on the other hand, was used for organizing these ideas into a coherent response which would respect grammatical and syntactic conventions. Although the discussions appeared to be rambling and disorganized at times, the written responses which resulted generally appeared to be coherent and well organized. 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.

The discussion appears to be quite jumbled with many incomplete ideas, interruptions, and vague and uncertain references. However, the written responses of these four students suggest that concept learning was occuring below these surface language fragments.

- 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 high-ability students were as productive as other students during the talk sessions, the role of writing appeared to be particularly important to them 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 that might be possible solutions for the problem. However, the written response which 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 and 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:	OK. //
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 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 predated upon, all of the other ideas can be considered satisfactory responses to the question. Interestingly, all of the ideas which were suggested during the preceding discussion have been incorporated into his response, but with more coherence and structure. The responses of his peers were not nearly so coherent or well-organized.

In summary, talk or discussion appears to be important for extending and distributing knowledge among peers. Asking questions, hypothesising, explaining, and formulating ideas together all appear to be important mechanisms during these discussions. Writing, on the other hand, appears to be an important tool for organizing these rudimentary ideas and for constructing knowledge which is coherent and well-structured. However, as Thaiss has observed, writing appears to work only if talking works with it (1988).

4.4.2 *The Limits of Talk*

The use of talk or discussion may be a useful strategy for extending and distributing knowledge among students, but there are limits and pitfalls to its use in the science classroom. First, peer discussion will not be effective if students have no knowledge about the problem task. Second, erroneous ideas may surface during these discussions and be distributed to all of the students in the group. Third, a student may share good ideas with the group, but these may be acquired only partially by some or all of the other students. Evidence of incomplete understandings in spite of good discussions were found in some of the transcripts and written responses. Fourth, some students appeared to benefit from these peer discussions, as witnessed by their written responses, but they were not able to apply these ideas correctly to the posttest question based on similar concepts. Fifth, the role of student status in distributing knowledge was also apparent. Occasionally, good ideas from some students, generally low-achievers, were rejected while erroneous ideas which were proposed by other students were uncritically accepted.

There appear to be limits to the extent that students can explain ecological principles using their own observations and constructed explanations. Although students were generally able to construct satisfactory explanations during most of the problem sessions, occasionally they were not able to invent even an adequate explanation despite working with peers. This appeared to be particularly true in the session on plant and animal adaptations. An excerpt from the transcript of this session was based on the following question:

> Many animals species found in the taïga and the tundra have adaptations to snow. Explain how the snowshoe hare, the lemming, the caribou and the weasel are all adapted for life in the snow (diagrams which highlighted these adaptations accompanied the text).

Lisa: What's a lemming?

Elliot: It's that (pointing to the diagram).

. . .

David: It's a rodent? It's in the rodent family?

. . .

Lisa: Well, they have big thick claws for / like / digging / digging in the ground and for defence.

David: It's probably an animal that hibernates and digs a hole in the ground.

. . .

David: Since it has claws, it can probably dig into the ground to go /

Lisa: Hibernate.

David: *Yeah.*

Melissa: And also for protecting itself.

David: Yes.

The lemming's claws are an adaptation for digging in the hard-packed tundra snow cover and are probably not used for defence against its predators. Furthermore, they are not hibernators. However, lacking basic information about this species and its environmental characteristics, the students attempted to construct a reasonable explanation given what they collectively knew about the problem. If the problem statement had included some basic facts about the lemming and about tundra conditions, the students likely would have been able to develop a satisfactory explanation.

While discussing the case of the snowshoe hare, students grappled with how the feet, which were shown in a diagram, might be considered an adaptation for life in the snow.

Melissa: The feet are / like snowshoes. You know when you want to walk on snow and you do not want to sink.

. . .

David: Why does it have just four toes? // Are the feet covered in fur or is that just a shadow (examining the diagram)?

. . .

Spongy, perhaps.

Elliot: Furry.

Lisa: To hop.

. . .

Its like their / its like cushions under their feet. Its like soft / so <u>they can hop better</u>.

Elliot: <u>It's *soft* when they walk</u>. It's *soft* when they walk.

David: Yes / they can't make a lot of noise because they will be caught real quick.

Melissa: *Yeah* / That's why they don't have nails / so that you don't hear the clicking sound.

Although all four students explained that the size of the hare's feet allowed them to float over instead of sink in deep snow, two of them also added the erroneous idea that the soles of the feet were spongy so that the hare could hop better. Clearly, at some point input from the teacher, or some other knowledgeable person, is necessary to dispel these erroneous ideas. Alternatively, students might be expected to consult other sources, including print-based resources, to confirm or to reject some of these tentative ideas.

Occasionally, good ideas surfaced in the peer groups that were later integrated into the written responses of most of students, but only as incomplete explanations. 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 four or five months. Summer lasts about two or three [months] / *so* they don't have time to grow and to regrow again.

The written responses which were completed by the four students in this group varied in appropriateness.

- These leaves [needles] are long and thick [walled] and they are solidly anchored onto the branch, so that the tree does not lose its leaves like deciduous trees. Pines do not have the time to lose their leaves, and to regrow them. There isn't enough sunlight for this process (David).
- The needles are long, and I predict that nearly all of the trees were dry. The conifers have a process of slow growth, and that is why they don't fall in winter because they can't reproduce them (Lisa).
- They are needles and not leaves. They do not have the time to reproduce so they never fall (Elliot).

• The soil is acid [acidic]. The needles are hard and do not have the time to fall because the deciduous trees are soft and do have the time to fall (Melissa).

David's response is an adequate explanation for the adaptation of conifers. However, the other three students have either misinterpreted the ideas which he presented in the discussion or given only incomplete explanations using kernels of these ideas.

Occasionally, good ideas which were presented by students, generally low-achievers or girls, were ignored or rejected in favour of erroneous ideas which were presented by others. For instance, in the following excerpt, the students were asked to explain whether or not chihuahuas, bulldogs and german shepherds all belonged to the same species.

Bill: No, I think that it's false / because there are different / there are different dogs. / It's similar to / the larvae / there are many kinds of larvae [the student is refering to the larvae of aquatic insects that students had observed in an earlier lab investigation].

Derek: But they [the dogs] are all one species / and that's three different / three, even four / different / no, three / different dogs / not species / that live together.

Kelly: I think that it's true because species is like / a dog is a species, a cat is a species, and all that / and there is like / it would be a family or something like that (...)

Bill: Yes, but it's still not a single species.

Edmond: A species is all of the same kind.

Kelly: OK. And family would be all that (pointing to the dogs).

Edmond: OK. *So,* false.

Although two of the three students present (one student was absent during this session) initially argued that the three dogs all belonged to the same species, they later relinguished and accepted Bill's idea of a family of dogs consisting of three separate species. The written responses of the students also confirmed their acceptance of this erroneous explanation.

In most cases, good ideas were shared during the peer discussion and these were later incorporated into the written responses which were completed immediately following the discussion. In some cases, however, students were not able to apply these ideas later to a novel yet similar problem on the posttest. One of the problems in the session on niche and habitat asked students to explain how a hunter might use his knowledge of bear ecology. An excerpt of the discussion of this question by one of the talk-and-writing groups was discussed earlier. Interestingly, the low-achieving student in the two groups which were videotaped were not able to apply this knowledge to a similar problem on the posttest. The question asked students to explain how a fisherman might be able to use his knowledge of the ecology of the jackfish. Neither student attempted this question on either the immediate and delayed posttests.

In summary, talk or discussion may be an effective strategy for enhancing learning in the science classroom. However, there are limitations and pitfalls which must be considered when implementing this strategy. Students, collectively as a group, must have some information or knowledge before they can benefit from sharing with peers. The task must also be appropriate for students and cast in such a manner that discussion is encouraged. Redundancy might be used to ensure that all group members have a complete understanding of the problem being discussed. For instance, group members might be asked to individually explain to the others who would ensure that the responses demonstrated a complete understanding. Webb (1982, 1985) had found that the act of explaining enhanced learning. The propagation of erroneous ideas might be controlled by requiring students to further research the problems using a variety of resources, or by conducting whole-class discussions to scrutinize and evaluate ideas.

A number of studies have reported that low ability students have considerable difficulty applying science concepts (Department of Education and Science, 1985;

Lapointe, Mead, & Phillips, 1989). This was also observed in the present study. One possible solution would be to use a series of related problems, all applying the same fundamental concepts: McGinley and Tierney's notion of "traversing a topical landscape" (1989, p. 249). Hopefully, these students would learn to generalize the concepts to novel but similar instances.

The role of student status in biaising group deliberations had been noted by Linn and Burbules (1993) and was also observed in the present study. Possibly by combining these small groups into increasingly larger deliberative bodies might resolve this problem by forcing students to justify their thinking before more of their peers. However, the success of this strategy is not assured. Another approach would involve requiring students to consult authoritative references to support their explanations. A final approach would involve a whole-class discussion with the teacher assuming an authoritative stance with regards to the content.

4.5 Summary

The hypotheses which framed this study addressed the effects of treatment, gender, and the interaction of treatment and gender. Although treatment effects were not observed at the immediate posttest when aggregate measures were employed, exploratory analyses using various multiple choice measures did show that these effects may have been washed out because of the small sample size. Nonetheless, treatment effects were present at the delayed posttest for simple, integrated and total knowledge. *Post hoc* pairwise comparisons suggest that discussing problem tasks with peers before writing explanations for them enhanced the retention of science learning over time.

Gender effects were observed for all three aggregate knowledge measures at both times-of-test. However, a gender-treatment interaction was observed for just the simple knowledge delayed posttest aggregate measure. *Post hoc* pairwise comparisons suggest that girls showed better retention of simple facts and concepts over time when they had talked with peers. Boys, however, showed better retention when they, not only talked with peers, but also wrote down explanations to the various problem tasks.

Contrasts involving various combinations of treatment and control groups suggest that talking with peers, either with or without writing afterwards, is important for constructing knowledge, whereas writing is important for the long-term retention of this knowledge over time.

5. SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Following a brief overview of the study and of the findings, this chapter will discuss implications of the current research and provide suggestions for future research.

5.0 Summary and Discussion

The ultimate purpose of this study was to serve as a template for new research. To this end, trends in the data were investigated in order to formulate hypotheses and questions and to recommend procedures that might be used in a follow-up study. The more immediate or proximate purpose was to investigate the role of talk and writing, alone and combined, on learning in science.

5.0.0 *Overview of Study*

A mixed factorial design, with a pretest and two posttests (immediate and delayed), was employed to test the effects of independent variables (gender and treatment) on dependent measures which included multiple choice tests, essay questions, and concept maps. A *post hoc* analysis of data was also conducted to investigate the effect of ability on treatment. Simple, integrated and total knowledge scores were established for all of the measures. Aggregate scores comprised of totals for each of these kinds of knowledge were also analysed. All of the measures were found to be valid and reliable for the study. In addition, some peer discussions were videotaped and later analysed to determine how students acquire knowledge while interacting with other students.

Forty-three eighth-grade students in two intact classrooms in a Franco-Manitoban school were randomly assigned to four treatment and control groups, all stratified for gender and ability. The treatment groups included a talk-only group, a writing-only group, a talk-and-writing group and a control group. During a six week unit on ecology, students were separated into these four treatment and control groups for five problem-solving sessions, which focused on different key concepts in ecology each time. With the exception of the control group, the tasks that were assigned in the sessions involved explaining applications of these ecological concepts. The talk-only group discussed these problems with peers. The students in the writing-only group individually responded in writing to these same problems. Students in the talk-and-writing group discussed the problems with peers before individually responding in writing. In comparison, the control group received descriptive tasks based on the same ecological concepts.

Statistical treatment of the data involved a series of analyses of covariance using pretest scores as covariate each time. Hypothesis testing involved analyses based on the aggregate scores only for each kind of knowledge (simple, integrated, and total). Other exploratory analyses were based on the three separate measures, as well as the aggregate scores, for both the immediate and delayed posttest results. *Post hoc* analyses using Least-Significant-Difference pairwise comparisons were also examined for trends in the data. In addition, contrasts comparing either the two groups which 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 possibility of interactions between student ability and experimental treatment was assessed using bar graphs of adjusted mean aggregate scores at both immediate and delayed posttests.

5.0.1 *Overview of the Findings*

Three null hypotheses guided this study. Although the study was exploratory in nature, these still served a useful role in focusing the analysis.

Hypothesis Testing

Hypothesis One. The first hypothesis stated that there will be no significant difference ($p < .05$) 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 found for the main effects of treatment or for the interaction of treatment and gender when the aggregate simple knowledge scores were analyzed. As was the case for every test of hypothesis, significant differences were found between male and female students for this measure. At the delayed posttest, significant differences were apparent for treatment, for gender, and for gender-treatment interaction. However, excluding outliers in this analysis suggested that the main effect for treatment might actually be a gender-treatment interaction. Moreover, pairwise comparisons, both with and without outliers, suggested that this appeared to favour the use of talk, alone or combined with writing, for the retention of science facts over time. Nonetheless, since some of these comparisons favoured females whereas others favoured males, the findings may simply be an artifact of a small - *n* study. The issue of whether talk, alone or combined with writing, enhances the learning of facts for all students, or just for male or female students was therefore left unresolved.

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 males again outperforming females. At the delayed posttest, differences were observed for the effects of both gender and treatment. In addition, when outliers were excluded from the analysis, significant differences were found for the treatment-gender interaction. *Post hoc* analysis of pairwise contrasts, both with and without outliers, suggests that students who discussed problems with peers before individually writing explanations for them (TW group) showed better retention of integrated knowledge over time than students who had not interacted with peers. The *post hoc* analysis without the outliers for the gender-treatment interaction also suggests that female students who used the talk-and-writing strategy outperformed those female students who did not benefit from peer discussions. However, differences between females in the TW group and those in the talk-only group were not significant.

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 (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 males again outperforming females. At the delayed posttest, both analyses, with and without outliers, indicated that differences exist for treatment, for gender,and for the gender-treatment interaction. Students using the talk-and-writing strategy again outperformed those who did not benefit from peer discussion (W and C). Moreover, female students who initially discussed the problems with peers (TW and T) learned more than those who worked individually at their desk.

Discussion. The results suggest that peer discussion combined with analytical writing may enhance the retention of science knowledge over time for most students, but appears to have little effect on immediate learning. Contrasts with those students who worked individually on similar (W group) or related tasks (C group) were particularly striking. However, the differences between those students who simply discussed the problems in peer groups with those who also used analytical writing were not as apparent. 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. Since differences were measured after a delay of six weeks, then differences should have been observed immediately following the unit. However, differences did favour the talk-and-writing group, then the talk-only group, the writing-only group and the control group in that order. However, they 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 and total knowledge). In fact, when the outliers were excluded from this series of analyses, all three multiple choice posttest-one measures were significant for the main effects of treatment.

The issue of a small - *n* study is confirmed by some of the contrasts in which one or more treatment groups were compared with other groups, for instance the TW+T≠W+C contrast. When the two groups which used talk were contrasted with the other two groups which did not benefit from peer interaction, significant differences were now observed for four measures, one of them being the total aggregate score. Moreover, the two other aggregate measures were marginally significant: the simple knowledge aggregate score ($p = .09$) and the integrated knowledge aggregate score ($p = .06$). More will be said about these contrasts later.

Answering the Research Questions

The literature suggests that talk or peer discussion is important for mediating the initial construction of knowledge, whereas writing primarily exerts its effect on the retention of this constructed knowledge over time. The immediate posttest was thus designed to measure the initial learning or constructed knowledge, whereas the delayed posttest assessed the retention of this learning. In addition, gender and ability may be important intervening variables, enhancing the use of talk or writing by one more subgroups. The six research questions specifically addressed these issues.

Writing or Talking. The analyses of covariance and *post hoc* contrasts unanimously rejected the idea that writing-alone enhances learning more than peer discussion. This finding suggests that talk is important for knowledge construction in science learning. Working individually, many students probably lacked the knowledge base for constructing adequate explanations for the problem tasks, and consequently did not learn as much alone as those students who discussed these problems with peers.

Talk-and-Writing or Just Writing. 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 have previously constructed this new knowledge, and interacting with peers appeared to enhance the construction process. Only then did writing appear to exert a positive effect on learning, enhancing retention over time.

Talk-and-Writing or Just Talking. The results are far from conclusive on the issue of whether talk combined with writing is superior to talking-only. Four *post hoc* contrasts using intact data favoured the TW group over the T group, compared with only one contrast in the other direction. However, after the outliers were excluded from the analysis, the comparisons were even weaker. This issue will have to be resolved in a future study.

Comparisons with the Control Group. Although the control group ranked fourth for half of all measures and third for the remaining measures when adjusted mean scores were compared between the four conditions, *post hoc* contrasts were less definitive. The contrasts suggested that diferences were significant between both groups using peer discussion (TW and T) and the control group for about a quarter of all measures. In addition, despite the writing-only group ranking ahead of the control group for three-quarters of the measures, the differences between these two groups were non-significant. One might argue that the writing-only group essentially served as another control in the experiment.

Contrasts between groups confirmed that peer discussion was important for knowledge-construction. When the TW and T groups were contrasted with the W and C groups at the immediate posttest, four of twelve measures were significantly different ($p < .05$) and another two were marginally significant ($p < .10$). Excluding outliers from the analysis or using separate error term estimates essentially did not affect this finding.

Contrasts between groups also confirmed that writing after peer discussions promoted the retention of learned material over time. When the TW group was contrasted with the other three groups (TW, T, and C) on delayed posttest measures, half of the contrasts for the effect of treatment favoured the use of talk combined with writing. When outliers were excluded from the analysis, the number of significant contrasts increased to seven ($p < .05$) and one other contrast was marginally significant ($p < .10$). These findings suggest that writing is important for the retention of science knowledge over time, but that talk or peer discussion is essential for the initial construction of this knowledge. Writing only seems to work if talk works with it.

Interactions with Gender. Although only one measure was significant for gender-treatment interactions when the intact data were analysed, after excluding outliers five measures were significant and one other approached significance. All six of these measures were from the delayed posttest. For the five measures which were significant, *post hoc* pairwise comparisons suggested that female students demonstrate better retention of science learning when instruction has included peer discussion, alone or combined with writing. Contrasts involving various combinations of treatment groups (TW+T≠W+C and TW≠T+W+C) confirmed the role of peer discussion for enhancing retention of learning over time. This suggests that a large -*N* study might show significance for many of the measures.

Interactions with Ability. Visual inspection of bar graphs representing the adjusted mean scores for all conditions suggested that peer discussion is particularly important for students of low or average ability. These students seemed to benefit the most from sharing knowledge. In comparison, high-ability students appeared to do better when they individually responded to the problem tasks in writing. However, since the talk-and-writing group always ranked right behind the writing-only group on the delayed posttests, one could argue that talk may enhance the retention of science knowledge over time even for high-ability students. In fact, for both integrated and total knowledge aggregate scores, the talk-only group ranked in between the talk-and-writing group and the control group. Since the high-ability students tended to act as peer tutors by often explaining during discussion sessions, talking may have promoted the retention of integrated, complex knowledge over time for this group of students.

Peer Discussion

The peer discussions which were videotaped showed that talking allowed students to extend the knowledge available to them, collectively, and to redistribute it. Asking questions, explaining, hypothesizing, and formulating ideas together were all important mechanisms for constructing knowledge. Talk was used for interpreting the problem tasks and for generating, clarifying, and evaluating ideas. Writing, on the other hand, was used for organizing these idea fragments into a coherent and structured response.

However, there are limits to peer discussion. If the group, collectively, does not possess even basic ideas about the problem task, then peer discussion will be generally ineffective as a learning strategy. Incomplete or poor understanding, or an inability to apply the acquired knowledge later on can also result in spite of discussions with peers. An uncritical acceptance of ideas, particularly those suggested by high-status students, is another problem that can hinder the effectiveness of group discussion.

5.1 Recommendations

In the present section, the trends which were apparent in the data will be briefly summarized, then specific suggestions regarding hypotheses, questions, methods and instruments which might be implemented in a future study will be described.

5.1.0 *Implications of the Current Research*

Lacking confidence in the data because of the small size of the sample, firm conclusions were considered to be inappropriate on the basis of this preliminary study. Nonetheless, trends were evident in the data and tentative conclusions have been proposed.

Talk for Constructing Knowledge

Peer discussion appeared to be an important mechanism for knowledge construction. When the two groups using talk were contrasted with the writing-only and control groups at the immediate posttest, all three knowledge scores based on the multiple choice measures were significantly different at $p < .01$. In addition, the aggregate total knowledge scores were also significantly different at $p < .05$. Moreover, the two other aggregate measures were marginally significant at $p = .06$ and $p = .09$ for integrated and simple knowledge, respectively. Although differences between these groups on the essay questions and the concept maps still favoured those students using talk, the contrasts were nonsignificant.

Writing for Retention

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, all six concept mapping and aggregate measures were significant at $p < .01$. Moreover, excluding the outliers from the analysis resulted in one other measure, the integrated knowledge measure using the essay questions, being significant at $p < .05$, and another for total knowledge with the essay questions marginally significant at $p = .07$. Although the multiple choice measures did not favour the talk-and-writing group, they still confirmed the role of peer discussion in enhancing learning, as the TW+T≠W+C contrast using total knowledge scores was significantly different at $p = .028$, whereas the other two total scores for simple and integrated knowledge were marginally significant at $p = .065$ and $p = .097$, respectively.

5.1.1 *Suggestions for Future Research*

Size of Sample

Since interactions with gender and with ability appear to be important mediating variables in the use of talk and writing for learning science, a future study should include a larger sample of students. To achieve an n of 15 students in each factorial cell, a sample of 120 or 180 students would be required for investigating interactions between treatment and gender or treatment and ability, respectively.

Questions and Hypotheses

A future study might address the following questions:

1. Do middle school students learn more science when peer discussion is employed as a classroom strategy?
2. Is the retention of this knowledge better over time if students have used a talk-and-writing strategy?
3. Do female students learn more science when peer discussion is employed as a classroom strategy?
4. Is the retention of this knowledge better over time if female students have used a talk-and-writing strategy?
5. Do low- and average-ability students learn more science when peer discussion is employed as a classroom strategy?
6. Is the retention of this knowledge better over time if low- and average-ability students have used a talk-and-writing strategy?
7. Do high-ability students learn more science when individual analytic writing tasks are employed as a classroom strategy?
8. Is the retention of this knowledge better over time if high-ability students have used a talk-and-writing strategy?

The hypotheses which could guide a future study might include:

1. There will be no significant difference ($p < .05$) in mean multiple choice test scores, a measure of simple knowledge, at the immediate posttest between the two groups using talk or peer discussion and the writing-only and control groups.
2. There will be no significant difference ($p < .05$) in mean multiple choice test scores, a measure of simple knowledge, at the delayed posttest between the talk-and-writing group and the other three conditions combined.

3. There will be no significant difference ($p < .05$) in mean essay question scores, a measure of integrated knowledge, at the immediate posttest between the two groups using talk or peer discussion and the writing-only and control groups.
4. There will be no significant difference ($p < .05$) in mean essay question scores, a measure of integrated knowledge, at the delayed posttest between the talk-and-writing group and the other three conditions combined.
5. There will be no significant difference ($p < .05$) in mean concept mapping scores, a measure of total knowledge, at the immediate posttest between the two groups using talk or peer discussion and the writing-only and control groups.
6. There will be no significant difference ($p < .05$) in mean concept mapping scores, a measure of total knowledge, at the delayed posttest between the talk-and-writing group and the other three conditions combined.

Instrumentation

One of the weaknesses of the present study was the inability to distinguish between simple and integrated knowledge. Although most of the multiple choice questions appeared to be valid for measuring either simple knowledge, such as facts and concepts, or integrated knowledge, such as applications, the expert panel thought that some of them could be interpreted either way. In a future study, the multiple choice questions could be restricted to factual and comprehension questions that are specifically based on the unit of study. Instead of using questions which have been used in previous national or international assessments that may or may not be perfectly matched to the study unit, questions would be developed for actually assessing the facts and concept in the unit of study. This means that the multiple choice test could be used as a measure of simple knowledge in all analyses.

The essay questions would also be tied more closely to the unit of study, but they would still be evaluated using a holistic approach. However, the questions should not be evaluated for simple knowledge, as this appeared to be an artificial construct in this experiment. These scores could thus be considered a measure of integrated knowledge in the analyses.

For the concept mapping, most of the key concepts in the ecology unit would be given to students in the test protocol. An expert panel would determine which propositions are valid and how they would be scored. The concept maps, which would be scored for the presence of both concepts and propositions, could thus be considered a measure of total knowledge. However, their real importance would be as records of the knowledge structure of students. These maps could thus be analysed qualitatively to assess how students' understanding of the content domain changed with the instructional unit.

Methods

In the present study, the students had been instructed about the different types of explanations based on Kourany's typology. Although the students had reported that they were already familiar with these types of explanations, many of the verbal and written explanations produced during the problem sessions were still unsatisfactory or incomplete. King (1994) used a direct teaching approach for showing elementary students how to generate explanations. This approach could be blended with the current approach in any future study to ensure that student explanations were satisfactory.

Summary

Firm conclusions were considered to be inappropriate on the basis of this preliminary study. Although not all of the statistical tests were significant, trends were evident in the data. Peer discussion appeared to be an important mechanism for knowledge construction, whereas peer discussion combined with writing appeared to enhance the retention of science knowledge over time. Minor changes in instrumentation and method should ensure that any future study will be stronger than the present study. However, the most important change would be enhancing the size of the sample. With a larger *N* , many of the statistical analyses and *post hoc* pairwise comparisons and multi-group contrasts might be significant.

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), New Directions for Teaching and Learning Series, N° 49. 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* (3): 315-343.

Alvermann, D. E. (1991). The discussion web: A graphic aid for learning across the curriculum. *The Reading Teacher, 45,* 92-99.

Alvermann, D. E., Dillon, D. R., & O'Brien, D. G. (1987). *Using discussion to promote reading comprehension* . Newark, DE: International Reading Association.

Alvermann, D. E., O'Brien, D. G., & Dillon, D. R. (1990). What teachers do when they say they're having discussions of content area reading assignments: A qualitative analysis. *Reading Research Quarterly, 25*, 296-322.

Ammon, P. & Ammon, M.S. (1990). *Using student writing to assess and promote understandings in science.* Center for the Study of Writing: Occasional Paper No. 16. (ERIC Document Production Service No. ED 316 864)

Anderson, C. W. (1987). Strategic teaching in science. In B. F. Jones, A. S. Palincsar, D. S. Ogle, & E. G. Carr (Eds.), *Strategic teaching and learning: Cognitive instruction in the content areas* (pp. 73-91). Alexandria, VA: Association for Supervision and Curriculum Development.

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, CO: JAI Press.

Anson, C. M. (1988). Toward a multidimensional model of writing in the academic disciplines. In D. A. Jolliffe, (Ed.). *Advances in writing research: Writing in academic disciplines, Vol. 2* (pp.1-33). Norwood, NJ: Ablex.

Antaki, C. (Ed.). (1988). *Analysing everyday explanations: A casebook of methods.* London: Sage Publications.

Applebee, A. N. (1981). *Writing in the secondary school: English and the content areas* (NCTE Research report No. 21). Urbana, IL: National Council of Teachers of English.

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

Applebee, A. N. (Ed.). (1984b). *Contexts for learning to write: Studies of secondary school instruction.* Norwood, NJ: Ablex.

Applebee, A. N., Langer, J. A., & Mullis, I. V. S. (1986). *Writing: Trends across the decade.* National Assessment of Educational Progress, Princeton, NJ: Educational Testing Service.

Ausubel, D. P., Novak, J. D., & Hanesean, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York: Holt, Rinehart & Winston.

Baier, K. (1980). The Meaning of life. In E. D. Klemke, R. Hollinger, & A. D. Kline (Eds.), *Introductory readings in the philosophy of science* (pp. 338-369). Buffalo, NY: Prometheus Books.

Baker, D., & Leary, R. (1995). Letting girls speak out about science. *Journal of Research in Science Teaching, 32* (1): 3-27.

Barenholz, H., & Tamir, P. (1992). A comprehensive use of concept mapping in design instruction and assessment. *Research in Science and Technological Education*, 10, 37-52.

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

Barnes, D. (1990). Oral language and learning. In S. Hynds & D. L. Rubin (Eds.), *Perspectives on talk and learning* (pp. 41-54). Urbana, IL: National Council of Teachers of English.

Bernard, R. M., & Naidu, S. (1992). Post-questioning, concept mapping and feedback: A distance education field experiment. *British Journal of Educational Technology, 23* (1): 48-60.

Beyer, B. K. (1982). Using writing to learn social studies. *The social studies, 73* (33): 100-105.

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

Brey, C. (1984). Les travaux pratiques de reformulation. *Langue française, 64* : 68-80.

Bridges, D. (1990). The character of discussion: A focus on students. In W. W. Wilen (Ed.), *Teaching and learning through discussion: The theory, research and practice of the discussion method* (pp. 97-112). Springfield, IL: Charles C. Thomas.

Britton, J. (1970). *Language and learning.* Coral Gables, FL: University of Miami 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.

Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). *The development of writing abilities (11-18)* (Schools Councils Research Studies). London: MacMillan Educational.

Bromberger, S. (1966). Why-questions. In R. G. Coldny (Ed.), *Mind and cosmos: Essays in contemporary science and philosophy.* Pittsburgh: University of

Pittsburgh Press.

Brophy, J., & Alleman, J. (1991). Activities as instructional tools: A framework for analysis and evaluation. *Educational Researcher, 20* (4): 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.* Hillsdale, NJ: Lawrence Erlbaum.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18* (1): 32-42.

Bruffee, K. A. (1984). Collaborative learning and the 'conversation of mankind'. *College English, 46* (7): 635-652.

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

Bruner, J., & Haste, H. (Eds.). (1987). *Making sense: The child's construction of the world.* New York, NY: Methuen.

Cazden, C. B. (1986). Classroom discourse. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 432-463). New York, NY: Macmillan.

Champagne, A. B., Gunstone, R. F., & Klopfer, L. E. (1985). Instructional consequences of students' knowledge about physical phenomena. In L. H. T. West & A. L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 61-90). Orlando, FL: Academic Press.

Cliburn, J. W., Jr. (1987). Helping students understand physiological interactions: A concept mapping activity. *The American Biology Teacher, 49* (7): 426-427.

Cohen, E. G. (1986). *Designing groupwork: Strategies for the heterogeneous classroom.* New York, NY: Teachers College Press.

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

Cole, N. S. (1990). Conceptions of educational achievement. *Educational Researcher, 19* (3): 2-7.

Coley, R. J. (1989). The gender gap in education: How early and how large? *ETS Policy Notes, 2*, 1-8.

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

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 453-494).

Hillsdale, NJ: Lawrence Erlbaum.

Comley, N. R., Hamilton, D., Klaus, C. H., Scholes, R., & Sommers, N. (1984). *Fields of writing: Readings across the disciplines.* New York: St. Martin's Press.

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science, and programming. In C. B. Cazden (Ed.), *Review of research in education.* Washington, DC: American Educational Research Association.

Connolly, P. & Vilardi, T. (Eds.). (1989). *Writing to learn mathematics and science.* New York, NY: Teachers College Press.

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.

Costa, A. L. (1990). Teacher behaviors that promote discussion. In W. W. Wilen (Ed.), *Teaching and learning through discussion: The theory, research and practice of the discussion method* (pp. 45-77). Springfield, IL: Charles C. Thomas.

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

Cuban, L. (1989). The "at-risk" label and the problem of urban school reform. *Phi Delta Kappan,70* (10), 780-801.

Cushman, K. (1992). Conversation in classrooms: Who are seminars for? *The Harvard Education Letter*, *8* (2): 1-4.

Davis, B. G., Scriven, M., & Thomas, S. (1987). *The evaluation of composition instruction, 2nd ed.* New York, NY: 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: Literacy Publications.

Department of Education and Science. (1975). *Language for life* (Bullock Committee Report). London: Her Majesty's Stationary Office.

Department of Education and Science. (1985). *Science at age 15, Science report for teachers: 5.* London: Her Majesty's Stationary Office.

Dillon, J. T. (1988). *Questioning and discussion: A multidisciplinary study.* Norwood, NJ: Ablex.

Donlan, D. (1974). Teaching writing in the content areas. *Research in the Teaching of English* , *8*, 250-262.

Driver, R. (1981). Pupils' alternative frameworks in science. *European Journal of Science Education*, *3*, 93-101.

Driver, R. (1989). Changing conceptions. In P. Adey, J. Bliss, J. Head, & M. Shayer (Eds.), *Adolescent development and school science* (pp. 79-103). New York, NY: The Falmer Press.

Durst, R. K. (1987). Cognitive and linguistic demands of analytic writing. *Research in the Teaching of English, 21*, 347-376.

Durst, R. K. (1989). Monitoring processes in analytic and summary writing. *Writing Communication, 6* (3): 340-363.

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* (4): 375-394.

Duschl, R. A. (1994). Editorial policy statement and introduction. *Science Education, 78* (3): 203-208.

Dyson, A. H., & Freedman, S. W. (1991). *Critical challenges for research on writing and literacy: 1990-1995* (Technical Report No. 1-B). Center for the Study of Writing, Berkeley, CA; Center for the Study of Writing, Pittsburgh, PA. (ERIC Document Reproduction Service No. ED 335 676)

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

Edwards, D., & Mercer, N. (1987). *Common knowledge: The development of understanding in the classroom.* London, UK: Methuen.

Elbow, P. (1983, September). Teaching thinking by teaching writing. *Change* , pp. 37-40.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28*, 122-128.

Erickson, G., & Farkas, S. (1991). Prior experience and gender differences in science achievement. *The Alberta Journal of Educational Research, 37*, 225-239.

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, NY: Teachers College, Columbia University.

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* (9): 985-1001.

Ferguson, J. (1984). *Who turns the wheel?* Proceedings of a workshop in the science education of women in Canada. Ottawa, ON: Science Council of Canada.

Fillion, B. (1979, Winter). Language across the curriculum. *McGill Journal of Education,* pp. 47-60.

Fillion, B. (1985). Language across the curriculum. In T. Husen & T. N. Postlethwaite (Ed.), *The international encyclopedia of education: Research and studies, Vol. 5* . Oxford, UK: Pergamon Press.

Finley, F., & Stewart, J. (1982). Representing substantive structures. *Science*

Education, 66 (4): 593-611.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59* (2): 117-142.

Fowler, H. L. (1990). Writing across the curriculum: A call for pedagogical change in the secondary school (Doctoral dissertation, University of North Carolina at Greensboro, 1989). *Dissertation Abstracts International, 50*, 2767A.

Freedman, R. L. H. (1994). *Open-ended questioning: A handbook for educators.* Menlo Park, CA: Addison-Wesley.

Fulwiler, T. (1987). *Teaching with writing.* Montclair, NJ: Boynton/Cook.

Gall, J. P., & Gall, M. D. (1990). Outcomes of the discussion method. In W. W. Wilen (Ed.), *Teaching and learning through discussion: Theory, research and practice of the discussion method* (pp. 25-44). Springfield, IL: Charles C. Thomas.

Gall, M. D., & Gall, J. P. (1993, April). *Teacher and student roles in different types of classroom discussions.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. ED 359 256)

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*(9): 1039-1056.

Gere, A. R. (Ed.). (1985). *Roots in the sawdust: Writing to learn across the disciplines.* Urbana, IL: National Council of Teachers of English.

Gere, A. R. (1990). Talking in writing groups. In S. Hynds & D. L. Rubin (Eds.), *Perspectives on talk and learning* (pp. 115-128). Urbana, IL: National Council of Teachers of English.

Gilbert, S., & Orok, B. (1993). L'abandon scolaire. *Tendances sociales canadiennes, 30* :2-7.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. *Learning and instruction, 1*: 129-144.

Glass, G. V., & Hopkins, K. D. (1984). *Statistical methods in education and psychology, 2nd ed.* Englewood Cliffs, NJ: Prentice-Hall.

Glasser, W. (1990). The quality school. *Phi Delta Kappan, 71* (6): 425-435.

Glatthorn, A. A. (1991). Restructuring schools: Curriculum reform. In K. M. Kershner & J. A. Connolly (Eds.), *At-risk students and school restructuring* (pp.77-84). Philadelphia, PA: Research for Better Schools.

Goodlad, J. I. (1984). *A place called school: Prospects for the future.* New York, NY: McGraw-Hill.

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.

Gunstone, R. F., Gray, C. M. R., & Searle, P. (1992). Some long-term effects of uninformed conceptual change. *Science Education, 76*, 175-197.

Haggerty, S. M. (1991). Gender and school science: Achievement and participation in Canada. *The Alberta Journal of Educational Research, 37*, 195-208.

Halliday, M. A. K. (1975). *Learning how to mean: Exploration in the development of language*. London, UK: Edward Arnold.

Hart, D. (1994). *Authentic assessment: A handbook for educators.* Menlo Park, CA: Addison-Wesley.

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.

Hegarty-Hazel, E., & Prosser, M. (1991a). Relationship between students' conceptual knowledge and study strategies—part 1: Student learning in physics. *International Journal of Science Education, 13*, 303-312.

Hegarty-Hazel, E., & Prosser, M. (1991b). Relationship between students' conceptual knowledge and study strategies—part 2: Student learning in biology. *International Journal of Science Education, 13* (4): 421-429.

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, NY: The Free Press.

Hempel, C. G. (1966). *Philosophy of natural science.* Englewood Cliffs, NJ: Prentice-Hall.

Hempel, C. G. (1987). Statistical explanation. In J. A. Kourany (Ed.), *Scientific knowledge: Basic issues in the philosophy of science*, (pp. 44-50). Belmont, CA: Wadsworth.

Hempel, C. G., & Oppenheim, P. (1987). Studies in the logic of explanation. In J. A. Kourany (Ed.), *Scientific knowledge: Basic issues in the philosophy of science*, (pp. 30-43). Belmont, CA: Wadsworth.

Hendry, G. D., & King, R. C. (1994). On theory of learning and knowledge: Education implications of advances in neuroscience. *Science Education, 78* (3): 223-253.

Hewson, P. W., & Stewart, J. (1994). Learning section—Editorial policy statement. *Science Education, 78* (3): 213-215.

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D.

(1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77* (1): 95-111.

Howard, J. (1983). *Writing to learn.* Washington, DC: Council for Basic Education.

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, NY: Routledge.

Hurd, P. D. (1990). Historical and philosophical insights on scientific literacy. *Bulletin of Science, Technology and Society, 10* (3): 133-136.

Hurd, P. D. (1991). Why we must transform science education. *Educational Leadership, 49*, 33-35.

Hynd, C. R., McWhorter, J. Y., Phares, V. L., & Suttles, C. W. (1994). The role of instructional variables in conceptual change in high school physics topics. *Journal of Research in Science Teaching, 31* (9): 933-946.

Hynd, C. R., Qian, G., Ridgeway, V. G., & Pickle, M. (1991). Promoting conceptual change with science texts and discussion. *Journal of Reading, 34* , 596-601.

International Association for the Evaluation of Educational Achievement (IEA). (1988). *Science achievement in seventeen countries: A preliminary study.* Oxford, UK: Pergamon.

Jegede, O. J., Alaiyemola, F. F., & Okebukola, P. A. O. (1990). The effect of concept mapping on students' anxiety and achievement in biology. *Journal of Research in Science Teaching, 27* (10): 951-960.

Jensen, V. (1987). Writing in college physics. In T. Fulwiler (Ed.), *The journal book* (pp. 330-336). Portsmouth, NH: Boynton/Cook-Heinemann.

Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 23-37). NY: Praeger.

Johnson, D. W., & Johnson, R. T. (1991). *Learning together and alone: Cooperative, competitive, and individualistic learning, 3rd ed.* Englewood Cliffs, NJ: Prentice-Hall.

Johnson, D. W., Johnson, R. T., Holubec, E. J., & Roy, P. (1988). *Circles of learning: Cooperation in the classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.

Kagan, S. (1992). *Cooperative Learning.* San Juan Capistrano, CA: Kagan Cooperative Learning.

Kamman, C. F. (1990). Writing across the curriculum: Implications for preservice teacher education (Doctoral dissertation, Florida Atlantic University, 1990). *Dissertation Abstracts International, 51* , 1989A.

Kempa, R. (1986). *Assessment in science.* Cambridge: Cambridge University Press.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31* (2): 338-368.

King, L. H. (1993). High and low achievers' perceptions and cooperative learning in two small groups. *The Elementary School Journal, 93* (4): 399-416.

Kirkpatrick, L. D. & Pittendrigh, A. S. (1984). A writing teacher in the physics classroom. *The Physics Teacher, 22* (3): 159-164.

Klemke, E. D., Hollinger, R., & Kline, A. D. (Eds.). (1980). *Introductory readings in the philosophy of science.* Buffalo, NY: Prometheus Books.

Knoblauch, C. H., & Brannon, L. (1983). Writing as learning through the curriculum. *College English, 45* (5): 465-474.

Knoblauch, C. H., & Brannon, L. (1984). *Rhetorical traditions and the teaching of writing.* Upper Montclair, NJ: Boynton/Cook.

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

Lamb, H., & Purves, A. C. (1988). The reflective composition. In T. P. Gorman, A. C. Purves, & R. E. Degenhart (Eds.), *The IEA study of written composition I: The international writing tasks and scoring scales* (pp. 172-192). Oxford: Pergamon.

Langer, J. A. (1984). The effects of available information on responses to school writing tasks. *Research in the Teaching of English, 18*, 27-44.

Langer, J. A. (1986a). Learning through writing: Study skills in the content areas. *Journal of Reading, 29*,400-406.

Langer, J. A. (1986b). *Children reading and writing: Structures and strategies.* Norwood, NJ: Ablex.

Langer, J. A., & Applebee, A. N. (1986). Reading and writing instruction: Toward a theory of teaching and learning. In E. Z. Rothkopf (Ed.), *Review of Research in Education, Vol. 13* (pp. 171-194). Washington, D.C.: American Educational Research Association.

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

Langone, J. (1990). Emerging Viruses. *Discover, 11* (12): 63-68.

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.

Lehman, J. D., Carter, C., & Kahle, J. B. (1985). Concept mapping, vee mapping and achievement: Results of a field study with black high school students. *Journal of Research in Science Teaching, 27* (7): 663-674.

Lehr, J. B., & Harris, H. W. (1988). At-risk, low achieving students in the classroom.

Washington, D. C.: National Education Association.

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

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics* (pp.278-292). Palo Alto, CA: Stanford University Press.

Linn, M. C. (1990). Summary: Establishing a science and engineering of science education. In M. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, A. diSessa, & E. Stage (Eds.), *Toward a scientific practice of science education* (pp. 323-341). Hillsdale, NJ: Lawrence Erlbaum.

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

Linn, M. C., & Hyde, J. S. (1989). Gender, mathematics, and science. *Educational Researcher, 18*, 17-19, 22-27.

Linn, M. C., & Songer, N. B. (1993). How do students make sense of science? *Merrill-Palmer Quarterly, 39* (1): 47-73.

Lloyd, C. V. (1990, November-December). *The enactment of literacy in high school Biology classrooms: Two case studies.* Paper presented at the annual meeting of the National Reading Conference, Miami, FL. (ERIC Document Reproduction Service No. ED 337 747)

Lock, R. (1992). Gender and practical skill performance in science. *Journal of Research in Science Teaching, 29* , 227-241.

Lomask, M., Baron, J. B., Greig, J., & Harrison, C. (1992, March). *ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge of science.* A symposium presented at the annual meeting of the National Association for Research in Science Teaching. Cambridge, MA.

Lonning, R. A. (1993). Effect of cooperative learning strategies on student verbal interactions and achievement during conceptual change instruction in 10th grade general science. *Journal of Research in Science Teaching, 30* (9): 1987-1101.

Maheady, L., Harper, G. F., & Mallette, B. (1991). Peer-mediated instruction: A review of potential applications for special education. *Reading, Writing, and Learning Disabilities, 7* (2): 75-103.

Malone, J., & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. *School Science and Mathematics, 84* (3): 220-231.

Manitoba Education. (1986). *Manitoba assessment: Final report.* Winnipeg, MB: Queen's printers.

Manitoba Education and Training. (1993a). *Physics 300: Final report.* Winnipeg, MB: Queen's printers.

Manitoba Education and Training. (1993b). *Manitoba science assessment 1990: Summary report.* Winnipeg, MB: Queen's printers.

Marshall, J. D. (1984). Process and product: Case studies of writing in two content areas. In A. N. Applebee (Ed.), *Contexts for learning to write* (pp. 149-168). Norwood, NJ: Ablex.

Marshall, J. D. (1987). The effects of writing on students' understanding of literary texts. *Research in the Teaching of English, 21*(1): 30-63.

Martin, J. R. (1970). *Explaining, understanding, and teaching.* NY: McGraw-Hill.

Martin, N., D'Arcy, P., Newton, B., & Parker, R. (1976). *Writing and learning across the curriculum 11-16.* London: Ward Lock Educational.

Mason, C. L. (1992). Concept mapping: A tool to develop reflective science instruction. *Journal of Research in Science Teaching, 76* (1): 51-63.

Mason, C. L., & Kahle, J. B. (1989). Student attitudes toward science and science-related careers: A program designed to promote a stimulated gender-free learning environment. *Journal of Research in Science Teaching, 26*, 25-39.

Mastropieri, M. A., & Scruggs, T. E. (1992). Science for students with disabilities. *Review of Educational Research, 62* (4): 377-411.

McCulley, G. A. (1986). Research in writing across the curriculum: Beginnings. In A. Young & T. Fulwiler (Eds.), *Writing across the disciplines: Research into practice* (pp. 42-48). Upper Montclair, NJ: Boynton/Cook.

McGinley, W., & Tierney, R. J. (1989). Traversing the topical landscape: Reading and writing as ways of knowing. *Written Communication, 6* (3): 243-269.

McLaughlin, M. W., & Talbert, J. E. (1993). Introduction: New visions of teaching. In D. K. Cohen, M. W. McLaughlin, & J. E. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 1-10). San Francisco: Jossey-Bass.

McMillan, J. H., & Schumacher, S. (1989). *Research in education: A conceptual introduction, 2nd ed.* New York, NY: Harper & Row.

McTeague, F., Payne, R., Graham, A., & Murray, J. (1981). An investigation of secondary student writing across the curriculum and some suggestions for school language policies. In I. Pringle & A. Freedman (Eds.), *Teaching, writing, learning.* Ottawa: Canadian Council of Teachers of English.

Meloth, M. S., & Deering, P. D. (1994). Task talk and task awareness under different cooperative learning conditions. *American Educational Research Journal, 31* (1): 138-165.

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.

Murray, D. M. (1982). *Learning by teaching: Selected articles on writing and teaching.*

Montclair, NJ: Boynton/Cook.

National Center for Improving Science Education. (1989). *Getting started in science: A blueprint for elementary school science education.* Washington, DC: The Network, Inc.

Newell, G. E. (1983). Learning from writing in two content areas: A case study/protocol analysis (Doctoral dissertation, Stanford University, 1983). *Dissertation Abstracts International, 44*, 412A.

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* (4): 291-302.

Newman, D., Griffin, P., & Cole, M. (1984). Social constraints in laboratory and classroom tasks. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 172-193). Cambridge, MA: Harvard University Press.

Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school.* Cambridge: Cambridge University Press.

Newmann, F. M. (1988). Can depth replace coverage in the high school curriculum. *Phi Delta Kappan, 69* (5): 345-348.

Novak, J. D. (1980). Learning theory applied to the Biology classroom. *American Biology Teacher, 42* (5): 280-285.

Novak, J. D. (1984). Application of advances in learning theory and philosophy of science to the improvement of chemistry teaching. *Journal of Chemical Education, 61* (7): 607-611.

Novak, J. D. (1990a). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27* (10): 937-949.

Novak, J. D., (1990b). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science, 19*, 29-52.

Novak, J. (1991). Clarify with concept maps. *The Science Teacher, 58* (7): 45-49.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn.* Cambridge: Cambridge University Press.

Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education, 67* (5): 625-645.

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal, 28* (1): 117-153.

O'Flahavan, J. F., Hartman, D. K., & Pearson, P. D. (1988). Teacher questioning and feedback practices: A twenty year retrospective. In J. E. Readence & R. S. Baldwin (Eds.), *Dialogues in literacy research,* Thirty-seventh yearbook of the National

Reading Conference (pp. 183-208). Chicago, IL: National Reading Conference.

O'Loughlin, M. (1991, April). *Beyond constructivism: Toward a dialectical model of the problematics of teacher socialization.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Odell, L. (1980). The process of writing and the process of learning. *College Composition and Communication, 31*, 42-50.

Okebukola, P. A. (1992). Concept mapping with a cooperative learning flavor. *The American Biology Teacher, 54* (4): 218-221.

Okebukola, P. A., & Jegede, O. J. (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education, 72* (4): 489-500.

Oliver, J. S., & Simpson, R. D. (1988). Influences of attitude toward science, achievement motivation, and science self concept on achievement in science: A longitudinal study. *Science Education, 72*, 143-155.

Orpwood, G. W. F., & Souque, J.-P. (1984). *Science education in Canadian schools, summary of background study 52.* Ottawa, ON: Ministry of Supply and Services Canada.

Palincsar, A. S. (1986). The role of dialogue in scaffolded instruction. *Educational Psychologist, 21*, 71-98.

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* (5): 643-658.

Pankratius, W. J. (1990). Building an organized knowledge base: Concept mapping and achievement in secondary school physics. *Journal of Research in Science Teaching, 27* (4): 315-333.

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* (3), 265-277.

Pearce, D. L. (1984). Writing in content area classrooms. *Reading World, 23* (3): 234-241.

Pechmann, J.H.K., Scott, D. E., Semlitsch, R. D., Caldwell, J. P., Vitt, L. J., & Gibbons, J. W. (1991). Declining Amphibian Populations: The Problem of Separating Human Impacts from Natural Fluctuations. *Science, 253* : 892-895.

Perkins, D., & Blythe, T. (1994). Putting understanding up front. *Educational Leadership, 51* (5): 4-7.

Pines, A. L. (1985). Toward a taxonomy of conceptual relations and the implications for the evaluation of cognitive structures. In L. H. T. West & A. L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 101-116). Orlando: Academic Press.

Powers, D. T. (1990). The effects of hands-on science instruction on students' cognition structures as measured by concept maps. (Doctoral dissertation, Kansas State University, 1990). *Dissertation Abstracts International, 51* (6): 1895-A.

Prawat, R. S. (1989a). Promoting access to knowledge, strategy, and disposition in students: A research synthesis. *Review of Educational Research, 59* (1):1-41.

Prawat, R. S. (1989b). Teaching for understanding: Three key attributes. *Teaching and Teacher Education, 5* (4): 315-328.

Prawat, R. S. (1993). The value of ideas: Problems versus possibilities in learning. *Educational Researcher, 22* (6): 5-16.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elementary School Journal, 89* (3): 301-342.

Purves, A. C. (1984). In search of an internationally-valid scheme for scoring compositions. *College Composition and Communication, 35* (4): 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, NY: Longman.

Reif, F., & Larkin, J. H. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching, 28* (9): 733-760.

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

Resnick, L. B. (1991). Shared cognition: Thinking as social practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 1-20). Washington, DC: American Psychological Association.

Resnick, L. B., & Klopfer, L. E. (Eds.). (1989). *Toward the thinking curriculum: Current cognitive research* (1989 ASCD Yearbook). Alexandria, VA: Association for Supervision and Curriculum Development. (ERIC Document Reproduction Service No. ED 328 871)

Rice, D., Ryan, J., Samson, S. (1993, April). *A year-long middle school study of concept maps for instruction and assessment correlated with traditional achievement measures.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Rivard, L. P. (1994). A Review of writing to learn in science: Implications for practice and research. *Journal of Research in Science Teaching, 31* (9):969-983.

Rivard, L. P., & Yore, L. D. (1992, March). *Review of reading comprehension instruction: 1985-1991.* Paper presented at the annual meeting of the National

Association for Research in Science Teaching, Boston, MA. (ERIC Document Reproduction Service No. ED 354 144)

Rogan, J. M. (1988). Conceptual mapping as a diagnostic aid. *School Science and Mathematics, 88* (1): 50-59.

Rogoff, B. (1984). Introduction: Thinking and learning in social context. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 1-8). Cambridge, MA: Harvard University Press.

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

Ross, B., & Munby, H. (1991). Concept mapping and misconceptions: A study of high-school students' understandings of acids and bases. *International Journal of Science Education, 13* (1): 11-23.

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

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

Roth, K. J., Anderson, C. W., & Smith, E. L. (1987). Curriculum materials, teacher talk and student learning: Case studies in fifth grade science teaching. *Journal of Curriculum Studies, 19* :527-548.

Rubin, D. L. (1990). Introduction: Ways of talking about talking and learning. In S. Hynds & D. L. Rubin (Eds.), *Perspectives on talk and learning* (pp. 1-17). Urbana, IL: National Council of Teachers of English.

Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans.* New York, NY: Oxford University Press.

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 students 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). NY: Macmillan.

Scardamalia, M., & Bereiter, C. (1989). Conceptions of teaching and approaches to core problems. In M. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 37-46). New York, NY: Pergamon Press.

Schmid, R. F., & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *The Journal of Educational Research, 84* (2): 78-85.

Schoenfeld, A. H. (1989). Ideas in the air: Speculations on small group learning, environmental and cultural influences on cognition, and epistemology. *International Journal of Education Research, 13* (1): 71-88.

Schumacher, G. M. & Nash, J. G. (1991). Conceptualizing and measuring knowledge change due to writing. *Research in the Teaching of English, 25* (1)67-96.

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

Science Council of Canada. (1984b). *Science education in Canadian schools, volume III: Case studies of science teaching.* Ottawa, ON: Ministry of Supply and Services, Canada.

Sensenbaugh, R. (1989). Writing across the curriculum: Evolving reform (ERIC/RCS). *Journal of Reading, 32*, 462-465.

Shanahan, T. (1988). The reading-writing relationship: Seven instructional principles. *The Reading Teacher, 41*(7): 636-647.

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

Shore, L. S., Hakerem, G., & Hickman, P. (1993, April). *Using concept maps to compare expert and novice understanding of the scientific applications of fractal geometry.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Skolnick, J., Langbort, C., & Day, L. (1982). *How to encourage girls in math and science.* Englewood Cliffs, NJ: Prentice-Hall.

Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin, 94* (3): 429-445.

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

Slavin, R. E. (1993). Ability grouping in the middle grades: Achievement effects and alternatives. *Elementary School Journal, 93* (5): 535-552.

Slesnick, I. L. (1985). *Biology.* Glenview, IL: Scott, Foresman and Company.

Smail, B. (1987). Organizing the curriculum to fit girls' interests. In A. Kelly (Ed.), *Science for girls?* (pp. 80-88). Milton Keynes, PA: Open University Press.

Solomon, J. (1987). Social influences on the construction of pupils' understanding of science. *Studies in Science Education, 14*, 63-82.

Stanley, L. C. (1991). A two-tiered writing program for the technology curriculum.

Teaching English in the two-year college, 18 : 109-113.

Stenhouse, D. (1986). Conceptual change in science education: Paradigms and language-games. *Science Education, 70*, 413-425.

Stewart, J., Van Kirk, J., & Rowell, R. (1979). Concept maps: A tool for use in biology teaching. *The American Biology Teacher, 41* (3): 171-175.

Stinner, A. (1992). Science textbooks and science teaching: From logic to evidence. *Science Education, 76*(1): 1-16.

Strenski, E. (1984). The write stuff. *The Science Teacher*, *51* (5): 58-61.

Stuart, H. A. (1985). Should concept maps be scored numerically? *European Journal of Science Education, 7*(1): 73-81.

Sullenger, K. S. (1991). Science teachers' perspectives about writing in science (Doctoral dissertation, University of Georgia, 1990). *Dissertation Abstracts International, 52,* 126A.

Tamir, P. (1989). Some issues related to the use of justifications to multiple-choice answers. *Journal of Biological Education, 23* (4): 285-292.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Science, 12*, 435-502.

Thaiss, C. (1988). The future of writing across the curriculum. In McLeod, S. H. (Ed.), *Strengthening programs for writing across the curriculum* , Volume 36 in the series New Directions for Teaching and Learning (pp. 91-102). San Francisco, CA: Jossey-Bass.

Thall, E., & Bays, G. (1989). Utilizing ungraded writing in the chemistry classroom. *Journal of Chemical Education, 66* (8): 662-663.

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

Thornburg, D. G. (1991). Strategy instruction for academically at-risk students: An exploratory study of teaching "higher-order" reading and writing in the social studies. *Reading, Writing, and Learning Disabilities, 7* (4): 377-406.

Tighe, M. A. (1991). *Teaching composition across the curriculum in Southeastern Alabama and in Suffold County, England* . Paper presented at the annual meeting of the Conference on College Composition and Communication, Boston, MA. (ERIC Document Reproduction Service No. ED 331 094)

van Fraassen, B. (1987). The pragmatics of explanation. In J. A. Kourany (Ed.), *Scientific knowledge: Basic issues in the philosphy of science* (pp. 65-76). Belmont, CA: Wadsworth.

Van Nostrand, A. D. (1979). Writing and the generation of knowledge. *Social Education, 43*, 178-180.

VanOrden, N. (1987). Critical-thinking writing assignments in general chemistry.

Journal of Chemical Education, 64 (6): 506-507.

von Wright, G. H. (1971). *Explanation and understanding.* Ithaca, NY: Cornell University Press.

Vitale, M. R., Romance, N. R., Parke, H., & Widergren, P. (1994, March). *Teaching for student conceptual understanding in science: Research implications from an interdisciplinary perspective.* Paper presented to the annual meeting of National Association for Research in Science Teaching. Anaheim, CA.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT 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* (10): 1033-1052.

Wandersee, J. H. (1990). Concept mapping and the cartography of cognition. *Journal of Research in Science Teaching, 27* (10): 923-936.

Weaver, W. (1964). Scientific explanation. *Science, 143* : 1297-1300.

Webb, N. M. (1982). Student interaction and learning in small groups. *Review of Educational Research, 52* (3): 421-445.

Webb, N. M. (1985). Student interaction and learning in small groups: A research summary. In R. Slavin, S. Sharon, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmusck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 147-172). NY: Plenum Press.

Webb, N. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21-40.

Wehlage, G., Smith, G., & Lipman, P. (1992). Restructuring urban schools: The new futures experience. *American Educational Research Journal, 29* (1): 51-93.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research in teaching, 3rd ed.* (pp. 315-327). New York, NY: Macmillan.

Weiss, R. H., & Walters, S. (1980, April). *Writing to learn.* Paper presented at annual meeting of the American Educational Research Association, Boston, MA.

Wells, G. (1986). *The meaning makers: Children learning language and using language to learn.* London, UK: Hodder & Stoughton.

Wertsch, J. V. (1991). A sociocultural approach to socially shared cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 85-100). Washington, DC: American Psychological Association.

Whitehead, A. N. (1929). *The aims of education and other essays.* New York, NY: The Macmillan Company.

Wilen, W. W. (Ed.). (1990). *Teaching and learning through discussion: The theory, research and practice of the discussion method.* Springfield, IL: Charles C. Thomas.

Wilkinson, L., Hill, M., & Vang, E. (1992). *SYSTAT: Statistics, Version 5.2 Edition.* Evanston, IL: SYSTAT.

Willerman, M., & Mac Harg, R. A. (1991). The concept map as an advance organizer. *Journal of Research in Science Teaching, 28* (8): 705-711.

Wood, E., Fler, C., & Willoughby, T. (1992). Elaborative interrogation applied to small and large groups. *Applied Cognitive Psychology, 6,* 361-366.

Yager, R. E. (1989). The power of a current issue for making school programs more relevant. *Social Science Record, 26* (2): 42-43.

Yager, R. E. (1991). The constructivist learning model: Towards real reform in science education. *The Science Teacher, 58* (6): 52-57.

APPENDIX A:

Instructional Material used by the Teacher for the Nature of Scientific Explanation

L'explication scientifique

Expliquer, c'est faire connaître les raisons ou les causes d'un phénomène. Une bonne explication comprend une description des relations entre le phénomène à expliquer (l'effet) et les événements qui l'ont précédé ou les conditions qui existent (les causes).

Nous pouvons différencier quatre types d'explication en sciences:

- LA COMPOSITION — Une explication peut se fonder sur la composition d'un objet. Une explication des propriétés d'un objet pourrait comprendre une description des propriétés des parties qui le composent.

  Par exemple, nous expliquons les propriétés d'un neurone en décrivant ces parties (synapses, dendrites, etc.) et en expliquant comment ces parties accomplissent leur rôle dans le système nerveux.

- LA FONCTION — Une explication de la fonction d'un objet pourrait comprendre une description du fonctionnement des parties qui le composent.

  Par exemple, l'adaptation des poissons à une vie aquatique pourrait s'expliquer par une description du fonctionnement de leurs branchies, de leurs nageoires, etc.

- L'ÉVOLUTION — Une explication peut se fonder sur l'évolution ou le développement d'un objet dans le temps.

  Par exemple, une explication de la vitesse des cyclistes dans le Tour de France actuel pourrait comprendre une description des progrès qu'ont subis les vélos depuis une vingtaine d'années.

- LA TRANSITION — Une explication des changements dans un objet pourrait comprendre une description d'un ou des événements qui ont perturbé l'objet dans le passé.

  Par exemple, une explication de la pollution des Grands Lacs pourrait comprendre une description des matières rejetées dans les égouts depuis une trentaine d'années par les municipalités et les industries qui entourent ces lacs.

APPENDIX A:

Instructional Material used by the Teacher for the Nature of Scientific Explanation

Quel type d'explication?

Dis quel type d'explication est representé par chacun des énoncés suivants:

- *une explication de la STRUCTURE d'un objet;*
- *une explication de la FONCTION d'un objet;*
- *une explication de l'ÉVOLUTION dans un objet; ou*
- *une explication de la TRANSITION dans un objet.*

1. Explique comment le feu peut être bon et mauvais pour une forêt.

   TYPE: TRANSITION

2. Explique comment les conifères (par ex., les pins et les épinettes) sont adaptés pour survivre au nord du Manitoba, où il y a peu de précipitation et une courte saison de croissance pour les plantes.

   TYPE: COMPOSITION

3. Explique comment le corps humain se protège contre des maladies infectieuses.

   TYPE: FONCTION

4. Explique comment la rivière Seine pourrait changer au cours du prochain siècle.

   TYPE: ÉVOLUTION

5. Y a-t-il beaucoup de microbes dans le sable du désert? Explique.

   TYPE: COMPOSITION

6. Comment les animaux de montagne réagissent-ils à l'arrivée de l'hiver?

   TYPE: TRANSITION

7. Pourquoi les reptiles sont-ils rares dans l'Arctique?

   TYPE: FONCTION

8. Comment les heures de lumière du jour (photopériode) varient-elles avec les saisons au nord du cercle arctique?

   TYPE: ÉVOLUTION

APPENDIX B:

Explanatory Tasks for the Problem-Solving Session on Ecosystems, Populations and Communities

QUESTIONS À DÉVELOPPEMENT: Écosystème, population et communauté

1. Un écosystème est formé de l'ensemble des être vivants, du milieu dans lequel ils vivent et des relations qui existent entre ces êtres et leur milieu. Explique comment la ville de Winnipeg peut être considérée comme un écosystème.

____

____

____

____

____

____

____

________________________________ au verso

2. Une bûche morte est-elle un écosystème? Explique.

____

____

____

____

____

________________________________ au verso

3. Quels facteurs abiotiques seraient parmi les plus importants pour la vie d'une grenouille? Explique l'importance de chacun de ces facteurs abiotiques pour la vie de cette espèce.

____

____

____

____

____

________________________________ au verso

APPENDIX B:

Explanatory Tasks for the Problem-Solving Session on Ecosystems, Populations and Communities

4. Dis si chacun des énoncés suivants est vrai ou faux. Explique ta réponse.

a) Toi et tes camarades de classe appartenez à la même population.

____

____

____

____

____________________________________ au verso

b) Tous les insectes dans un jardin forment une seule population.

____

____

____

____

____________________________________ au verso

c) Les caniches, les bouledogues et les bergers allemands appartiennent à la même espèce.

____

____

____

____

____________________________________ au verso

5. Comment ferais-tu pour déterminer le nombre de pissenlits sur une pelouse sans toutefois les compter un à un?

____

____

____

____

____

____________________________________ au verso

APPENDIX C:

Simple Learning Tasks for the Problem-Solving Session on Biomes

EXERCICE EN ÉCOLOGIE: Les biomes

1. Relie chaque terme de la colonne A à l'énoncé de la colonne B qui lui convient le mieux.

Colonne A	*Colonne B*
• organisme ____	a) une science
• écosystème ____	b) fine couche du globe terrestre où se trouvent la vie
• écologie ____	c) un être vivant
• biome ____	d) le raton laveur
• espèce ____	e) un espace vert dans la ville
• biosphère ____	f) la taïga

2. Décris deux caractéristiques qui permettent de différencier un biome d'un autre.

   a) ______________________________

   b) ______________________________

3. Identifie le biome (toundra, taïga, forêt décidue, forêt pluviale tropicale, désert et prairie) dans lequel chaque espèce ci-dessous est généralement retrouvée.

Plante	*Animal*
• un cactus ____________	• le Jaguar ____________
• un érable ____________	• le Lemming ____________
• des herbes ____________	• le Loup ____________
• un sapin ____________	• le Bison ____________
• un lichen ____________	• le Scorpion ____________

4. Identifie le biome (toundra, taïga, forêt décidue, forêt pluviale tropicale, désert et prairie) décrit par chacun des énoncés suivants:

   ________ a) une région où les arbres perdent leurs feuilles chaque automne.

   ________ b) une région aride où les précipitations annuelles ne dépassent pas 20 cm.

   ________ c) une région connue sous divers noms autour du monde: pampa, steppe et savanne.

__________ d) une région où les étés sont frais, les hivers froids et les précipitations annuelles sont de 40 à 100 cm.

__________ e) une région où les températures sont chaudes et les pluies abondantes tout au cours de l'année.

__________ f) une région ayant plus de précipitations que les déserts mais pas assez pour la survie des arbres.

__________ g) une région où les jours sont chauds et les nuits fraîches.

__________ h) une région froide avec peu de lumière pendant une bonne partie de l'année.

__________ i) une région caractérisée par le frêne, le chêne, l'érable et le bouleau.

__________ j) une région caractérisée par des mousses et des lichens.

__________ k) une région caractérisée par des conifères et des marécages.

__________ l) une région caractérisée par une grande diversité végétale.

5. Vrai ou faux: évalue chacun des énoncés suivants.

__________ a) Le climat de la toundra est extrêmement froid, venteux et sec.

__________ b) La toundra reçoit beaucoup de neige pendant l'hiver.

__________ c) Il n'y a pas de fleurs sauvages aux couleurs éclatantes dans la toundra.

__________ d) Les reptiles ne sont pas retrouvés dans la toundra.

__________ e) Les températures moyennes dans la toundra sont plus élevées que celles dans la taïga.

__________ f) La taïga reçoit beaucoup plus de neige que la toundra.

__________ g) Les plantes qui caractérisent la taïga sont les conifères comme l'épinette et le sapin.

__________ h) La forêt décidue se trouve dans des régions moins froides que la toundra et la taïga.

__________ i) Les prairies reçoivent généralement plus de précipitations que les forêts décidues.

__________ j) C'est le manque d'eau, bien plus que la chaleur, qui donne naissance aux déserts.

APPENDIX D:

Instructions for Supervising the Problem-Solving Sessions

DIRECTIVES POUR LES CINQ SESSIONS[1]

GROUPE W (writing-only)

- Distribue le texte "**Questions à développement**" à chaque élève dans le groupe (vérifie la liste et inscris les noms sur les textes).

- Identifie les élèves absents et note leur nom:

______________________________ ______________________________

______________________________ ______________________________

- Chaque élève doit travailler individuellement. Tu peux clarifier les questions et les énoncés sans donner trop d'indices.

- Chaque élève doit remettre le texte au complet à la fin de la période. Encourage les élèves à répondre à autant de questions que possible dans le temps alloué.

- Note comment la session s'est déroulé. As-tu des suggestions? Veuillez les écrire sur la feuille pour commentaires.

GROUPE C (contrôle)

- Distribue le texte "**Exercice en écologie**" à chaque élève dans le groupe (vérifie la liste et inscris les noms sur les textes).

- Identifie les élèves absents et note leur nom:

______________________________ ______________________________

______________________________ ______________________________

- Chaque élève doit travailler individuellement. Tu peux clarifier les questions et les énoncés sans donner des indices.

- Chaque élève doit remettre le texte au complet à la fin de la période. Encourage les élèves à répondre à autant de questions que possible dans le temps alloué.

- Note comment la session s'est déroulé. As-tu des suggestions? Veuillez les écrire sur la feuille pour commentaires.

---

[1] Les cinq sessions sont: (1) les biomes; (2) les adaptations; (3) écosystème, population et communauté; (4) habitat et niche; et (5) reseaux alimentaires.

APPENDIX D:

Instructions for Supervising the Problem-Solving Sessions

GROUPES T (talk-only)

- Distribue le texte "**Questions à développement**" à chaque élève dans le groupe (vérifie la liste et inscris les noms sur les textes).

- Tu peux clarifier les questions et les énoncés sans donner trop d'indices.

- Demande aux élèves de discuter autant de questions que possible dans le temps alloué en respectant la démarche présentée pour la discussion. Les élèves **N'ÉCRIVENT PAS** leurs réponses à ces questions. Tout se passe à l'oral seulement.

- Identifie les élèves absents et note leur nom:

______________________ ______________________

______________________ ______________________

- Chaque élève doit remettre le texte au complet à la fin de la période. Encourage les élèves à répondre à autant de questions que possible dans le temps alloué.

- Note comment la session s'est déroulé. As-tu des suggestions? Veuillez les écrire sur la feuille pour commentaires.

GROUPES TW (talk & writing)

- Distribue le texte "**Questions à développement**" à chaque élève dans le groupe (vérifie la liste et inscris les noms sur les textes).

- Identifie les élèves absents et note leur nom:

______________________ ______________________

______________________ ______________________

- Tu peux clarifier les questions et les énoncés sans donner trop d'indices.

- Demande aux élèves de discuter autant de questions que possible dans le temps alloué en respectant la démarche présentée pour la discussion. Les élèves **DOIVENT INDIVIDUELLEMENT ÉCRIRE** leurs réponses à ces questions.

- Chaque élève doit remettre le texte au complet à la fin de la période. Encourage les élèves à répondre à autant de questions que possible dans le temps alloué.

- Note comment la session s'est déroulé. As-tu des suggestions? Veuillez les écrire sur la feuille pour commentaires.

APPENDIX E:

Three-Step Procedure for Discussing Explanatory Tasks

Discussion en sous-groupes: la démarche

1re étape: faire un remue-méninges — QUE PENSEZ-VOUS?

Tour à tour, chaque élève dans le groupe devrait proposer une explication. Les pairs ne devraient pas poser de jugement sur les idées proposées par chacun et chacune (Linn & Burbules, 1993).

2e étape: clarifier et étoffer les idées — SONT-ELLES CLAIRES ET COMPLÈTES?

Tour à tour, chaque élève dans le groupe est invité à:
- élaborer une des idées proposées;
- clarifier une des idées proposées;
- poser une ou des questions sur une des idées proposées;
- répondre à une des questions posées (Webb, 1989).

3e étape: évaluer les idées proposées — S'AGIT-IL DE BONNES EXPLICATIONS?

Tour à tour, chaque élève dans le groupe est invité à analyser, évaluer et critiquer les idées proposées ou les explications. Une bonne explication devrait accomplir un des rôles suivants:

- Une ou des causes du phénomène sont mises en évidence.
- La relation entre le phénomène et les conditions ou événements qui l'ont précédé sont mis en évidence.
- La description des parties ou des composants explique les propriétés de l'objet.
- La description du développement de l'objet dans le passé explique sa condition actuelle.
- La description du fonctionnement des parties composant l'objet explique son rôle ou sa fonction.
- Le changement dans l'objet est expliqué par une perturbation subie dans le passé.

4e étape: rédiger l'explication — COMMENT EXPLIQUER LE PHÉNOMÈNE?

- Chaque élève dans le groupe devrait rédiger son explication individuellement.

APPENDIX F:

Guidelines for Discussing the Explanatory Tasks

Comportements souhaités pendant les discussions[1]

1re étape: faire un remue-méninges — QUE PENSEZ-VOUS?

1. Encouragez la participation de tous et chacun.
2. Partagez vos idées. Chacun et chacune doivent tenter une explication du phénomène.
3. Écoutez attentivement ce que les autres disent. Notez les points que vous aimeriez faire ressortir plus tard.

2e étape: clarifier et étoffer les idées — SONT-ELLES CLAIRES ET COMPLÈTES?

4. Élaborez les idées déjà proposées en y ajoutant d'autres éléments (Johnson, Johnson, Holubec, & Roy, 1988).
5. Posez des questions quand vous n'avez pas bien compris les idées proposées par vos pairs.
6. Répondez aux questions posées par vos pairs.
7. Donnez votre raisonnement.

3e étape: évaluer les idées proposées — S'AGIT-IL DE BONNES EXPLICATIONS?

8. Ne critiquez pas les personnes mais les idées (Johnson, Johnson, Holubec, & Roy, 1988).
9. Comparez et évaluez les idées proposées.

4e étape: rédiger votre explication — COMMENT EXPLIQUER LE PHÉNOMÈNE?

10. Ne parlez pas à vos pairs pendant la rédaction de l'explication.

---

1 Basé sur Bridges (1990); Cohen (1986); Costa (1990); Johnson, Johnson, Holubec, & Roy (1988); Meloth & Deering (1994); Palincsar, Anderson, & David (1993).

APPENDIX G:

Written Prompt for Scaffolding Peer Discussions

DISCUSSION EN SOUS-GROUPES: LA DÉMARCHE

1. Faire un remue-méninges
   — **QUE PENSEZ-VOUS?**

2. Clarifier et étoffer les idées
   — **SONT-ELLES CLAIRES ET COMPLÈTES?**

3. Évaluer les idées proposées
   — **S'AGIT-IL DE BONNES EXPLICATIONS?**

4. Rédiger l'explication
   — **COMMENT EXPLIQUER LE PHÉNOMÈNE?**

APPENDIX H:

Transcription Codes*

Bold type	To mark emphatic speech: "John: **This one** / this one shows less variation. /"
[ ]	Square brackets can be used to add words that would facilitate the comprehension of the transcript: "John: **This one** [biome No. 1] / this one shows less variation /"
( )	Parentheses to indicate nonverbal cues and actions: "Carole: This one (pointing to one of the climatograms)?"
<u>Underlining</u>	To mark simultaneous or overlapping speech: "John: There is less / less <u>variation</u> Paul: <u>OK</u>, biome 1 looks like it receives more precipitation."
Comma & period	"," and "." to indicate breaks in the flow of speech.
Question mark	"?" if the context allowed the speech act to be interpreted as a question.
(...)	To indicate that words were undeciphered. "Bill: (...) but it's biotic."
/	Pause of less than 2 seconds.
/ /	Pause of greater than 2 seconds.
Italics	English words used during the discussion which was conducted in French "Derek: *Well,* I don't think so".

Example of a Transcript

Derek:	*Well,* I don't think so, because / there is nothing // living.
Bill:	(...) but it's biotic.
Kelly:	I think that the opposite is true. Because abiotic means dead / but there are all kinds of bugs and all that which ...
Derek:	Which live inside.

---

* Based on W.-M. Roth & A. Roychoudhury (1992). The social construction of scientific concepts or the concept map as conscription device and tool for social thinking in high school science. *Science Education, 76* (5): 531-557 and on D. Edwards & N. Mercer (1987). *Common understanding in the classroom.* London: Methuen. Pseudonyms were used throughout the dissertation for the analysis of transcripts and other qualitative data sources.