LEARNING, KNOWLEDGE BUILDING, AND SUBJECT MATTER KNOWLEDGE IN SCHOOL SCIENCE

by

Jan C.W. van Aalst

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Curriculum, Teaching, and Learning
Ontario Institute for Studies in Education of the University of Toronto

©Copyright by Jan C.W. van Aalst 1999
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Learning, Knowledge Building, and Subject Matter Knowledge in School Science
Doctor of Philosophy, 1999 Jan C.W. van Aalst
Department of Curriculum, Teaching, and Learning
Ontario Institute for Studies in Education of the University of Toronto

ABSTRACT

Following a theoretical analysis of constructivist approaches to collaborative learning, a curriculum development model, the Learning to Knowledge Building Model, is proposed. Two empirical studies of student work with Computer Supported Intentional Learning Environment (CSILE) are then presented; these lend support to the model and explicate in detail the nature and extent of the knowledge developed by elementary school students. The first study is a content analysis of a database developed by a combined Grade 5/6 class as part of a unit on heat and matter, conducted after the children completed their work; the analysis assumes the point of view of a subject matter specialist in the field of the students' inquiry. The second study was conducted while the students' investigation was in progress, and takes the point of view of curriculum coverage; it involved a teacher who used a different model of database use, as well as different subject matter. The proposed LKB model is based on a distinction Bereiter and Scardamalia (1996a) have made between learning and knowledge building (i.e., progressive collaborative problem solving); its aim is to support the design and planning of curriculum units and classroom practices in which knowledge building is central. An important feature of the model is the attention given to ensuring that students learn to evaluate their knowledge and to ask the questions that can advance shared knowledge. Among the findings of the first study are: (a) students who wrote more notes that explicated their commonsense knowledge early in the unit, by means of mixed framework notes, tended to write more notes of high scientific merit later; (b) some of students in this category tried out their ideas in diverse problem contexts; and (c) they tended to dominate the discussions they started. The second study provides additional insight into the role of the teacher, and the potential role of subject matter specialists, in knowledge building; it also provides a complex picture of the class's performance and the extent to which the various probes used revealed different aspects of this complexity.
For Alida
ACKNOWLEDGMENTS

It is a pleasure to thank the network of family, friends and colleagues who helped one way or another to make the completion of this dissertation possible.

During two extended visits to Holland I could work without worries. Thank you Pa, Ma, Erik, Ger, and Carla for the various ways in which you assisted me during those visits; in Canada, Ed, Elizabeth, Joe, Diane, Leo, Juliet, Larry, Eva, Wanja, and Debbie were there for me. A special thank-you is due to the students who developed the databases I studied, as well as to their teachers ("John" and "Paul" in the text). Clearly I could not have done this work without your effort! I also wish to thank Lorin Peters, the first physics teacher I taught with, and my students at Bishop O’ Dowd High School in Oakland, California, for their role in my journey toward educational research.

My committee, consisting of Derek Hodson, Erminia Pedretti, John Ross, and chaired by Marlene Scardamalia, got their job done efficiently and effectively. Thank-you all! Derek, your teaching made a strong impact on my thinking about the history, philosophy and sociology (HPS) of science. HPS has become an important component of the dissertation and a strong professional interest. Marlene, I could not have wished a better supervisor — passionate about learning and research, intellectually stimulating, and always encouraging, patient and enjoyable to work with. I also feel most fortunate, Marlene, for the many opportunities you created for me to learn about the enterprise of educational research: not the studies, but writing large-scale proposals, reviewing others’ work, networking, and fighting for a new view of education.

I am grateful to Robbie Case for his course on conceptual and cognitive change. Carl Bereiter’s writing and the interactions I had with him were most helpful in my coming to understand a number of issues. I also wish to thank Jud Burtis, Kevin O’Neill, and Leo Elshof for several reliability checks, and members of the CSILE group for many interesting discussions and other ways they assisted my work: Mary, Jim, Kate, Susana, Stephanie, Caroline, Ben, Nancy, Earl, and Clare. Chronologically last — as he must be — I am grateful to the external examiner, Wolff-Michael Roth. Thank you, Michael, for your thoughtful and constructive comments on the dissertation.
Alida, I started to work with some of the ideas described in this dissertation eight years ago, when I began to use “active learning” approaches in my own science classes. The last two years, I spent full-time on this project. Most of the time it was fun, but at other times I would have preferred to watch TV, read a book, or visit friends. What was most fun was to see my ideas, which were rough at first, become better ideas, and then to see them fit into a comprehensive set of ideas — seeing how it all fit together. This happened by trying them out in various situations, talking to other people about them, presenting them at conferences, and just sitting quietly, thinking through arguments other people might have against my ideas. I got a big buzz out of thinking about what I think are important problems, and getting a little further with them. Of course, an important message of the dissertation is that we should create similar opportunities for students to improve their ideas!

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ........................................................................................................... 1
  1.1 Intentional Learning and Knowledge Building ................................................................. 2
  1.2 Physics Education Research ............................................................................................. 5
  1.3 Overview of the Dissertation ............................................................................................ 6
  1.4 Concluding Remarks ........................................................................................................ 8

CHAPTER 2: LEARNING AND KNOWLEDGE BUILDING ......................................................... 10
  2.1 Constructivism and Science Education ............................................................................. 10
  2.2 Philosophical Perspectives on Constructivism ................................................................. 13
    2.2.1 Empiricism, inductivism, and realism ............................................................................. 13
    2.2.2 Hypothetico-deductivism and progressive research programs ......................................... 13
    2.2.3 Socio-historical and socio-cultural considerations .......................................................... 15
    2.2.4 Explanative coherence ................................................................................................ 18
    2.2.5 Popper's three worlds ................................................................................................. 20
    2.2.6 Summary .................................................................................................................. 21
  2.3 Psychological Perspectives on Constructivism ................................................................. 23
    2.3.1 Piagetian stage theory ................................................................................................. 23
    2.3.2 The role of working memory ......................................................................................... 24
    2.3.3 The learning paradox ................................................................................................... 24
    2.3.4 Connectionist models ................................................................................................. 25
    2.3.5 Situated learning .......................................................................................................... 26
    2.3.6 Summary .................................................................................................................. 27
  2.4 Constructivism and Conceptual Change ........................................................................... 28
    2.4.1 Initial knowledge and cognitive conflict ....................................................................... 29
    2.4.2 Epistemologies ............................................................................................................. 31
    2.4.3 Interpreting students’ utterances .................................................................................. 33
    2.4.4 Macroscopic and microscopic explanation ..................................................................... 35
    2.4.5 Strike and Posner’s Conceptual Change Theory ............................................................ 36
    2.4.6 Conceptual change across ontological categories ......................................................... 37
    2.4.7 Summary .................................................................................................................. 40
4.3 Basic Literacy Skills and Writing Output ........................................ 88
  4.3.1 CTBS scores ............................................................................. 88
  4.3.2 Writing output and quality ....................................................... 89

4.4 Case Studies ................................................................................. 90
  4.4.1 Matt ......................................................................................... 91
    4.4.1.1 Establishing a research question ......................................... 91
    4.4.1.2 Helping others advance their understanding ......................... 92
    4.4.1.3 Researching the role of density in the interaction between heat 
            and matter ......................................................................... 93
    4.4.1.4 End-of-unit knowledge claims ............................................ 94
    4.4.1.5 Summary .......................................................................... 95
  4.4.2 Emma ...................................................................................... 96
    4.4.2.1 Early contributions (Tl) ....................................................... 96
    4.4.2.2 Convection, conduction, and radiation ................................. 97
    4.4.2.3 Helping others advance their understanding ......................... 97
    4.4.2.4 End-of-unit knowledge claims and summary ....................... 98
  4.4.3 Jerry ....................................................................................... 99
    4.4.3.1 Establishing a research question ......................................... 99
    4.4.3.2 Researching fire .................................................................. 100
    4.4.3.3 End-of-unit knowledge claims ............................................ 101
    4.4.3.4 Summary .......................................................................... 102
  4.4.4 Sandy ..................................................................................... 102
    4.4.4.1 Early writing ..................................................................... 102
    4.4.4.2 Later writing ...................................................................... 103
    4.4.4.3 End-of-unit knowledge claims and summary ....................... 104
  4.4.5 Summary ................................................................................ 104

4.5 Concept Articulation ..................................................................... 105
  4.5.1 Scientific terms usage ............................................................. 105
    4.5.1.1 Semantic aspects of term usage ......................................... 106
    4.5.1.2 Use of terms in notes .............................................................. 107
  4.5.2 Concept articulation by students with LFC achievement .............. 110
  4.5.3 Summary ................................................................................ 113

4.6 Discussion notes .......................................................................... 114
  4.6.1 'Is gas affected by gravity?' .................................................... 115
  4.6.2 The role of learner achievement and note type ......................... 116
  4.6.3 The role of power .................................................................... 122
  4.6.4 Checklist data .......................................................................... 123
  4.6.5 Summary ................................................................................ 125

4.7 Questioning and Commenting ..................................................... 125
  4.7.1 Questions ............................................................................... 125
  4.7.2 Comment notes ....................................................................... 127
CHAPTER 5: STUDY 2, A VIEW FROM CURRICULUM COVERAGE ....................... 141

5.1 Background ......................................................................................... 142
  5.1.1 The LKB model, assessment, and evaluation ................................. 142
  5.1.2 Three probes of students’ knowledge ........................................... 145
  5.1.3 The subject matter ......................................................................... 146

5.2 Method ................................................................................................. 148
  5.2.1 Subjects ........................................................................................ 148
  5.2.2 Materials and procedure .............................................................. 148
  5.2.3 Data and measures ....................................................................... 152
    5.2.3.1 Test scores ............................................................................. 154
    5.2.3.2 Conclusion note structure ..................................................... 154
    5.2.3.3 Conclusion note knowledge claims ....................................... 155
    5.2.3.4 Interview knowledge claims ................................................. 156
    5.2.3.5 Thinking-Type-in-Use ........................................................... 157
    5.2.3.6 Logical argumentation ......................................................... 158

5.3 Basic Skills and Overall Achievement .................................................. 159

5.4 Case Studies ....................................................................................... 163
  5.4.1 Group G1: Empiricism and the nature of science ......................... 164
    5.4.1.1 Content .............................................................................. 164
    5.4.1.2 Ideas about knowledge building ......................................... 168
    5.4.1.3 The nature of science ........................................................ 169
    5.4.1.4 Summary ............................................................................ 172
  5.4.2 Group B1: The science of rockets ............................................... 173
    5.4.2.1 Cars as rockets .................................................................... 173
    5.4.2.2 Knowledge building ........................................................... 175
    5.4.2.3 Dynamic lift ........................................................................ 177
    5.4.2.4 Summary ............................................................................ 178
  5.4.3 Discussion ...................................................................................... 179
5.5 Test Item Analyses ........................................................................................................... 180
5.6 Concept Articulation in the CSILE Transcript .......................................................... 185
  5.6.1 Scientific terms usage ............................................................................................. 185
  5.6.2 Distinguishing energy, power, and force ............................................................. 186
  5.6.3 Summary ................................................................................................................ 188
5.7 Discussion Notes, Questions, and Comments .......................................................... 188
  5.7.1 Discussion notes ..................................................................................................... 188
  5.7.2 Questions ................................................................................................................ 191
  5.7.3 Comments .............................................................................................................. 191
  5.7.4 Summary ................................................................................................................ 191
5.8 Discussion ....................................................................................................................... 192
5.9 Concluding Remarks .................................................................................................... 195

CHAPTER 6: IMPLICATIONS FOR CONTINUOUS IMPROVEMENT OF KNOWLEDGE
BUILDING PRACTICE ........................................................................................................... 197

  6.1 Summary and Discussion of Results .......................................................................... 197
    6.1.1 Theoretical underpinnings of the LKB model ....................................................... 197
    6.1.2 Studies 1 and 2 and the LKB model ................................................................. 199
      6.1.2.1 Phases 1 and 2 ............................................................................................... 199
      6.1.2.2 Phase 3 ........................................................................................................... 203
  6.2 Comparison with some other approaches that stress inquiry .................................. 205
    6.2.1 Fostering Communities of learners ..................................................................... 205
    6.2.2 ThinkerTools ....................................................................................................... 208
  6.3 Elaboration of the LKB model ................................................................................... 210
  6.4 Teachers, Researchers, and Subject Matter Specialists ......................................... 216
    6.4.1 Modeling knowledge building ............................................................................ 216
    6.4.2 Multicultural research ........................................................................................ 218
  6.5 Conclusion and Implications for Teaching and Research ..................................... 222
    6.5.1 The Matthew effect ............................................................................................. 223
    6.5.2 Data analysis and assessment ............................................................................ 225

REFERENCES ....................................................................................................................... 228

APPENDIX A: CONCEPT ARTICULATION FOR STUDENTS WITH HFC
ACHIEVEMENT(STUDY 1) ...................................................................................................... 251
# LIST OF TABLES

| Table 3.1: | Participation of boys and girls in class meeting (Feb. 5, 1997) .......... 68 |
| Table 4.1: | Overview of component analyses, variables, and reliability measures, study 1 ................................................................. 78 |
| Table 4.2: | Allan's writing profile for T2, showing three pieces of evidence supporting and one refuting the hypothesis that Allan's writing passed the test of acceptability to a subject matter specialist .......... 80 |
| Table 4.3: | Characteristics of notes written by Grade 5/6 students with Low Focus on Concepts and High Focus on Concepts achievement levels ......................................................................................... 90 |
| Table 4.4: | The number of note entries with terms that could be used in scientific discourse, for students with Low- and High Focus on Concepts achievement levels, early and late in the unit .......... 108 |
| Table 4.5: | Concept Articulation by students with Low Focus on Concepts achievement ......................................................................................... 111 |
| Table 5.1: | Science vocabulary based on science kit from local school board ... 149 |
| Table 5.2: | Overview of component analyses, variables, and reliability measures, study 2 ......................................................................................... 153 |
| Table 5.3: | Summary of quantitative measures for force and energy unit .......... 159 |
| Table 5.4: | Question 4-A on written test, force, energy, and power .......... 181 |
| Table 5.5: | The number of note entries with terms that could be used in scientific discourse ......................................................................................... 186 |
| Table 5.6: | Articulation of the relationship of energy, power, and force .......... 187 |
| Table B.1: | Scale used to grade written test ......................................................................................... 257 |
| Table B.2: | Responses to Question 4-B, arguments about gravity and gases .... 259 |
LIST OF FIGURES

Figure 2.1: Example of a connectionist network ......................................................... 26
Figure 2.2: The Learning to Knowledge Building Model ........................................... 43
Figure 3.1: Example of a discussion note in CSILE, version 1.5 ....................... 54
Figure 3.2: Knowledge Map at the end of John's "Heat" unit in 1994/95 .......... 57
Figure 4.1: Conceptual framework for examining "Heat and matter" database. .......................................................................................................................... 73
Figure 4.2: Note entries of 'What is heat?' discussion note ................................ 118
Figure 4.3: Entries of 'How is heat made?' discussion note ................................ 120
Figure 4.4: Percentage of entries contributed to 29 discussion notes created by students with Low Focus on Concepts achievement, and 25 created by students with High Focus on Concepts' achievement ... 124
Figure 4.5: The number of questions of each of three form levels ................. 126
Figure 4.6: Mean percentage of a student's notes in each of five note categories, early and late in the unit ................................................................. 129
Figure 4.7: Mean percentage of a student's notes that were Mixed Framework, for students with Low- and High Focus on Concepts achievement levels, early and late in the unit. .................. 130
Figure 4.8: The mean contribution individual students made to the collective body of Mixed Framework notes, expressed as a percentage of the total number of notes in the database, early and late in the unit ................................................................. 132
Figure 5.1: Standard group scores of No. of notes, test scores, conclusion note structure, conclusion note knowledge claims, and interview knowledge claims ................................................................. 160
Figure 5.2: Knowledge claims made during end-of-unit group interviews. ..... 162
Figure 5.3: Percentage of responses to Question 2 of test that are of each of seven Thinking-Types-in-Use. ................................................................. 183
Figure 5.4: Types of arguments for students who answered yes and no to the question, "Is gas affected by gravity?" ........................................ 184

Figure 5.5: Susan’s What is energy? discussion note............................... 190

Figure 6.1: Support for Phases 1 and 2 from theory and studies................ 200

Figure 6.2: Support for Phase 3 from theory and studies .......................... 203

Figure 6.3: Two-unit LKB model ............................................................. 215
CHAPTER 1

INTRODUCTION

This dissertation reports on explorations into teaching and learning with Computer Supported Intentional Learning Environment (CSILE), a communal database system designed by Marlene Scardamalia, Carl Bereiter, and their co-workers (Scardamalia et al. 1989; Scardamalia & Bereiter, 1993) to support what they have called intentional learning (see also Bereiter & Scardamalia, 1987a, 1987b, 1991; Scardamalia, Bereiter, & Lamon, 1994). Since CSILE was introduced in schools more than a decade ago, its theoretical underpinnings, its design features, and the ways in which it is used in schools, have all undergone continual development. For example, intentional learning has evolved into the idea of knowledge building (Bereiter & Scardamalia, 1993) and the Knowledge Society Network (Scardamalia & Bereiter, 1996). These ideas are gradually informing an educational theory (see Bereiter, in preparation; Bereiter & Scardamalia, 1996a). The aims of the dissertation are:

- To understand better theoretical and practical issues connected with knowledge building — progressive inquiry — in classrooms in which CSILE is used, and in which the teacher is a co-inquirer with the students.

- To formulate a curriculum development model that is to articulate what sorts of activities and issues need to be considered in designing a progressive inquiry unit, while being sufficiently flexible to be useful with a variety of technologies aimed at supporting collaborative learning.

To achieve these aims, the dissertation begins with an analysis of theoretical issues relevant to knowledge building. This analysis concludes with a preliminary presentation of a curriculum development model, the Learning to Knowledge Building (LKB) model. Following Bereiter and Scardamalia (1996a), I make a distinction between learning (traditionally associated with the acquisition of externally developed knowledge) and knowledge building (progressive inquiry on communal problems). This analysis is followed by studies of two CSILE databases developed by Grade 5/6 classes; these studies serve to identify — from a practitioner's point of view — strong and weak knowledge building practices in two classrooms. The LKB
model (as presented in chapter 2) and the studies inform theorizing in the final chapter of the dissertation.

In the remainder of this brief introductory chapter, I describe intentional learning and knowledge building, relate the dissertation to my prior work, and give an overview of the chapters that follow.

1.1 Intentional Learning and Knowledge Building

I want to begin from a context that extends beyond K-12 science education. There are several reasons for this. First, one of the most interesting "affordances" (Gibson, 1977) of Internet-based collaborative learning technologies is the possibility of making accessible to students and teachers in schools the expertise of the community. Once, when I taught broad-based communication technology, I had an interesting discussion with a parent, who happened to be a graphic artist. Being somewhat outside my own field, I had made a comment that it was difficult to predict from what was on a computer screen how an inkjet print would turn out. He noted this was a problem that graphic artists were struggling with as well, and could not really help me with it. But he did give me some tips that helped me in my teaching of unfamiliar subject matter. I had thought how wonderful it would be if teachers could tap into expertise such as he possessed more easily.

A pilot study (Lamon, Reeve, & Scardamalia, 1997; van Aalst, 1997) conducted as part of an effort in the direction of making multiple communities of learners accessible to classrooms, the Knowledge Society Network (Scardamalia & Bereiter, 1996), revealed a number of noteworthy effects. Elementary school teachers concerned with a unit on the science of flight, several Grade 5/6 students, members of the CSILE research team, and university students who had just completed a course on fluid dynamics, shared a database, using an Internet-based version of CSILE (WebCSILE). Each of these groups used the database for its own purpose, without necessarily being concerned with helping the other groups advance their knowledge. However, one would expect that children would benefit from interacting with an adult, as in tutoring relationships (O'Neill, 1997), provided that the adults were willing to spend time to answer questions from the children.
Surprisingly, perhaps, interactions with the children also led to advancement of the university students' inquiry, which was aimed at understanding better the content of their fluid dynamics course. A pre-service teacher who co-taught the unit was a licensed pilot, and the mix of practical and theoretical knowledge about flight in the database led to deeper understanding of flight for him, as well as for others who shared the database. Of course, opportunities like this to share expertise with school children need not be limited to specialists — model plane hobbyists, for instance, have much expertise that could be tapped this way.

Second, challenges such as understanding better social aspects of collaborative learning, including the role of the teacher, have appeared in a wide variety of contexts — across subject matter domains, with children as well as adults, and with different collaborative learning technologies (van Aalst & de Jong, 1997). They are not limited to learning in elementary schools. Finally, explanations based on knowledge from everyday practices can play an important role in conceptual change by providing elements that can become useful in scientific explanations (diSessa, 1993; Minstrell, 1992).

Collaboration, while important, is only secondary in the design goals of CSILE, viz. to provide permanent support for what Bereiter and Scardamalia (1987a, 1993) have called ‘intentional learning’ and ‘knowledge building’.

Intentional learning describes what Bereiter and Scardamalia (1993) call “the career path” of people who are on their way to becoming experts. Such people: (a) know the limits of understanding and skill of their fields, (b) engage in progressive problem solving by continually reinvesting freed-up information processing capacity in new learning, and (c) focus on the goal of advancing the field. Intentional learning is similar to what Salomon and Globerson (1987) have called ‘mindfulness’. It characterizes work in the learned disciplines as well as learning in a variety of everyday goal-oriented situations. One example is operating a business. Someone aspiring to operate a business usually begins by learning such things as sales, customer service, advertising, and market research. He or she will not master such skills out of context but by “learning the business.” With time, the nature of the business may change. For instance, the growing use of the Internet for conducting business has forced banks to offer financial services via the Internet. Laws intended to protect the environment have had an impact on manufacturing
processes and the cost of producing certain products. The business operator cannot afford to ignore such developments, and is constantly attempting to react to the changing conditions in ways that can give the firm an advantage over the competition (or at least keep it viable). In addition, the business operator tries to develop an advantage over the competition, even when such changing conditions do not yet exist. Other learning situations to which intentional learning is relevant arise from hobbies such as playing a musical instrument and photography. Not every CEO, cellist, or amateur photographer learns this way, but some do. Surprisingly perhaps, experts in a variety of fields do — chess masters (Chase & Simon, 1973; de Groot, 1965), scientists (Chi, Feltovich, & Glaser, 1981; Larkin, McDermott, Simon, & Simon, 1980; see van Heuvelen, 1991 for a review), writers, musicians, and so on.

But intentional learning, as it has been described here, is inadequate for accounting for all expertlike learning. While it may describe the work of the CEO, it does not sufficiently stress social factors to describe the work of individual employees within a firm construed as a ‘learning organization’ (Argyris, 1994; Senge, 1990). To be sure, intentional learning is social. For instance, it involves knowing the edge of the competition’s competence. But it is not social in the manner that team play and science are social. Knowledge building (Bereiter & Scardamalia, 1993, 1996a) is a concept that addresses this shortcoming. It is intentional learning situated in a community (or society). Applied to science education, it may be characterized as follows:

- Knowing the current limits of the collective understanding of the learning community (and of learning communities competing with it).

- Whenever a problem becomes understood at one level, reinvesting the information processing capacity that becomes available to learn about the problem at deeper levels or in more varied situations.

- Focusing on the goal of advancing collective knowledge of the learning community.

With knowledge building come several responsibilities for the individual learner: (a) to use intentional learning to develop specialized knowledge to a level acceptable to the learning community, (b) to share that knowledge with that community, and (c) to participate as a member of that community in the process of knowledge validation.
1.2 Physics Education Research

I became interested in constructivist approaches to learning physics early in my teaching career, in 1989, when I realized that even students who excelled at all the tasks assigned to them did not always develop a very deep conceptual understanding of the subject matter. They had simply learned to complete those tasks competently. This is a striking feature of learning in school (Bereiter & Scardamalia, 1993) that persists into university physics education. For instance, diSessa (1993) has given an account of an interview with what he called an "excellent" first-year student taking physics at MIT, who did not have problems with concept acquisition, but nevertheless did not use the sort of analyses that physicists use (e.g., invoking the central concepts of the domain). DiSessa concluded that:

the fact that even this relatively sophisticated novice could be satisfied with an analysis on the basis of a relatively noncentral p-prim such as the spring scale rather than moving to more physically central analyses is a strong indicator of the continued shallowness of novice explanation. It also indicates the difficulty of building the deep explanatory network of experts, deep in the sense that experts can always be more careful, retreating at need to sanctioned and richly encoded notions such as physical quantities and laws. (1993, p. 161)

(P-prims are discussed in section 2.4.2.)

Between 1990 and 1994, I worked (as a teacher) on an inquiry-based approach to physics education that made use of Microcomputer-Based Laboratory (MBL) tools (Mokros & Tinker, 1987; Thornton & Sokoloff, 1990), which was an adaptation and extension of Workshop Physics (Laws, 1991) for use in high school settings (van Aalst, 1992, 1994). Somewhat later, following Hodson's (1990) lead, I began to argue for more open-ended inquiry (van Aalst, 1995), incorporating modeling tools such as Maple \( V^\text{TM} \) computer algebra software and Interactive Physics\(^\text{TM} \). When I first learned about the CSILE project, I became interested in the potential CSILE might have to support discourse around open-ended inquiry in physics education. That is why I set out to examine aspects of classroom work with CSILE in science education.
1.3 Overview of the Dissertation

Theoretical analysis: Chapter 2 consists of a theoretical study of constructivism based, in part, on my classroom work with MBL and my reading of the literature on expertise, philosophy of science, and students' ideas about natural phenomena. The chapter mentions two problems with constructivism that have previously been identified in the literature (Matthews, 1994; Nola, 1997a; Osborne, 1996), and makes an attempt to disentangle them. It also examines Chi's (1992) theory of conceptual change across ontological categories in light of Klaassen's (1995) reinterpretation of the meaning of students' utterances in protocol studies. The analysis concludes with the proposal of the Learning to Knowledge (LKB) model. Briefly, this model has the following features:

- An important role for the public debate of those aspects of commonsense knowledge (i.e., knowledge resulting from everyday experience) that can become useful as part of scientific explanations (diSessa, 1993, Minstrell, 1992).

- *Learning differentiated from knowledge building* (Bereiter & Scardamalia, 1996a), with a shift from the former to the latter during a unit. Learning is assumed to be goal-oriented and teacher-directed; knowledge building is assumed to be goal-oriented, open-ended, and student-directed.

- An important role for intentional learning.

Chapter 3 describes aspects of the development of CSILE and research on teaching and learning with CSILE, and contextualizes the empirical studies.

The role of the studies (chaps. 4 and 5) is two-fold. First, the analyses lend support to the LKB model, in addition to the argument of Chapter 2. Second, the studies explore classroom work with CSILE and ways of investigating it. The data were obtained from two Grade 5/6 classrooms taught by different teachers, but in the same school. Both studies are qualitative, although some of the data are quantified and analyzed statistically. A preliminary description of the studies and their mutual relationship follows.

*Study 1 (chap. 4):* In study 1, I take the point of view of a subject matter specialist, by which I have the following background in mind (in the context of the dissertation):
Formal training in the scientific field of the topic to be studied, including awareness of recent or current research. A Masters degree in science, and limited research experience.

Several years of experience teaching the topic of interest. This implies knowledge of what students typically tend to be able to accomplish at various grade levels with the topic, as well as familiarity with the literature on concept acquisition.

Some experience with learning environments that have constructivist underpinnings, as well as experience with introducing new educational practices into a classroom/school.

The study is an analysis of the text contributed to a CSILE database by a combined Grade 5/6 class while it studied the interaction of heat and matter; it was conducted after the class had concluded its work. An important component of the analysis is the rating of notes in the database on the basis of subject matter knowledge and knowledge of the literature on children’s ideas about science in this domain. The research questions are:

- To what extent does children’s writing, using CSILE and having available a teacher committed to knowledge building values, indicate scientific understanding, as judged from CSILE transcripts and knowledge of the literature on children’s ideas in science?

- What is the relationship between the knowledge building strategies the students used in the database and the scientific understanding they articulated?

Among the findings are:

- Approximately 1/3 of the class was judged as having provided evidence for knowledge building and/or for advanced understanding of some topics.

- Students who wrote more scientific explanations than other students late in the unit began writing early, particularly notes that contain commonsense as well as approximations to scientific ideas. Such notes will be called Mixed Framework notes.

- There was differential participation between such students in discussions, depending on who started the discussion.

- There were occasions when students were close to articulating an advanced understanding, but either did not know this or else had no access to information that could have allowed them to do this.

*Study 2 (chapter 5):* This continues to evaluate classroom work with CSILE as science education by exploring some of the issues left out or suggested by Study 1 —
with different subject matter, a different teacher, and a different Grade 5/6 class. The study uses the AAAS (1993) benchmarks and locally designed curriculum objectives to examine the knowledge students revealed at the end of their work with CSILE. Goals of the study include:

- Obtaining first-hand experience in a classroom with a teacher who used a traditional learning model, rather than one based on intentional learning.
- Exploring additional probes of concept acquisition articulation and knowledge building.
- To better understand the role of the teacher, and the potential role of outside specialists to support the class's knowledge building activities.

The purpose of chapter 6 is to examine the LKB model in light of studies 1 and 2. The chapter discusses pedagogical issues of knowledge building in more detail than chapter 2 does; it also examines how the LKB model fits with other work on progressive inquiry, and proposes a number of problems that can be studied in classroom-based work with Knowledge Forum™, the second-generation version of CSILE that was introduced into classrooms in September, 1997.

1.4 Concluding Remarks

It seems important to underscore that the purpose of the dissertation is not primarily to argue for the educational effectiveness of CSILE, but rather for processes of continual improvement that are fundamental to its design and use. It aims to create a general structure of a curriculum unit aimed at knowledge building, and forms a basis for continual development. The CSILE project represents one particular way of addressing the issues, but ways are consistent with the LKB model (e.g., Brown & Campione, 1994, 1996; Roth, 1995, 1998; Roth & Bowen, 1995; White, 1993a, 1993b). This generality is a strength of the LKB model, but it limits its practical usefulness — it remains to develop from the model a specific instructional program that can be tested. It also requires underscoring that the studies focus on the database discourse, and as such, do not examine in detail the face-to-face processes of knowledge building (for discussions, see Latour, 1987, Roth & Duit, in press). I made this choice in order to learn more about the extent to which data
naturally produced by the database discourse (e.g., the text the children enter into the database), supplemented by a small amount of researcher presence in the classroom could inform the next steps toward more effective knowledge building practices. Future ethnographic studies, using the LKB model as a scaffold for continual improvement, may develop a fuller picture of knowledge building processes.
CHAPTER 2

LEARNING AND KNOWLEDGE BUILDING

This chapter examines constructivism critically and develops the Learning to Knowledge Building (LKB) model. The argument proceeds as follows. I state two problems identified in the history, philosophy, and sociology (HPS) of science literature in section 2.1. This is followed by a selective review of philosophy of science (section 2.2), psychology (section 2.3), and the treatment of constructivism in the science education literature (section 2.4). In section 2.5, I present the LKB model.

2.1 Constructivism and Science Education

The Twentieth Century has produced much interesting work in epistemology. Early in the century, scientific developments such as quantum mechanics led to scrutiny of the realist interpretation that had been given to science. In what sense, for instance, could one say that an electron was real? The philosophy of science that resulted no longer argued for science as a method for learning “Nature’s secrets,” but for science as a human enterprise — observations were theory- and value-laden, knowledge advances were negotiated in communities of scientists, and scientists did not use some method without fail, but stayed committed to their ideas in the face of counter-evidence. In short, science, however useful to society, had become seen as a social construction. At the same time, in the West, Piaget argued on the basis of a long research program that children developed — or constructed — complex logical structures from simpler ones. It is in this context that a rather diffuse set of educational ideas referred to as ‘constructivism’ developed. As Solomon (1994) has pointed out, three factors were particularly influential in the rise of constructivism in science education: (a) Kelly’s theory of personal constructs, (b) the notion of “Children’s Science” (Driver, 1983; Driver & Bell, 1986), and (c) the idea of the social construction of knowledge. She noted that Kelly’s theory of personal constructs led to “personal constructivism” in New Zealand, a movement that had much in common with Children’s Science. In cognitive science, ‘constructionism’ predated
these developments with the work on LOGO (see Harrel & Papert, 1991, Papert, 1980).

The science education literature on constructivism frequently stresses the importance of direct experience and the interpretation of that experience:

To learn science from a constructivist philosophy implies direct experience with science as a process of knowledge generation in which prior knowledge is elaborated and changed on the basis of fresh meanings negotiated with peers and teacher. (Watts, 1994, p. 51)

Several authors have taken issue with two aspects of the literature on constructivism in science education (Matthews, 1991; 1994; Nola, 1997a, Osborne, 1996): (a) its inherent relativism, and (b) a lack of attention to constructive processes. Perhaps the most extreme relativist position is that of von Glasersfeld:

The word “knowledge” refers to a commodity that is radically different from the objective representation of an observer-independent world which the mainstream of Western tradition has been looking for. Instead “knowledge” refers to conceptual structures that epistemic agents, given the range of present experience within their tradition of thought and language, consider viable. (1989, p. 124) [emphasis in original]

A problem with this position is that what is “viable” may not be interesting or valuable: on what basis are we to judge whether or not a student has done an adequate job in learning about a topic? Moreover, there is no place for truth in von Glasersfeld’s “radical constructivism” (Propositional logic is build on propositions which have truth values.)

A second source of relativism in constructivism is Feyerabend’s (1975) work on scientific method. He argued that no methodological rules exist that scientists use without exception in their work. However, that does not imply that there is no method at all to science: there are heuristics (see section 2.2.2). A third source of relativism is Kuhn’s (1970) notion that scientific knowledge is validated through negotiation by communities of scientists: science, astrology, and religion provide distinct frameworks that communities use to make sense of the natural world. The influence of this source of relativism has become more important in the 1990s as the nature of science is once again becoming problematic in history, philosophy, and sociology of science (Matthews, 1998). This time the controversy is over insights deriving from advances in the sociology and anthropology of science. For example,
Solomon (1994) has observed that Collins and Pinch, and Bloor and others “were softening the rigorous image of scientific thinking until no hard line at all could be drawn … between ideas held by the learning child and those of the practicing scientist” (p. 10). Matthews has observed that feminism is no longer a political movement but is developing new epistemologies. And arguments have been made for greater recognition of the ways cultural factors influence what a students may make of Western science (see Cobern, 1996; Jegede, 1994).

In view of these developments it is necessary to state my position concerning the nature of science and its implications for science education. I assume that two conditions must be met before an approach to understanding nature can be counted as a (mature) science. First, it must be conceptual. It must do more than describe phenomena in “observational terms” (Driver, Squires, Rushworth, & Wood-Robinson, 1994), it must provide a super-ordinate, explanatory, framework. Ancient Chinese science (see Needham, 1969) does not meet this condition, although it is a common ancestor, with Greek science, of modern science. The second condition has to do with the process by which knowledge progresses, that is, its epistemology. As Bereiter (in preparation) points out, a crucial aspect of science is knowledge improvement. Scientific claims unlike claims based on religion, are always provisional: while some statements may at a given time not be questioned (i.e., taken on faith), they may be later if substantial doubt about their validity accumulates.

Grandy (1997) and Nola (1997a) have advised to examine constructivism by distinguishing between its epistemic and psychological aspects; accordingly, in section 2.2 I review relevant epistemology, and in section 2.3 child development and learning. The ultimate goal of the discussion that follows is to articulate a theoretical perspective from which to construct (a provisional version of) the Learning to Knowledge Building (LKB) model.
2.2 Philosophical Perspectives on Constructivism

2.2.1 Empiricism, inductivism, and realism

In antiquity and medieval times empiricism was concerned with "saving the phenomena," with explaining observations. *Inductivism*, introduced by Francis Bacon as the method of science, took empiricism a step further and assumed that it is possible to generalize singular statements to universal statements. Before Bacon, theoretical science had been possible (it simply did not yield truth), but now it had become problematic: it was not inductive. One approach to saving theoretical science as knowledge was logical empiricism (or logical positivism), which insisted that theoretical ideas were reducible to observational terms. Another fundamental assumption of modern science became its realism. The referents of scientific theories were assumed to have a physical reality: terms like electron and photon were not just theoretical constructs, they were assumed to refer to *things-themselves*.

Two problems with empiricist/inductivist science are that (a) multiple theories may explain the same phenomenon, and (b) things are not necessarily as we see them. The first of these is known as the Duhem-Quine thesis (Duham, 1906/54) and the second as "blissful empiricism" (Science Council of Canada, 1984, p. 24). In the 1930s, science was criticized because its referents could not be observed (in the sense that term was then understood); this led first to positivism, then logical positivism, and finally to hypothetico-deductivism (see Duschl, 1990, p. 34). Logical positivism had as its crucial elements empirical evidence and propositional logic; in philosophy it was short-lived, but its impact in science education remained visible in those "discovery learning" curricula in which students were encouraged to observe or discover natural phenomena and scientific concepts "without any understanding of the fundamental concepts or principles needed for seeing and discovering" (Duschl, 1990, p. 33).

2.2.2 Hypothetico-deductivism and progressive research programs

Karl Popper's theory of the validation of scientific knowledge was about constructing and testing hypotheses; it was a reaction to logical positivism and did not imply that data precedes theory. It was a *normative* theory that did not so much make claims about what scientists did in practice as it attempted to delineate the conditions under which one could say one had scientific knowledge. Popper did not
insist that a scientific theory be capable of being singled out in a positive sense. Rather, he required that "its logical form be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical scientific system to be refuted by experience" (1968, p. 41). A refutation, he argued, proceeds along the following lines:

With the help of other statements, previously accepted, certain singular statements — which we may call "predictions" — are deduced from the theory. ... From among these statements, those are selected that are not derivable from the current theory, and more especially those which the current theory contradicts. Next we seek a decision as regards these ... and other derived statements by comparing them with the results of practical applications and experiments. If this decision is positive ... if the singular conclusions turn out to be acceptable, or verified, then the theory has, for the time being, passed its test: we have found no reason to discard it. But is the decision negative, if the conclusions have been falsified, then their falsification also falsifies the theory from which they were logically deduced. (p. 33)

Popper did not imply that every scientific statement is tested before it is accepted, only that it is capable of being tested (p. 48).

The trouble with Popper's normative theory was that scientists did not live by it. In a long-standing debate about the rationality of science with Lakatos, Feyerabend used an analysis of the Copernican revolution to argue that science is irrational — that scientists do not follow a set of rules (i.e., a method) without exception. He concluded a long discussion of the replacement of the Ptolemaic world view by the Copernican one with

... and he will perhaps see the merits of a different view which asserts that, while pre-Copernican astronomy was in trouble (was confronted by a series of refuting instances and implausibilities) the Copernican theory was in even greater trouble (was confronted by even more drastic refuting instances and implausibilities); but that being in harmony with still further inadequate theories it gained strength, and was retained, the refutations being made ineffective by ad hoc hypotheses and clever techniques of persuasion. (1975, p. 143)

Where Feyerabend saw irrationality, Lakatos saw a kind of rationality. According to Lakatos's view, scientists remain committed to a line of research as long as there is progress; he made the unit of analysis the progressive research programme, with its positive heuristics (telling us what paths to follow) and its negative heuristics (telling us what paths to avoid).
All scientific research programmes may be characterized by their "hard core." The negative heuristic of the programme forbids us to direct the *modus tollens* at this hard core. Instead, we must use our ingenuity to articulate or even invent "auxiliary hypotheses," which form a *protective belt* around this core, and we must redirect the *modus tollens* to these. It is this protective belt of auxiliary hypotheses which has to bear the brunt of tests and get adjusted and re-adjusted, or even completely replaced, to defend the thus hardened-core. A research programme is successful if all this leads to a progressive problemshift; unsuccessful if it leads to a degenerating problemshift. (Lakatos, 1970, p. 133)

The positive heuristics are not hard rules that guarantee success, but they do allow a scientific method. Lakatos stressed that we must be careful not to abandon a research program early in its life if it is unable to overtake a powerful rival (1970, p. 157); and scientists may stay with established theories that are in trouble. Dutch (1982) has extended Lakatos's theory by introducing an inner hard-core and an outer soft-core: movement can be only *into* the hard-core (not out of it), but for the soft-core motion in both directions is possible.

### 2.2.3 Socio-historical and socio-cultural considerations

Lakatos’s progressive research program is an important analytical invention because it frames individual studies in a *historical context*: a single rejected hypothesis may be brushed aside, particularly early in a research program, but if hypotheses are consistently rejected, the research program may become regressive, and scientists may abandon it. As, such, Lakatos continued a line of argument introduced by Thomas Kuhn in *The structure of scientific revolutions* (1962/70). Kuhn’s thesis was that what is to count as scientific progress is decided by communities of scientists in a process involving consensus and dissensus. To support his argument, he introduced several influential ideas: the notion of the ‘paradigm’ and the ‘paradigm shift’, and a distinction between ‘normal science’ and ‘revolutionary science’. He argued that a science begins in a pre-paradigm state — he cited chemistry before Lavoisier, and electricity before Franklin — in which there are no established theories and methods. The history, the basic ideas, and the successes of a particular theory are usually reviewed whenever the theory, or one of its applications, or a modification to it, is discussed. As a result of debate, some of these theories may be "shaken out" and a theory emerges as the survivor; the training of new scientists, methods of research, and even the kinds of problems that may be addressed by the science become more clearly defined (through negotiation in the scientific community). Kuhn referred to this situation as the first paradigm.
Kuhn referred to research in a well established paradigm as 'normal science', which consists of "puzzle solving." An example of such puzzle solving is theoretical and experimental research into high temperature superconducting materials, following its initial manifestation in Ba-La-Cu-O ceramic materials (Bednorz & Müller, 1986); which has not led to new theories that form the basis of condensed matter physics. As Kuhn points out, however, there have been several periods in the history of science when the number of unsolved puzzles increased dramatically. An example of this occurred at the end of the Nineteenth Century — Balmer’s unexplained empirical formula for the spectrum of hydrogen (1885), Hertz’s discovery of the photoelectric effect (1887), and Becquerel’s discovery of radioactivity (1896), to name a few (see Pais, 1986, 1991). In such a period, some scientists will step outside the established methods and attempt to solve these puzzles. Eventually, outside the paradigm a situation similar to the pre-paradigm state develops. Analogical reasoning and ad hoc assumptions are important strategies in such a period (e.g., Perrin’s 1903 model of the atom similar to the solar system, and Planck’s 1900 assumption of quantized radiation). With time, some coherence begins to form in the variety of ideas and explanations that exist, and the field becomes reconstructed from "new fundamentals." There now begins to be an alternative to the existing paradigm. Kuhn called science in this period ‘revolutionary science’. He considered competing paradigms to be incommensurable and saw a switch of a scientist’s commitment from one to the other — a paradigm switch — as a gestalt switch. As Latour (1987) comments, Kuhn’s paradigm switch has a considerable likeness to a political revolution. Kuhn argued that paradigm shifts could occur if evidence in favor of the new theory accumulated and, in addition, it left open possibilities for future research.

The idea of a paradigm shift has been popular with science educators because it provides a framework for thinking about what children must accomplish to learn science (Carey, 1986; Duschl, Hamilton, & Grandy, 1992; Garrison & Bentley, 1990; Posner et al, 1988; Strike & Posner, 1992). Whether learning science for children is more like a paradigm shift or more like a transition to a first paradigm is not clear. Kuhn’s framework also underscores that science is a social construction (or institution). His paradigm shift has been severely criticized by philosophers of science. He used this term rather indiscriminately, sometimes implying world view and at other times theory. The shift implied was, others suggested, more piecemeal than Kuhn argued (Laudan, 1984; Nersessian, 1989). Gerald Holton, a former
colleague of Kuhn at Harvard University, was strongly opposed to the notion of a paradigm shift: he saw, instead, "an evolutionary process in which the current theory-determiners are forever linked in sometimes subtle ways to their mentors, their backgrounds, and, consequently, to some new parameters of scientific integrity" (Holton, 1986, quoted in Loving, 1992, p. 53).

Kuhn elaborated on the paradigm shift in the 1970 edition of *The structure of scientific revolutions* by introducing a 'disciplinary matrix', consisting of four parts: (a) formal logical components, (b) beliefs in particular model components, (c) value components, and (d) exemplar components. Changes in conceptions occur via changes to parts and not the whole (see Duschl & Gitomer, 1991, p. 843). (According to Nola [1997b] he discontinued using the term 'paradigm shift' himself.) In addition, Kuhn stressed the importance of consensus, but did not explain in detail how it is reached. Other authors have taken this issue up. For instance, Longino (1994) proposed that four conditions must be met for consensus to be called knowledge: (a) there must be publicly recognized forums for criticism of evidence, of methods, and assumptions and reasoning; (b) there must be uptake of criticism; (c) there must be publicly recognized standards by reference to which theories, hypotheses, and observational practices are evaluated and by appeal to which criticism is made relevant to the goals of the inquiring community; and (d) communities must be characterized by equality of intellectual authority (Longino, 1994, p. 144-145). Laudan (1984, cited in Duschl & Gitomer, 1991, p. 844) has argued for a triadic notion of change involving theory, methods, and aims. The triadic network abandons the hierarchical models of change which suggest that a change in theory automatically results in changes in method and aims; it allows for the possibility that scientists working in a discipline may alter theoretical commitments, but still maintain existing methodological and axiological commitments developed from a previous theoretical framework.

The social discourse by which scientific advances are produced has received considerable attention in the last two decades, and is beginning to provide a fuller theoretical framework for understanding science. Much of this sociological and anthropological research uses a socio-historical perspective and ethnographic methods to study scientists and other practitioners in action (Dunbar, 1995; Hutchins, 1991; Latour, 1987; Lave & Wenger, 1991; Pera, 1994). Woodruff and Meyer (1997) have observed that (virtually) all scientists belong to two types of
community: micro- and macro-communities. Such micro-communities are the biology laboratories studied by Dunbar, who observed that successful laboratories had a high degree of "mutuality" (Damon & Phelps, 1989) in their conversations, where members could complete each other’s sentences:

Many members of the laboratory reasoned about the research, and often the results of one person’s reasoning became the impetus to another person’s reasoning. This resulted in rapid reconceptualization of problems and in significant changes in all aspects of the way the research was conducted. Situations in which group problem solving occurs provide a rich example of the way that cognitive and social mechanisms interact. (Dunbar, 1995, p. 388, quoted in Woodruff & Meyer, 1997, p. 29)

Thus, within laboratories social networks can function to support problem solving cognitively (several scientists discuss a shared problem), and emotionally. In the situation Dunbar describes in the above quote, there is a sense that the laboratory is trying to improve its knowledge.

Latour’s (1987) analysis of laboratories is important for several reasons. First, the social dynamics between laboratories are necessarily more competitive than within laboratories: — laboratories must compete with each other for funding and recognition. Very often, the director of a laboratory must make decisions for political reasons that are in the long-term interest of the laboratory. An example of this is Marie Curie’s reluctance to publicize and research quite evident health hazards associated with radioactivity before funding for the Radium Institute was firmly established (see Quinn, 1995). Also important for education is that science entails much more than working at a lab-bench or theorizing. Funding must be obtained, results disseminated, and so on, so that the lab scientists can continue to do science; Latour argues that organizations like the National Science Foundation are doing science, just as the bench scientists are. An implication of this for science education is that doing science can take many forms, but nearly all require substantial knowledge of the claims and methods of science.

2.2.4 Explanative coherence

Although understanding socio-cultural and socio-historical phenomena are central to a full understanding of knowledge building, I choose to study them by representing them in the LKB model by components, or via the complexity of the components. In this section, therefore, I continue to examine knowledge from a
rational perspective (for a related discussion, see Case, 1996a). Kant held that there are two incommensurate sources of knowledge — formal knowledge and knowledge derived from experience. Formal knowledge was true *a priori*. In fact, Kant argued that space is real, but

at the same time, it [Kant's argument] teaches the *ideality* of space, when things are viewed as only the mind can view them, as they are by themselves, apart from the activity of human sense. ... The *formal* idea of phenomena in space is a critical reminder that there is nothing of which one is aware that is a thing-itself (that is, something apart from man's perception of it). Space is not the form of things-themselves. The phenomena of which we are aware tell us nothing about things as they are apart from us, and in experience nobody ever asks about them as such.¹

Thus knowledge of things-themselves is impossible. What are the other possibilities? Foundation theories of justification assert that there are some infallible truths, but according to Lehrer (1990) there are far too few infallible statements to support commonsense knowledge. Another possibility is explanatory coherence theory, which does not allow foundational statements. In this theory, statements are accepted on the basis that they cohere with other statements, and the best argument is the one that explains the most. A widespread criticism of coherence theory has been its inherent circularity (Lehrer, 1990). Both the statements to be explained and those that do the explaining are accepted on the same criterion — coherence with other statements. Thagard and Verbeurgt (1998) have replied to that criticism that coherence-based inference does not depend on steps, but proceeds by simultaneous evaluation of multiple elements. According to them, foundational statements are not needed. They have used connectionist methods (see section 2.3.4) to attempt to make the idea of coherence more precise, interpreting it in terms of the connection strengths and activation levels of connectionist models: the best argument is represented by the network that satisfies the most constraints. Explanative coherence theory is quite powerful in practice. Thagard (1992) has used it to simulate several classic arguments from the history of science. In one case, the rise of plate tectonics, his algorithm explained both why plate tectonics was not accepted widely by scientists when it was introduced in the 1920s, and why it was in the 1960s and 1970s. But that simulation also points to a limitation — if the set of statements used is too small, it may fail to yield the argument that constitutes a conceptual advance.

¹See *An Immanuel Kant Reader* (Blakney, 1960, pp. 27-28).
For Lehrer (1990), the way out is a hybrid theory consisting of explanatory coherence theory and aspects of foundation and externalist theories. He uses the idea of acceptance for the purpose of avoiding error, arguing that we may accept statements if to do so would seem more "reasonable" than to reject them. As he puts it, "though we do not know of the existence of physical objects, we may reasonably infer their existence is the simplest and best explanation of why we experience sense data we do" (p. 91). But in his view, "external judges" are crucial. He suggests a probabilistic game in which one attempts to beat a skeptic, a position that is influenced by Bertrand Russell. Finally, as O’Gorman (1989, p. 51) notes, "when theories are in conflict, each theorist has only his own theory to go on in criticizing his opponents — all criticism is relative to one’s own theory."

2.2.5 Popper’s three worlds

A philosophical framework for constructivism can benefit considerably from a later achievement of Karl Popper. In his (1972) *Objective knowledge: An evolutionary approach* he distinguished between three worlds. ‘World 1’ is the space-time world occupied by things-themselves. ‘World 2’ is occupied by mental states (personal knowledge) and ‘World 3’ by what he called objective knowledge.

World 1 is accessible only by our senses. World 2 is shaped by three elements: (a) experience, (b) cognitive function, and (c) the ideas a person already has. People construct ideas about sense experience, and their ability to do this depends on the extent to which their cognitive functioning has matured/developed and on their current ideas. For example, dropping a piece of paper through some height, one may distinguish between laminar and turbulent flow of air, but the ability to do this depends on the ability to identify differences and similarities between the two kinds of flow (i.e. cognitive functioning), as well as on having theoretical ideas into which both kinds of flow can fit. An important point is that it is impossible to understand new sense data in the absence of existing ideas. Of course, while the constructive process always refers to sense experience, it does not necessarily require new sense experience — cognitive development can also result when a super-ordinate layer rises above an existing cognitive structure as a result of reflection. One begins to see that emphasis on direct (i.e., “hands-on”) experience in much constructivist writing is quite misleading. Rather, central to constructing understanding is *theory*
improvement (Bereiter & Scardamalia, 1996a, 1996b), thinking about theories, just as it is central to science (Bereiter, in preparation).

A crucial point about World 2 is that it is private. Discourse is possible only if some World 2 objects (mental states) are made the objects of arguments. That is what takes place in World 3. When people put their ideas in World 3, they distance themselves from them to some extent, so that the argument that follows is not personal — the point is to improve the knowledge. Therefore, objective knowledge, in the sense Popper used it, is neither necessarily true, nor unbiased. Consider as an example Richard Wagner's Ring of the Niblungs. At one level, there are the things-themselves (World 1 objects) we experience through the senses, such as a performance, the libretto, or a recording. At a second level, there are mental states associated with it: what Wagner intended, how he felt after its first performance, how he felt while he was composing it, my mood when I first saw it performed, the ideas about music I have derived from listening to it, and so on. Finally, it can be seen as a World 3 object — a musical work that stands out in a musical literature as having some things to say about human nature, and that makes unprecedented uses of leitmotifs. Of course, a performance can also be a World 3 object if viewed within a body of interpretations of what the work is about, what Wagner intended, and so on. The story Wagner had to tell about human nature in the Ring may not stand up to criticism and its musical elements may not be lasting, but they are discussible within appropriate conceptual frameworks.

2.2.6 Summary

How can the above discussion of philosophy of science inform the design of a model aimed at supporting knowledge building practices?

1. Inductive logic and progressive research programs: In individual studies, a hypothesis is usually tested against an alternate hypothesis. Therefore, an important goal of science education must be that children have opportunities to conceive, plan, and conduct experiments (Hodson, 1990; van Aalst, 1995), beginning from ill-defined problems (Roth & Bowen, 1995). However, they must also learn that an experimental study is part of a research program, and that a single rejected hypothesis often is not sufficient to abandon a line of research.
2. *Science is socially constructed:* The articulation of research problems, the methods by which they may be researched, by whom, for what purpose, and whether the results will count as an advance of scientific knowledge all are determined by culturally and historically determined discourse practices. A key element of these practices is the objectifying of knowledge claims in Popper’s World 3, where they can be scrutinized, debated, and improved by the scientific community. Students, to have authentic experiences of doing science (see Benzce, 1995; Hodson & Bencze, 1998), should have opportunities to engage in discussions of this kind, particularly discussions in which they formulate research questions that can advance the knowledge of the class, develop the class’s research program, and validate knowledge advances. If students work in small groups, they can experience both the affordances of micro-communities that Dunbar (1995) observed in laboratories, and the more competitive social phenomena of macro-communities described by Latour (1987).

3. *Tension between learning and doing science:* When students are learning what Bereiter and Scardamalia (1993) call the “formal knowledge” of science, they are not constructing new knowledge, but learning the knowledge a community has already produced. Here it does matter where students end up: ‘The rest mass of a proton is more than that of a neutron’ is not a statement a physicist would make in scientific discourse. In curriculum development, we must ask at any juncture if it is more important that students develop accurate understandings of formal knowledge or of the social phenomena of science — these cannot simultaneously be high priority goals. When the goal is that students learn about the social processes by which problems of understanding are articulated, researched, and the research findings debated in science, students must be allowed to end up drawing conclusions that are outside the formal knowledge of science. When the goal is that students learn the formal knowledge, the best we can offer, it seems to me, is what I have called “guided concept articulation” (van Aalst, 1995), but here the students do not articulate problems of understanding and validate knowledge advances.

4. *Explanative coherence:* This can provide a model for validating knowledge. One tries to fit new information with an existing theory, and makes modifications to either of these in such a way that the resulting theory is the most parsimonious and explains most.
2.3 Psychological Perspectives on Constructivism

In this section, I continue to develop a theoretical background for the LKB model, focusing on psychological considerations.

2.3.1 Piagetian stage theory

Piaget’s long-lasting interest in child development can be attributed to his early finding that the thought of young children was very different from that of adults, from which he concluded that it must undergo development. According to Piaget, in the sensorimotor period (0-2 years), the child applies intelligence to action and there is no thought (e.g., Piaget, 1967, p.78); language is for Piaget an example of representation and develops as a result of the child having acquired a sufficient amount of knowledge through action. Chomsky (in Piatelli-Palmarini, 1980), in contrast, assumed that sensorimotor experience merely triggers a native facility for language, which he compared with an organ (a heart).

Two lines of thought were crucial to Piaget’s thinking about the development of logical thought in children. First, in the history of mathematics one notices that mathematical structures develop in such a way that earlier structures are subsumed by later ones: Cartesian analytic geometry includes all aspects of Euclidean geometry (Piaget, in Piatelli-Palmarini, 1980, p. 150). This idea forms the basis for Piaget’s stages — it is not that the adolescent no longer applies operations characteristic of the concrete-operational thinker, but that he or she is now, in addition, capable of other operations. Thus, while in the concrete-operational period the learner uses two kinds of inversion, negation and reciprocity, the relationship between these operations is still lacking; in the formal-operational period, the learner realizes they are both part of a “whole structure” (the INRC group), uses all the operations in this group and understands their mutual relationships. Considering the idea that such a structure might be innate and simply “triggered” at perhaps age 15, Piaget could not see why, if it had been assembled at birth, this would take so many years (see Gruber & Vonèche, 1977, p. 473; Piatelli-Palmarini, 1980, p. 26). He opted for a constructivist interpretation in which the child develops this system as a result of his or her interaction with the world. The second line of thought refers to self-regulatory aspects of the theory of evolution and a dynamic interpretation of the concept of equilibrium. Within each stage, an identical process is at work in which
the learner moves toward an equilibrium position in which assimilation and accommodation compensate one another (Celletier, in Piatelli-Palmarini, 1980, p.70; Piaget, 1967, p. 114). Crucial to such a dynamic equilibrium is the notion of reversibility, which is expressed fully only in the formal-operational period.

2.3.2 The role of working memory

In The child and reality (1973), Piaget described a number of issues that were not addressed by his theory. He admitted that in some societies formal thought does not develop at all; that his theory attempted to explain a special case — logico-mathematical thought — and probably did not apply to the development of such things as language and perception (1973, p. 49); and that field effects and affective factors could impede or accelerate the development of propositional thought, although not its structure (p. 47). Nevertheless, many neo-Piagetian theories have retained developmental stages (see Bidell & Fisher, 1992; Case, 1985a, 1985b; Demetriou et al, 1992; Pascual-Leone, 1969).

A crucial factor not exploited by Piaget’s theory is that tasks differ according to the number of schemes a child must simultaneously hold in working memory for their completion, the so-called M demand of the task (Pascual-Leone, 1969). M capacity, the number of schemes a child can hold in working memory varies with age: children develop sufficient M capacity to cover executive function by the end of the sensorimotor period; thereafter it increases by one scheme every two years until it reaches seven at age 15 or 16. Pascual-Leone (cf. Chapman, 1981, p. 151) showed that a poorly explained result on a Piagetian task — the conservation of weight is observed approximately two years after the conservation of matter, although both tasks have the same logical structure — could be explained by M capacity and M demand. Scardamalia (1977) confirmed this finding for another horizontal décalage, concluding that “… when the M demand of the task is set relative to M capacity, subjects of all ages should perform qualitatively alike” (p. 29).

2.3.3 The learning paradox

Fodor raised an important problem at the Piaget-Chomsky debate; it has been called the “learning paradox” by Pascual-Leone (1976). The problem is this:
An inductive logic (that is, a theory of learning in the only sense in which there are theories of learning) can't tell you how the concept *miu* is acquired because it assumes the availability of that concept, when it assumes that *miu* occurs in the confirmed inductive hypothesis (Fodor, in Piatelli-Palmarini, 1980, p. 146).

Several partial solutions have been proposed. Schemes repeatedly held in working memory simultaneously can become assimilated to each other and are then activated together as one scheme (Case, see Juckes, 1991, p. 270; Pascual-Leone, see Chapman, 1981, pp. 146-147; Piaget, in Piatelli-Palmarini, 1980, p. 166). In addition, schemes can become assimilated to each other if the *conditions* under which they are activated are similar on repeated occurrences. The conditions might capture factors such as the learning situation, emotion, and interest. Bereiter (1985) noted that "humans have a need to use ... mental capacity" (p. 213), adding that this can be done by searching out other goals that might be served by the procedure at hand; the converse of this he called "piggy-backing," using existing but advanced capabilities that were developed for other purposes. In general, a new structure can rise from an existing one by changing the "activation values" of individual schemes and the weights that describe the effect of one scheme on another.

### 2.3.4 Connectionist models

For most of this century the mind has been characterized as a container of knowledge structures, often described as rules. Examples are Chomsky’s native facility for language, Piaget’s logical structures, and the treatment in the literature of misconceptions as knowledge structures (see section 2.4 for a review). The learning paradox may be seen as a consequence of this view of the mind. In the 1980s, however, an alternative characterization of the learning of rules was proposed: parallel distributed processing (PDP), also known as *connectionism* (Rumelhart, 1989; Rumelhart, Smolensky, McClelland, & Hinton, 1986; Smolensky, 1988). A connectionist model consists of a network of units (i.e., nodes) with connections between them (Figure 2.1). One layer of units is stimulated by inputs, and the activations of another layer is specified (i.e., constrained) by the rule-like behavior the network is to simulate; between these layers are so-called hidden units, the activation of which is not observable. In order to get such a network to respond in a rule-based manner, one presents the input layer with a series of stimuli exemplifying some sort of concept, and forces it to make a choice; the output layer then provides feedback on the adequacy of the choice. (This is the so-called backward propagation...
algorithm.) Such networks gradually learn the ability to correctly identify all instances of the concept presented to its input layer; they behave in much the same fashion as do young children (Case, 1996b) and are good at pattern recognition (Bechtel & Abrahamsen, 1991). A number of learning algorithms that are more efficient than backward propagation have been tested (for details, see Hinton, Daylan, Frey, & Neal, 1995; Shultz, 1991; Shultz & Schmidt, 1991). An important point is that no definite meaning (or knowledge) is associated with the units. Instead, knowledge is represented by the configuration of the network (units and connections); the same network can represent different concepts by rearranging the unit activation levels (via the connection strengths). Connectionist models can be used at multiple levels of analysis. For example, the simulations of explanative coherence (Thagard, 1992; Thagard & Verbeurgt, 1998) use a connectionist model in which the inputs represent propositions (or concepts) to be consolidated by the network. An explanation of connectionism that uses a frisbee and rubber band model, has been given by Bereiter (1991).

![Figure 2.1: Example of a connectionist network](image)

Dark circles represent activated, and open circles unactivated units; darker connections have higher connection weights. In this example, the input units are stimulated externally; the output units are constrained to be activated in a particular pattern. The hidden units are invisible, but crucial to producing the output pattern.

2.3.5 Situated learning

Vygotsky made social interaction a central issue in his theory of child development. While Piaget saw language as developing from experience and as an example of a representational system, Vygotsky saw language as developing
independently of practical activity and as a social tool: "The most significant moment of intellectual development ... occurs when speech and practical activity, two previously completely independent lines of development, converge" (1978, p. 24). New knowledge is first learned via social interaction, and then internalized. A key concept in Vygotsky's theory is the Zone of Proximal Development (ZPD). This is the difference between the current knowledge of the child and what the child can learn with some assistance. The cognitive functions within the ZPD are already present in an embryonic state (p. 86), so that the support from a teacher or more knowledgeable person needs to be sustained only temporarily. Schooling is key in the child's development because of its socializing function.

Vygotsky's theory (like Piaget's) has undergone substantial modification. As Case (1996a) has reviewed, current thought asserts that: (a) notational systems such as diagrams are vital intellectual resources, in addition to language (Cole, 1991; Hutchins, 1991, 1995; Olson, 1994); (b) physical and intellectual tools have specific, rather than general, effects on intellectual development; (c) informal learning (e.g., in recreational settings rather than in school) is important to intellectual development (Lave & Wenger, 1991; Rogoff, 1990); and (d) intelligence is distributed across groups (Brown & Duguid, 1991). Together, these perspectives stress that learning is situational. A physicist who normally uses software to compute integrals, will be in a lesser position to solve electromagnetics problems when such tools are not available (e.g., when not near a computer). In group situations, only some group members may have particular skills, so that the group is less functional when that person is absent. Finally, different discourse modalities (text, equations, diagrams) vary in their effectiveness to convey meaning (Lemke, 1998).

2.3.6 Summary

What are the implications of the foregoing for guiding progressive problem solving? The validation of knowledge in science and the Piagetian processes of assimilation and accommodation are psychologically similar processes. They both: (a) involve inductive hypothesis testing; (b) can be undermined by a lack of diversity of viewpoints, or too much diversity; and (c) can be undermined by social practices. As we have seen, M capacity can constrain the complexity of tasks that individuals can complete, but one can lower the M demand of a task if it is completed by means of joint activity in a group and/or the use of tools. For example, data entry into a
computer program has a smaller \( M \) demand if a person is available for reading the data items to the person typing them. If two people can both see information on a computer screen, neither person has to keep the information in short-term memory. The computer screen and the information it presents can help both persons focus their attention on the cognitive task to be completed (see Roth, McGinn, Wosyczyna, & Boutonne, in press). In addition, members of a group differ on experience and conceptual knowledge. Modeled as a connectionist network, the group with tools has a much larger configuration space (i.e., the number of distinct output patterns is larger for a given set of inputs), and can therefore be expected to be better suited to problem solving. Thus, there is psychological support for the idea that students should work in small groups, in addition to the philosophical support I have already discussed in section 2.2 (e.g., Dunbar, 1995). However, the effectiveness of this strategy can be undermined if the group does not take full advantage of the cognitive affordances of the group structure and the materials used. Large groups (e.g., a multi-cultural class) may be ineffective if the diversity of ideas is so great that it is difficult to consolidate the ideas into a coherent framework. Therefore, one might aim to have small groups consolidate their knowledge, stating about which ideas group members agreed, and which they were unable to resolve, and then to use this as the basis for a whole-class discussion (Goldberg & Bendall, 1995; Woodruff & Meyer, 1997). A goal of a model for curriculum development must be to make participant structures that are aimed at consolidating knowledge at these two levels explicit in the structure of the curriculum unit, and to promote social practices that make the best use of the cognitive affordances of these participant structures.

2.4 Constructivism and Conceptual Change

In this section, I selectively examine the literature on conceptual change in light of the previous discussion of philosophy of science and psychology, with an aim to explicate additional features of a model for curriculum development.
2.4.1 Initial knowledge and cognitive conflict

Ausubel's (1968) theory of "meaningful learning" stressed the importance of students' initial knowledge. Rather than the state of cognitive development of the learner, prior knowledge was assumed to constrain what a learner is capable of learning on a particular occasion. In the subsequent decades, students' initial knowledge in science has been studied extensively in a variety of content areas (e.g., Carmichael et al., 1990; Driver & Easley, 1978; Halloun & Hestenes, 1985; Hestenes, Wells, & Swackhamer, 1992; McCloskey, 1983; Mestre, 1991; Pfundt & Duit, 1994). Ausubel referred to initial ideas as 'preconceptions', but most researchers have stressed that initial ideas are incorrect and use the term 'misconception'. Driver and Easley (1978) concluded that students' ideas coexist with scientific ideas; they referred to an 'alternative framework':

In learning about the physical world, alternative interpretations seem to be the product of efforts of pupils to explain events and abstract commonalities they see between them. These may well be in keeping with their experience although they may be recognized as partial explanations of limited scope. (Driver & Easley, 1978, p. 62)

Throughout this dissertation, I will use the term 'commonsense explanation' to denote explanations that suggest knowledge arising from everyday experience that falls outside that validated by scientific communities (i.e., what Bereiter and Scardamalia [1993] call 'formal knowledge' and Roth [1998] 'canonical knowledge').

There is in the conceptual change literature a widespread use of terms that suggest that "misconceptions" should be "eradicated," "abandoned," "overcome," "relinquished" or "replaced" after instruction (e.g., Clement, 1982; Gunstone, Gray, & Searle, 1992; Hestenes, Wells, & Swackhamer, 1992; Strike & Posner, 1992). For instance, Mestre, in a review of physics education in high schools writes:

Unfortunately, because students have spent considerable mental effort constructing their "theories," and because these theories do explain and predict some subset of physical phenomena, students do not relinquish their misconceptions easily. (1991, p. 57)

---

2 I agree with Hewson (1996) that the term 'misconception' may still be used to refer to incorrect conceptions that have resulted from misinformation, for example, when a teacher says that red light has a greater (rather than smaller) frequency than blue light.
The following list is intended to illustrate a range of educational innovations, from physics education, that stress cognitive conflict aimed at replacing initial knowledge:

1. **Microcomputer-based laboratory (MBL) tools.** (Laws, 1991; Mokros & Tinker, 1987; Pfister & Laws, 1995; Thornton & Sokoloff, 1990) These allow the exploration of initial knowledge and the application of newly learned concepts by reducing the labor-intensiveness of traditional laboratory work. They also allow visual real-time representations that can illustrate hard-to-teach concepts (e.g., that the acceleration is not necessarily zero when the velocity is).

2. **Predict-Observe-Explain (POE).** (Champagne, Klopfer, & Anderson, 1980; Gunstone & White, 1981) These activities attempt to take fuller advantage of demonstrations and help students to focus on concepts. One approach that combines the POE with the use of MBL in university physics lectures is the Interactive Lecture Demonstration (ILD; Sokoloff & Thornton, 1997).

3. **Conceptual exercises.** A variety of multiple choice test items have been developed (Halloun & Hestenes, 1985; Hestenes, Wells, & Swackhamer, 1992; Hestenes & Wells, 1992) which are sometimes used as part of lectures (Mazur, 1997).

Such approaches have had substantial success, but conceptual change can be unstable (Gauld, 1986; White & Gunstone, 1989), students may not “see” a conflict or they are inexperienced in dealing with anomaly (Chinn & Brewer, 1993), or view discrepant events as exceptions rather than as reasons for conceptual change (Williams, 1989). White and Gunstone have described how Gauld performed a demonstration concerning the behavior of the current in simple electric circuits, after which most students performed well on a post-test:

... At this point Gauld might reasonably have concluded that the demonstration had brought about a revision of students’ beliefs, but ... he took a further step. Some three months after the demonstration he returned to the students and asked them again about the current in the wires. Most had reverted to their original beliefs, and when he asked those who held [a particular] opinion why they thought that, many referred to the demonstration. ... What had happened was that instead of reconstructing their beliefs in accord with their observations, they had reconstructed their memories of the observations in accord with their beliefs. Beliefs conquer memories. (p. 580)

---

3 Many constructivist approaches use a learning cycle. Students are first provided with opportunities to explore and become more aware of their initial knowledge; subsequently, cognitive conflict is induced and resolved by the introduction of a relevant scientific concept; and finally students are given opportunities to strengthen their understanding of the scientific concept by application to problem solving or further laboratory work (e.g., Clevenstine, 1987; Laws, 1991; Nussbaum & Novick, 1982; Renner, Stafford, Lawson, McKinnon, Friott, & Kellog, 1976; Renner & Stafford, 1979; Thornton & Sokoloff, 1990). *PSSC Physics* (Haber-Schaim & Dodge, 1986) and *Harvard Project Physics* (Rutherford, Holton, & Watson, 1981) are also in this tradition.
This is not to play down the value of cognitive conflict strategies to conceptual change, but an indication that sustained effort is likely to be required, in which the conceptual issues are addressed in diverse contexts. Chan, Burtis, and Bereiter (1997) concluded from a study of conceptual change in biology that although cognitive conflict strategies by themselves can be ineffective, they can be rendered more effective if followed by sustained knowledge building. Gunstone, Gray and Searle (1992) concluded that working with a POE strategy resulted in increased metacognition, and that this, of itself, was a valuable outcome of this strategy.

There is a more serious problem with the idea of replacing existing knowledge in that it is a strategy committed to a view of the mind as a container of knowledge structures (Bereiter, in preparation). Proponents of connectionist models would argue that what is in people's heads is not the content of science, but the capacity for producing scientific explanations from inputs, and proponents of socio-cultural perspectives that knowledge is co-constructed by people in a joint activity that makes use of mediational tools (Roth & Duit, in press). From these perspectives, conceptual change can be interpreted as an increased ability to produce scientifically valid explanations in such activities (provided that their aim is related to science).

2.4.2 Epistemologies

A second strand of the conceptual change literature deals more specifically with replacing commonsense epistemologies with a scientific one (van Heuvelen, 1991). Gil-Perez and Carrascosa (1994) and Halloun & Hestenes (1985) have argued that children should use an inductive epistemology to test scientific claims against their own. David Hammer has extended this approach by making epistemology core content of a first-year university physics course at Tufts University. Although students still study traditional content such as Newton's laws, the goal of learning content is subordinated to that of learning to use scientific epistemology. Students are required to commit themselves to a theoretical framework, which may be outside the "canons" (Roth, 1998) of science, and to attempt to use it consistently to make sense of new information. Such approaches to university science education are important because they do greater justice to the nature of science than courses that are defined by the canonical content of science alone.

---

The epistemological status of physics students' explanations has been studied extensively, particularly by McCloskey and diSessa. For example, McCloskey has noted that commonsense explanations by students often mirror historically significant theories, such as the medieval impetus theory. From a study in which he analyzed explanations offered in interviews by undergraduate physics students at Johns Hopkins University, he concluded that students used medieval theories. DiSessa (1993), however, concluded that impetus explanations by students do not constitute "a systematic and coherent theory" (p. 197). He argued that commonsense explanations are constituted by a loose collection of phenomenological primitives (p-prims). An example is the force as mover p-prim: an object moves in the direction in which it is pushed. DiSessa argues that this claim is so basic that it needs no explanation — almost anyone will accept it. But although it does explain when an object is pushed from rest, it does not explain circular motion and the observation that gyroscopes do not turn in the direction they are twisted. DiSessa argues that:

One starts with a very shallow explanatory system: many p-prims essentially serving as primitive explanations for various phenomena. Gradually, p-prims cluster and become organized as distributed encodings. ... The fundamental change in structure is that, instead of a very broad and shallow explanatory system, whatever p-prims are still used defer to or become part of the complex but few subsystems that are the encoding of the physical laws themselves. Instead of a slew of p-prims, only physical laws are explanatorily primitive at the highest levels of reliability. (pp. 142-143).

He gave a partial explanation of the mechanism by which the activation of p-prims is changed in terms of connectionist models (p. 180), devising a network of schemes representing the structural (rather than the functional) aspect of p-prims, which he called s-p-prims. The external activation of one of these s-p-prims is governed by prior activations as well as the activations of other s-p-prims through time-dependent weights.

From a pedagogical perspective, an advantage of p-prims is that they small (weakly associated) collections of them can be taken as theories to be scrutinized and improved as World 3 knowledge objects. Students activities are than focused on epistemological issues — developing more sophisticated ways of improving knowledge — rather than on the content of the claims. For explanations to be effective as knowledge objects to be improved, it would seem to be important that they can be brought into contact with explanations that reveal more elaborate
clusterings of the p-prims. James Minstrell has pioneered classroom discourse in which students have opportunities to construct scientific explanations along these lines (Minstrell 1992; Hunt & Minstrell, 1994).

2.4.3 Interpreting students' utterances

Klaassen has pointed out a semantic problem in his dissertation that arises in discourse between teachers and students (1995; Klaassen & Lijnse, 1996): they often do not mean the same to the participants in the discourse. He observed that, in his opinion, this has also led to frequent misinterpretation by researchers in the conceptual change literature (e.g., Chi, 1992; Clement, 1982; Gunstone & Watts, 1985; Halloun & Hestenes, 1985; McCloskey, 1983). Drawing extensively from work by Davidson (1984), he asserted that

[A] problem of interpretation arises if we do not take identity of meaning for granted, if we do not assume that a speaker uses his [sic] words as we do, if we do not want our interpretation to simply consist in the statement that a student says something that a scientist would not say. In short, the problem of interpretation necessarily and unavoidably emerges in the context of positive interpretation. So the problem is not only relevant to philosophers, like Davidson, but also to science educators. For what are pupils trying to say when in their talk they use words like 'force', 'energy' and the like? (p. 10)

He answers this question, in part, as follows:

Children and lay people generally do not use the expression 'to exert a force' when there are no agents involved. It makes no sense for them to say, e.g., "The ball exerted a force on the window," simply because the ball cannot be the agent of an action. ... Nevertheless, its motion can, just like our actions, cause something to happen to other objects. To be able to express the latter is, I think, the point of the expression 'to have a force' in relation to moving objects. So I take children's and lay people's holding true the sentence 'If an object is in motion, then there is a force in the direction of its motion' as an expression of the belief that the motion of an object can cause something to happen to other objects that are located in the direction of its motion. (1995, p. 14)

I can remember an argument with an 11th-grade student who insisted that the wall did not exert a force on a person leaning against it. After reading Klaassen's dissertation, I understood why I could not have won that argument. That students do not tend to attribute agency to inanimate objects was also discussed by diSessa (1993) in connection with p-prims.
Klaassen's point is important for pedagogy. In the classroom discourse he analyzed, students made statements that made sense in their framework, and that cohered with each other. For instance, they asserted consistently that inanimate objects do not exert forces on objects that are supported by them. Klaassen and Lijnse (1996) reject the use of conflictual situations in such cases:

The idea behind the use of conflict situations is to confront students with a discrepant event that will more or less force them to abandon, for instance, the static objects are barriers that cannot exert forces conception. ... On our analysis this conception represents the belief that static objects are barriers that cannot of themselves cause something to happen, and there is no reason to make students abandon this belief. (p. 126)

By the same token, they argue, analogical arguments in which a book is first supported by a spring, and then by progressively less "springy" materials (Clement, 1993; Minstrell, 1982) will not convince students that the table exerts a force on the book, although they are likely to agree with the teacher that the situations are similar (i.e., in each case, something is deformed). These authors relate this situation to the incommensurability of Kuhnian paradigms, and conclude that the teacher and student could have learned each other's conceptions — they could have translated for each other rather than attempted to convince the other of their own conception. But they go on to argue that it is not sufficient that the teacher and student learn to communicate with each other: there are good reasons for using the scientific notion of force, and it is a pedagogical challenge to help students see this. As in the example I discussed in the previous subsection, it is a matter of extending the student's conception — it must become applicable to a larger range of situations, and hence explain more. It would appear, then, that Klaassen's analysis, which is based on Davidson's semiotic theory (see Davidson, 1984), and diSessa's p-prim theory are congruent with each other.

In terms of the terminology I have been using, a pedagogical perspective consistent with Klaassen and Lijnse's position would consist of two parts: (a) the teacher and students learn to use each other's discourse practices, and (b) they attempt to alter the discourse practices in the direction of greatest explanatory coherence and parsimony. In my opinion, there are several advantages to such an approach.
It gives recognition to the validity of students' knowledge (within their own discourse practices), and focuses activities on theory improvement rather than replacement.

- It gives students practice in interpreting phenomena in multiple perspectives, an activity crucial to knowledge building (Bereiter & Scardamalia, 1993).
- It stresses that learning science involves learning particular discourse practices, for particular reasons — this suggests a role for direct, bi-directional, teaching (i.e., the students teach the teacher what things look like in their perspective).

2.4.4 Macroscopic and microscopic explanation

In the development of a new field in science, one usually begins at a macroscopic or phenomenological level. For example, thermodynamics explains a wide range of phenomena involving liquids, solids, gases and heat, using such variables as temperature, pressure and volume. Heat transport through a wall may be predicted on the basis of knowing the temperature on both sides of the wall and measured characteristics of the wall, such as its R-value. Another example is the exponential decrease of the activity of a radioactive substance, which is an adequate level of explanation for many medical applications of radioactivity. Microscopic theories go beyond this level of analysis to one in which one looks at the object of interest (e.g., a material) as a collection of particles, possibly with inter-particle interactions or interactions with external influences such as light. An example is the kinetic theory of gases. Microscopic theories can explain why radioactivity decays exponentially and can provide an interpretation of the idea of temperature.

Which level of analysis is appropriate (macroscopic or microscopic) depends on the purpose of the inquiry. As Tinkham (1985) writes in his textbook on superconductivity:

The BCS microscopic theory ... gives an excellent account of the data in those cases to which it is applicable, namely, those in which the energy gap \( \Delta \) is constant in space. However, there are many situations in which the entire interest derives from the existence of spatial inhomogeneity. ... In such situations, the fully microscopic theory becomes very difficult, and much reliance is placed on the more macroscopic Ginzburg-Landau theory. (p. 104)

Thus, although microscopic theories could not be developed without an extensive knowledge of the phenomena obtained from observation and theorizing at the phenomenological level, it is not true that microscopic theories supersede them.
The implication of this for science education is that inquiry at both levels should be encouraged (cf. Mortimer, 1995; NAEP, 1996). Unfortunately, there has been a tendency in science curricula to stress microscopic models, at the expense of macroscopic models. The problem with stressing microscopic models is not so much that students cannot successfully make sense of such models, but that they may miss opportunities to learn to build simple models (an important aspect of scientific inquiry), and opportunities to build more extensive phenomenological knowledge. It can probably not be over-emphasized that microscopic theories are not based on a few insights at the phenomenological level — they require a large amount of background knowledge. Thus, while some elementary- and middle school students are capable of understanding microscopic theories, emphasis on microscopic theories may be at the expense of knowledge of what such theories can explain.

2.4.5 Strike and Posner's Conceptual Change Theory

Another line of research has been to explicate the conditions under which students will be persuaded to commit themselves to scientific concepts and methods. In a Science Council of Canada (1984) report, Nadeau and Désautels concluded that “teenagers will not change their unconscious epistemological premises ... merely on the strength of what a teacher says or when confronted with a few minor contradictions” (p. 54) because they work for them in everyday life, and suggested four teaching principles. One of these stressed that students should develop insight into their own learning processes. A major Australian effort, the Project to Enhance Effective Learning (PEEL; Baird, 1986; Baird, Fensham, Gunstone, & White, 1991; Baird & Mitchell, 1986; Baird & Northfield, 1992) has developed such metacognitive strategies through action research.

Strike and Posner (1982) proposed a theory of conceptual change grounded in Kuhn's paradigm shift. Their theory was normative and epistemological; it was not psychological or cognitive. It was not concerned with learning but with the question, 'What does it take for people to change their commitment from one conceptual framework to another?' The theory was not empirically grounded in the initial knowledge research but could be applied to explain it. It gave four conditions that must be met for conceptual change to occur (Strike & Posner, 1992, p. 149):

- Existing ideas must be found to be unsatisfactory;
- The new idea must be intelligible;
• The new idea must initially be plausible;
• And it must be fruitful in further learning.

"Intelligible" implies that the new idea appears to be both coherent and internally consistent (Duschl, Hamilton, & Grandy, 1992). An important feature of the revised theory became the conceptual ecology, consisting of artifacts as anomalies, analogies, metaphors, epistemological beliefs, metaphysical beliefs, knowledge from other areas of inquiry, and knowledge from competing conceptions. Citing Eylon and Linn (1988), Strike and Posner noted that

what may be most important in instruction is understanding the factors in learners' conceptual ecologies that generate and maintain misconceptions. It is these factors that may account for both the durability of misconceptions and their commonalities across learners and even sometimes cultures. (p. 158)

Criticisms of the revised theory fell into three areas. First, it was not clear that the initial knowledge of students can be characterized as paradigmatic, that is, that the kind of shift required is a paradigm shift rather than the onset of the first paradigm: Knowledge might be iconic or representational (see Nercessian, 1992). Second, a conceptual framework is not clearly separable from the conceptual ecology: what is seen as problematic depends on what you look with. Third, the conceptual ecology was overly rational and must also include such things as interest and emotion (Pintrich, Marx, & Boyle, 1993).

2.4.6 Conceptual change across ontological categories

Chi (1992) has proposed a theory of conceptual change that distinguishes across conceptual change within ontological categories and "radical conceptual change" between them. Her theory makes two assertions: (a) conceptual change within ontological categories requires a different set of processes than conceptual change across ontological categories, and (b) learning physics is difficult because it requires conceptual change across ontological categories. Central to Chi's theory is a distinction (Sommers, 1963) made between propositions that could be true but are false and propositions that could not be imagined true except in a metaphorical sense (Keil, 1979). Thus, the proposition 'The cow was green' is probably false, but one can imagine painting a cow green. On the other hand, one cannot imagine painting an event green: events cannot be painted. Sommers called propositions like 'The flight was green' "anomalies." Chi argued that learning science involves such
anomalies because novices treat concepts such as force and current as material-based entities, rather than as "constrain-based events;" they also give animate attributes to inanimate objects. Evidence motivating her theory was: (a) limited success with instructional strategies designed to facilitate conceptual change (see section 2.3.2), and (b) striking similarity between ideas used by medieval scientists and by children (McCloskey, 1983). This evidence, Chi concluded from a review of the literature on concept acquisition (see Carmichael et al., 1990; Pfundt & Duit, 1994), was consistent across (a) studies, (b) concepts, (c) ages, (d) educational levels, and (e) historical periods. She inferred that medieval scientists' conceptions were material-based, and that radical conceptual change is at the core of scientific revolutions.

Chi asserted that 'force' is a constrain based event, and that students often interpret it as a material-based entity, and hence must accomplish an ontological shift — radical conceptual change — to learn the scientific force concept. Klaassen (1995) rejected this interpretation, and suggested that the problem of interpretation lies in the fact that the word 'force' functions grammatically as a noun. According to him, "students do not assign a meaning to the word 'force' in isolation at all, and, in particular, do not use it to refer to any entity whatsoever" (p. 16). A similar situation exists when statements like 'The molecules try to avoid the heat' are interpreted as saying that children attribute mental states to molecules. Talk like this is almost unavoidable and used by scientists, as in the utterance, 'An atom likes to be in the ground state'.

In the history of science there are ontological anomalies. The proposition, 'The electron split up, and one half traveled through each slit', is nonsense if applied to a classical particle but an exemplar of a quantum mechanical effect. Let us examine some of the steps involved for scientists for conceptual change from classical to quantum mechanics in more detail than I did in section 2.2.3. First, there were numerous anomalies, and early in the revolutionary science period, progress was greatly facilitated by mathematical and physical intuition and the use of analogy.

- Planck's quantum was an *ad hoc* assumption that resolved a mathematical singularity, the origin of which — the summation of a geometric series — was well

---

5In Chi's view, an electric current is a constraint-based event because it exists only when a condition is met, that is, when a potential difference exists between two points in a circuit (and the circuit is closed).
known to nineteenth century physicists (e.g., it is used in problems in classical electromagnetic theory).\textsuperscript{6}

- Perrin’s planetary model of the atom was in obvious close analogy with the solar system. The work of Bohr and Sommerfeld to develop this model was greatly facilitated by classical analytical formalism used in orbital mechanics (e.g., Hamilton-Jacobi theory), also well known to (theoretical) physicists of the period.

- Wave optics, developed, in part, by Fraunhofer and Fresnel in the early part of the nineteenth century, was well known at the time, as was its relationship to geometric optics.

The conceptual change to the so-called “old quantum theory” depended crucially on expert knowledge of classical methods. It had important successes, such as an explanation of the Balmer formula for the hydrogen spectrum and of the photo-electric effect, but also some failures. For instance, it could not deal with scattering problems and the spectra of large atoms. The ad hoc assumptions also remained unexplained. Beginning in 1925, quantum mechanics was developed in a remarkably short period, and it was applied to many problems with great success. Perhaps one of its greatest successes was the theory development in nuclear physics in the 1930s. However, its interpretation remained problematic for many physicists and philosophers, despite the interpretation of the wave function given by Born. The theory predicted successfully some effects that did not have counterparts in classical mechanics. And the matrix techniques used by Pauli in developing the matrix formulation of quantum mechanics were unfamiliar to most physicists. An important aspect of this period in the history is the role an intermediate theory played. Roschelle’s (1996) ideas of convergent conceptual change seems applicable here.

For students who begin to learn quantum mechanics, the situation is quite different: they have less complete knowledge of anomalies and classical methods, but more knowledge of matrix methods, and the benefit of access to many decades of thought about the ontological status of the theory.

\textsuperscript{6}Thagard refers to such an \textit{ad hoc} assumption as \textit{abduction}, which is between induction and deduction. One observes that $X$ explains $p$, and therefore assumes $X$. 
2.4.7 Summary

As at the end of sections 2.2 and 2.3, I summarize the implications of the preceding discussion for curriculum development in the context of a knowledge building epistemology.

1. Misconceptions versus p-prims: I have argued against the expectation that students will replace their current ideas by scientific ones as a result of discrepant ideas. From the point of view of knowledge building, small, weakly associated p-prims, interpreted as theories that can be developed in World 3, seems a better starting point. This position acknowledges that students' present knowledge is often valid, and should be valued as such. A curriculum development model must address the issue of helping students to get started with theory improvement, for instance, by providing ways by which students can become aware of the content of their theories, their strengths, and limitations.

2. The problem of language: We have seen that the meaning individuals give to statements can cause difficulty in classroom discourse aimed at conceptual change. It is my contention that this problem is primarily semantic rather than ontological. Before the value of scientific knowledge over commonsense knowledge can be addressed, students and the teacher must learn to communicate with each other, which requires that they learn what terms like 'force' and 'energy' represent in their respective explanations. This can be accomplished by bi-directional teaching early in the unit.

3. Empirical content of theories: In section 2.4.4, I argued that scientists use theories at varying levels of abstraction, depending on the purpose of the immediate analysis; in section 2.4.6, I illustrated that scientific theories are developed from a rich problem space (questions, empirical knowledge, methods, established theory). This suggests to me that an effort should be made to ensure that students theorize from an adequate problem space, and that they learn to develop a variety of types of theories — progressive discourse should enable students to develop more elaborate theories on the one hand, and more extensive knowledge of natural phenomena (and technology) on the other.
2.5 Conclusion: A Curriculum Development Model Aimed at Knowledge Building

2.5.1 Description of the model

In this section, I present a preliminary version of the Learning to Knowledge Building (LKB) model; it will be evaluated (and modified) in chapter 6, in light of studies 1 and 2. The model is a curriculum development model; it is not a cognitive model aimed at explaining learning processes or conceptual change. Nor is it a type of curriculum model that provides a standard against which practice is to be evaluated. Rather, the model conveys what a unit based on knowledge building epistemology might look like — early in the unit, and later. My concern is with the kinds of activities a teacher might consider in planning a unit in which students have opportunities to engage in "authentic science" (Bencze, 1996; Hodson & Bencze, 1998; Roth, 1995). Some components of the model are based on findings from social and socio-cultural studies of science (Dunbar, 1995; Kuhn, 1970; Lakatos, 1970; Latour, 1987), but the model does not yet address explicitly how to make them work. The design parameters/constraints I used for constructing the model are:

- A progressive research program (Lakatos, 1970) within the unit.
- Inductive logic as a basis for experiment design, where appropriate, within such a research program (Popper, 1959/68).
- Recognition of the validity of students' current knowledge (Klaassen, 1995; Klaassen & Lijnse, 1996), but encouragement to improve theories in World 3 (Popper, 1972) toward greater parsimony and explanatory coherence (Thagard, 1992).
- Small-group and large-group discourse (Dunbar, 1995; Woodruff & Meyer, 1997) aimed at consensus about problems, methods, and results of research (Kuhn, 1970).
- Sensitivity to the limitations of discourse reflected by the learning paradox (Pascual-Leone, 1976).
- Emphasis on a diversity of perspectives.
- A distinction between learning and knowledge building (Bereiter & Scardamalia, 1996a).
- Emphasis on an empirically and conceptually rich context for theory improvement, and on a balance between depth-of-understanding and breadth-of-coverage.

The last two items in this list suggest a role for direct, but cognitively engaging, teaching. The distinction Bereiter and Scardamalia (1996a) make between learning and knowledge building is as follows:

We have encountered teachers who would never stand in front of a class and explain the difference between restrictive and nonrestrictive clauses and who would be appalled at the thought of teaching young children how to sound out words; yet, when introducing a new software application to their students, they do not hesitate to deliver direct
lessons .... Somehow, teaching software use falls outside their strict constructivist philosophies. It is just something the students need to learn in order to get on with their constructivist work.

With a better-elaborated epistemology, such teachers might find that a great deal more of schooling fell into the category of "just something the students need to learn in order to get on with their constructivist work" — their constructivist work being knowledge building. The so-called "basic skill" components of reading, writing, and arithmetic might well fall into this category. So might various kinds of factual knowledge that are presupposed in reading materials students will use .... (p. 501)

Not all aspects of the model, as it is presented here, should be expected to be enduring. The model includes structural features (particularly the "phases") that are helpful for the process of developing the desired understanding of the structure of a unit aimed at progressive discourse, but they may become unnecessary at some point in this journey.

The model is shown in Figure 2.2. The arrows indicate 'results in/contributes to ...', so that conjectures in Phase 1 may result in questions, but they may also contribute directly to the preliminary research program. The model does not assume that commonsense knowledge is to be replaced by scientific knowledge, but it does assume an epistemology — the knowledge-building epistemology (Bereiter & Scardamalia, 1996b). Each of the phases of the LKB model is described.

2.5.1.1 Phase 1: Developing a knowledge base.

This phase is teacher-directed but goal-oriented. It serves two purposes:

- To ensure that the students have a sufficiently rich context for developing a research program around the topic to be studied. This includes knowledge of what others have learned about the topic before (e.g., a previous class), what resources are available, and how scientists have treated the topic. It is only familiarity with scientific ideas (not deep understanding of them) that is needed at this time.

- To ensure that student achievement has some measure of comparability with that of students in other learning situations. This may be accomplished by including locally used curriculum guidelines and benchmarks in what is considered when the research program is developed.
Figure 2.2: The Learning to Knowledge Building Model

The model reflects a knowledge building epistemology in which knowledge is made public for debate, scrutiny and improvement. Movement down from the top of the diagram indicates time progression, but the diagram is not drawn to a scale. Phase 1 stresses learning, Phase 2 the transition to knowledge building, and Phase 3 knowledge building itself.
Phase 1 has a fairly close correspondence to how scientists work. Although physicists specialize in small subfields (e.g., using field theories to explain the behavior of materials), they usually expend considerable effort to keep up with the progress of the field as a whole, and with other fields in their discipline. Many physicists read much and attend conferences and colloquia for this purpose — they develop ideas for studies from such activities. And when scientists have an idea for a study, especially if it means a new direction for them, they check what has been done on the problem before. Thus, they develop knowledge at two levels: deep knowledge in their specialty, and more superficial knowledge in other areas. It would be difficult to imagine an expert who could not situate his or her deep knowledge in a broader context. Students, too, need to develop knowledge at two levels — deep knowledge in their specialty and more superficial knowledge around that. Having only deep knowledge in a very limited area flies in the face of most notions of scientific literacy. Curriculum guidelines can play an important role in delineating the breadth of knowledge desired at a particular grade level.

How do I imagine Phase 1 to work in practice? The teacher chooses a topic to be studied from the prescribed curriculum, and designs or selects some activities that will allow students to develop a base of working knowledge relevant to this topic. (This knowledge will become more fully elaborated in subsequent phases.) The teacher, can, of course organize the curriculum in such a way that the topics for the year are embedded in a series of problems, close to students' interests. The activities should span the breadth of the prescribed curriculum for the topic in question. Examples of activities might include (no temporal sequence is implied):

- Practical work that can give students a common base of experience with the phenomena to the studied.
- Lessons that introduce scientific ideas.
- Eliciting from students what they think they already understand about the topic (e.g., by asking students to enter notes about this in a CSILE database).
- Independent reading.
- Examining what has been done on the topic before (e.g., by searching the Internet or local CSILE databases developed in previous years).
- Locating resources and checking their suitability (i.e., is the reading level appropriate?)
- A field trip to a science museum or a visit by a scientist.

Cooperative learning strategies can be used. Locating resources requires only a few students, and not all students need to do the same individual reading: students can
make a brief class presentation of what they found out, and write a summary note of it and the resulting discussion in a CSILE database. (Indeed, not all students need to do individual reading). So the class can work as a learning community from the beginning, but the teacher must organize things in such a way that what individual students learn becomes, as much as possible, known by the whole class. That is why Phase 1 is called teacher-directed. It should be clear that the activities are not ends, but serve the need of developing a fairly broad knowledge base. (Developing such skills as making a presentation to the class are incidental benefits of the activities, they do not constitute their purpose.) Also important in Phase 1 is explicit attention to knowledge building strategies. This can take the form of evaluating the knowledge the class has previously developed, and setting goals for improvement for the current unit. In a sense, this becomes a discussion on research method in the context of the class's previous research experience.

Toward the end of Phase 1, students should have many questions and ideas. Some questions are about surface features or factual information. How does heat affect matter? What is the chemical that turns sand into glass? What is the density of water? Some of these questions may be answered quickly. About others, students may have "theories" or hunches. I have called these conjectures because they tend not yet to be sufficiently well articulated to be amenable to hypothesis testing. Of course, conjectures, too, can lead to questions.

2.5.1.2 Phase 2: Transition to knowledge-building.

Phase 2 is a transition, and therefore has fuzzy boundaries. I have called it a phase to stress that explicit attention must be given by the teacher and the class to ensure the class makes a successful transition from learning to knowledge building — this does not happen automatically.

Some of the questions and ideas that have been generated by the end of Phase 1 are more promising than others for advancing the collective understanding of the class. Moreover, the scope of questions and ideas may be too broad to allow the formation of a coherent research program. The purpose of Phase 2 is to select the most promising questions and ideas from the available set, perhaps generate new questions and ideas, and gradually formulate a research program. The appearance of an early form of that research program marks the end of Phase 2.
Phase 2 might be viewed as analogous to doing pilot studies or developing a research proposal. Individual students or groups of students begin by selecting from the body of questions and ideas that have been generated a subset that they want to work on. Usually, this subset will consist of their own questions, but questions closely related to them, contributed by other students, can be included as well. Then they begin to work, clarifying their questions to the point that they suggest an experiment, reading, thinking, and discussing. Some questions might not sustain the students’ interest or inquiry, and after a period of time a small set of questions and ideas remains that the students wish to research. In a sense, the aim of Phase 2 is to ensure all students converge on research ideas that are sufficiently promising for the knowledge building phase to be successful. What students plan to do should be in their Zone of Proximal Development (Vygotsky, 1978). No one should remain stuck because their specialty turned out too difficult for them, or because they lacked motivation and chose not to participate actively in the unit. The teacher’s role is to ensure this by monitoring student activity (e.g., by reading the CSILE database and/or obtaining database use statistics). If students work in groups, there is likely to be some negotiation as part of the process of building a research agenda for the group, and the issue of power becomes important to understand (an example concerning “making glass” is given in chapter 4). Consensus should be a goal within the group — at least coming to a stage where the group can say, “We like these two approaches, but cannot decide one which one is best.”

2.5.1.3 Phase 3: Knowledge Building.

The groups have improved the set of questions they had at the end of Phase 1 in the following sense: they have selected a subset and have done enough research to refine the remaining questions into a proposal for further study — e.g., an experiment, building a model, or a study based on library research. Each group has in rough outline a research program that can advance its own understanding. At the beginning of Phase 3, public debate (face-to-face or electronic) now needs to occur, with the following aims:

- Each group must refine its research ideas. It is important to have students report ideas that seem to be working, and to facilitate criticism, as well as to have them report ideas with which they have not been able to make progress, including the reasons why they expected these ideas to work. Another group may have located the information necessary to make an idea a fruitful one, or may still come across it later in the unit.
The group research programs need to be consolidated into an overall research program for the class. It must be clear that a group is not on a track to researching something others in the class already know much about. And, if two groups are working on similar problems, an opportunity should not be missed to stimulate debate between them — e.g., two groups may test competing hypotheses. (A case of failure to do this is discussed in chapter 5.)

The improvement of research ideas may be facilitated in a number of ways. It seems important to vary what is done. Some possibilities are:

- Writing a proposal as a database entry, and having several students (each from another group) peer review it.
- Verbal presentations before the class.
- A poster session.

The middle part of Phase 3 is taken up by executing the research program. This may involve any combination of experimental work, technical design, theoretical work (e.g., based on reading), model building, and correlational studies (Bencze, 1996). Although discourse is likely to be within groups while the work is going on, drawing from the knowledge of other groups should always be considered. At the end of Phase 3, the newly developed knowledge of each group needs to be consolidated. Activities at this point can be structured in ways similar to those at the beginning of Phase 3, but it is important to end up with a product that makes clear what was learned, how it was learned, and why the class thought advances were made. This can take the form of conclusion notes, "published" notes, or whatever. Again, these products are not ends, but means to ensure that the knowledge advances are not lost and are available in some form to outsiders who may be interested in the topic subsequent to the completion of the unit.7

2.5.2 Discussion

Some additional discussion of the LKB model seems warranted at this point.

1. It takes too long to get to knowledge building: Figure 2.2 is not drawn to a time scale, although moving down in the diagram represents moving ahead in time. The figure stresses that explicit attention must be paid to a number of issues before open-ended inquiry can successfully proceed. Phase 1 involves a variety of learning strategies and is, overall, perhaps not very different from traditional practice in

---

7E.g., a future class that is to do a study of the same topic.
learning science (Hodson, 1990). Getting to the "Consolidated Research Program" need not take very long, but it is a delicate process that does not happen by itself, and must therefore be attended to by the teacher and students. The bulk of the time should be taken up by what comes after this (at least half the total time spent on the unit). I contend that each phase can be arranged in a way that will capture students' interest, and that, moreover, the model is a realistic representation of the work of scientists.

2. There is no learning about the value of science to society: The model can be used for any subject matter area for which its epistemology is appropriate. An holistic approach can be used. For example, in studying flight, one group of students might study the environmental impact of increasing commercial air traffic, while another group studies the designs of several types of aircraft, and yet another studies aspects of the history of commercial flight.

3. How is a teacher to know how far to aim in Phase 1? That depends on many factors — how much time the class can spend on the topic, the abilities of the students, the available resources, the teacher's previous experience with the unit, and so on. What the teacher exposes students to should be in their Zone of Proximal Development, but using primarily "hands-on" activities, at the expense of work of a conceptual nature, is to be avoided. Chan, Burtis, and Bereiter (1997) recommend maximizing cognitive conflict, followed by sustained knowledge building effort.

4. Knowledge building requires more teacher knowledge than traditional approaches to science education: Critics may argue that many teachers are likely to be ill-equipped to cope with progressive discourse. While this may be true, teachers can engage in progressive problem solving to improve practice. Action research (e.g., Carr & Kemmis, 1986; Elliott, 1991; Hodson, in preparation) would seem to be an appropriate way to achieve this, particularly if people who understand the subject matter and knowledge building epistemology are part of the action research team. A purpose of the Knowledge Society Network (Scardamalia & Bereiter, 1996) is to support progressive discourse of this kind.

5. Does a teacher use the model for every unit? Not necessarily. Teaching with the LKB model can be very time consuming, and with some units all that may be possible is Phase 1. However, I suggest that students should do at least one unit of
10-15 weeks in science this way each year, possibly with a second unit in another subject area such as history. (Or, for instance, two problem-based units that transcend traditional curriculum boundaries.)

6. For what learning situations is the LKB model appropriate? No definite answers are possible without testing the model, but the following remarks seem appropriate. Research on commenting indicates that students in a Grade 1/2/3 classroom that used CSILE could write comments that revealed an intent to help other students advance their understanding (Woodruff & Brett, 1993). Thus, even in these early grades aspects of knowledge building are manifest. Certainly by Grade 4, students can do collaborative research projects (see chap. 3). The LKB model could provide a framework for aspects of science courses in university, such as laboratory work or preparing a research paper, but I do not see the bulk of, say, introductory university physics, being addressed in this way. However, intentional learning is still a goal here, even if open-ended research is not (van Aalst & Marjoribanks, in preparation). Finally, as has already been implied, the model can be valuable in organizational learning (Argyris, 1994; Senge, 1990) and teacher education.
CHAPTER 3

CSILE, RESEARCH, AND CLASSROOM PRACTICE

The purpose of this chapter is to frame the studies reported in chapters 4 and 5 against the CSILE Project and prior empirical studies on classroom work with CSILE. Much of the data reviewed was obtained at a single middle class urban elementary school; for convenience, I have given the CSILE teachers at this school pseudonyms as follows: Linda (Grades 1/2/3), Ann-Marie (Grades 4/5), John (Grades 5/6), and Paul (Grades 5/6). John’s class of 1994/95 developed the database of study 1, and Paul’s class of 1996/97 that of study 2. John and Paul had taught with CSILE since the beginning of the CSILE Project in 1986. (No databases by Linda and Ann-Marie’s classes are analyzed for the dissertation.)

Section 3.1 focuses on CSILE itself. First, I review the research context from which CSILE was conceived. Why use CSILE for school science? I follow this with a description of the design features of CSILE and of several episodes in which both the software and classroom practices have undergone substantial enhancement. These episodes illustrate continual improvement of the sort that the dissertation is aimed to inform. Then I describe how students use the CSILE design features.

Section 3.2 reviews research on teaching and learning with CSILE at John and Paul’s school, including two prior dissertations (Cohen, 1995; Woodruff, 1995) that have compared discourse in face-to-face mode with that in electronic discourse with CSILE. Also important is the distinction between two models of CSILE use: the independent research model and the collaborative knowledge building model (Bereiter & Scardamalia, 1991). I will argue in the chapter that Paul has consistently used the first of these models and John the second.

Finally, section 3.3 presents and discusses data collected as part of study 2 to illustrate that the findings obtained from Paul’s classrooms in the past are consistent with his use of CSILE during the unit analyzed in study 2.
3.1 The Development of CSILE

3.1.1 CSILE and research into the writing process

In chapter 1, research into the nature of written composition by Bereiter, Scardamalia, and their co-workers was mentioned. Here, I provide some details of that research and describe the theoretical underpinnings of CSILE.

In *The psychology of written composition*, Bereiter and Scardamalia (1987a) distinguished between two models of writing: ‘knowledge telling’ and ‘knowledge transforming’. Writers who use the knowledge telling model face problems with transferring their knowledge to written form. In making revisions, they focus on local features of the text. They may rephrase a sentence to make it more cohesive, elaborate on a point, or add additional points. Writers who use the knowledge transforming model do this, but they also reflect on the text as a whole. Struggling with the rhetorical problems of making an argument effective, these writers may come to re-examine their thesis, and rethink what it is they need to say. They are concerned with the global features of the text. Thus Bereiter and Scardamalia recognized writing as a powerful vehicle for transforming knowledge. They found little evidence in protocol analyses for knowledge transforming in young writers. “Mature writers,” they write:

referred explicitly to specific language, to main ideas, structural elements, problems and goals. ... The protocols of immature writers consist almost entirely of references at a lower level, closer to the surface text. To judge from planning protocols, almost all the mental work of composing by young writers consists of operations involving discrete items of content or language. (Bereiter & Scardamalia, 1987a, p. 352)

Bereiter and Scardamalia chose to study writing because they thought that cognitive science had advanced sufficiently to inform such a study, and because they saw writing as an example of problem solving in a content-rich domain. At the time, most work on expertise had been in areas dominated by logico-mathematical reasoning (e.g., de Groot, 1965; Chi, Feltovich & Glaser, 1981; Larkin, McDermott, Simon, & Simon, 1980; see Bereiter & Scardamalia, 1993, for a review). The theory of writing they developed was a cognitive theory that should be able to inform problem solving in other knowledge-rich domains. From the early 1980s, Bereiter and Scardamalia began to investigate a number of regulating mechanisms that could be introduced into the executive procedure of composition and facilitate knowledge...
transformative writing. They called the use of these mechanisms 'procedural facilitation'. In reading the second half of the *Psychology of written composition*, one recognizes some of the design features of CSILE. For example, two strategies that became incorporated in *Thinking Types* in CSILE were sentence completion and planning. Completing sentences beginning with phrases such as 'We have learned that...', Bereiter and Scardamalia concluded, "served to keep students to the distinctive character of reflective writing and to keep them from slipping into the more familiar genre of argument or personal narrative writing" (1987a, p. 262). The idea of *intentional learning* also developed from the knowledge transforming model and procedural facilitation — Bereiter and Scardamalia saw writing in the knowledge transforming model as a special case of intentional learning.

The original aim of CSILE was to bring the knowledge transforming model of writing to life for young writers. It does this in several ways. First, writing into a database provides a more permanent record of students' ideas than class discussion or other forms of writing do. Students can, having read something, or having thought more about the subject, re-read a note written several weeks earlier and improve on it. In addition, a note in a database is not read with the author of the note co-present, as is the case with face-to-face discourse. That means that students must convey in their notes more of the context within which the note is to be understood than would be necessary in face-to-face discourse — electronic discourse introduces a need for solving complex rhetorical problems. Such problems include anticipating and responding to alternate viewpoints; for this reason they tend to involve high \( M \) demand. A psychological assumption underpinning writing with CSILE is that writing in a public space will help students cope with the high \( M \) demand to some extent: one may expect a group using a technology to record their ideas to have greater information processing capacity than an individual. Thus, the design aims of CSILE were originally primarily socio-cognitive; as the CSILE project evolved, *philosophical* reasons for objectifying knowledge in a communal space have received more emphasis in the Knowledge Society Network (Scardamalia & Bereiter, 1996).

### 3.1.2 Description of CSILE

CSILE is a networked, computer-based database system (Scardamalia et al. 1989, 1992; Scardamalia & Bereiter, 1993; Scardamalia, Bereiter, & Lamon, 1994) The
installation at John and Paul’s school consisted of a server and eight client stations in each of the two classrooms, resulting in a student-to-computer ratio of approximately four to one. Installations, however, vary considerably from school to school — some schools now have better student-to-computer ratios. Classroom activities are organized with the aim of giving each student *individual access* to a networked computer for 30 minutes a day. As Scardamalia, Bereiter, and Lamon (1994) noted, “the physical conditions — especially the fact that not all children can be working on CSILE at the same time — militate against the traditional schoolwork model, where all the students are doing the same thing at the same time” (pp. 208-209).

3.1.2.1 Design features.

There have been several versions of CSILE. I describe version 1.5 for Macintosh computers, which is the version used by John and Paul for the units I have studied. The features of CSILE 1.5 that support intentional learning are as follows.

- **Note type** (discussion, text, graphic). A *discussion note* is started by stating a problem; contributions added to the discussion of this problem are presented together with the problem statement, in chronological order; subdiscussion notes are also possible. In contrast, a *text note* is presented as a single composition, and can be used for commenting on other notes. A *graphic note* is similar to a text note in this respect, but consists of a graphic (and whatever text has been added to it using a text tool). Text notes can be co-authored. Figure 3.1 shows a discussion note.

- **Note status** (normal, draft, published). Notes of all types can be edited so that students can work on a note over an extended period of time. In addition, a note may be raised to a higher status, first from ‘normal’ to ‘draft for publication’, and then to ‘published note’. This feature is used by some teachers toward the end of a unit.

- **Thinking types**. These help students focus on the *goal* of their note in the discourse; they are also helpful for reading notes — i.e., understanding what the author of a note intended. The thinking types most commonly used are (others can be defined by the teacher): ‘My Theory’, ‘I Need to Understand’, ‘New Information’, ‘Comment’, ‘Plan’, and ‘What We Have Learned’. The student must declare a thinking type before a note can be saved.

- **Searching and linking**. This facilitates reflection, which Bereiter and Scardamalia (1987a), following Piaget, have taken to be “a dialectical process by which higher order knowledge is created through the effort to reconcile lower-order elements of knowledge” (p. 300). The database can be searched in a variety of ways (author, key word, topic, note type). Selected notes can be joined by a link; before the formation of a link can be completed, a justification for the link must be provided. This supports reflection as well.
Figure 3.1: Example of a discussion note in CSILE, version 1.5
Note entries are presented below the problem statement to preserve a sense of the discussion. This particular example is a subdiscussion.

3.1.2.2 Enhancements to CSILE.

The interface of a learning technology is a key factor in the effectiveness of that technology. Here, the development of several enhancements of early versions of CSILE is reviewed. These developments provide an example in which a teacher and researcher jointly used an electronic trace of database use for inquiry aimed at continual improvement.
Early versions of CSILE did not have *discussion notes* because text notes were thought to make for a more flexible knowledge medium, provided that they were supported by facilities to make links between notes and for searching the database in multiple ways. But in practice, “the overhead involved in moving from one note to another discouraged students from following, and extending, lines of thought” (Hewitt, Webb, & Rowley, 1994, p. 4). A 1992 pilot study by James Hewitt revealed that discussion notes contained a significantly higher number of epistemological terms, such as terms that make reference to a hypothesis, evidence, and opinion change (cited in Scardamalia, Bereiter, & Lamon, 1994). Hewitt and Webb (1993) concluded from a study in which *discussion notes* were compared with *text notes with links*, that discussion notes fostered greater collaboration and deeper explanations. However, they also observed that discussion notes provided a bias toward deep explanation. Contributions tended to extend *the current line of thought*, at the expense of raising new issues for discussion.

The *subdiscussion* feature of discussion notes, which made threaded discussions possible, was designed to address this problem. Hewitt, Webb, and Rowley (1994) examined the impact of this branching feature on the way students organized their discussion, and how the different interface styles affected attempts to address ‘I Need To Understand’ (INTU) queries. They compared a unit on human biology done by John’s classes of 1992 (prior to the introduction of branching) with one done in 1993 (subsequent to it). Students worked in groups, but had individual access to a computer. Hewitt, Webb, and Rowley concluded that at $p < .01$ branching did not affect the amount of question asking or reduce the frequency of unaddressed questions, but it did affect the following up of INTU entries. In 1993, the average number of follow-up entries was 2.23, compared with 1.13 in 1992. Dividing the 1993 INTU entries into *branched* and *unbranched* ones, they also found that the branching condition favored follow up — 3.84 follow up entries, over 0.69. Of course, these effects may have been partly due to the teacher or the presence of the researcher.

Tracking data¹ revealed a number of additional problems (Hewitt & Webb, 1995): (a) a small amount of opening other students’ notes (low “familiarity”), (b) a small

---

¹Programs run on the CSILE server can obtain such information as the number of notes opened by a student, the number of words written by a student, and the number of notes of a given thinking type.
amount of commenting, and (c) a predominance of searches by authorship. On average, a student in John’s class opened only 21 per cent of all of the class’s notes in 1993/94; that was more than any class for which data was collected. To address these problems, Knowledge Map, add-on client software to be used in conjunction with CSILE, was designed. Its design goals were: (a) to make note relationships and note clusters visible, (b) to emphasize the work of the community over the work of the individual, (c) to facilitate exploration, and (d) to make knowledge building and collaboration a goal (Hewitt & Webb, 1995). John and his 1994/95 class used Knowledge Map for two units, the second being the unit of chapter 4. Hewitt and Webb found that familiarity was up to 41 per cent in the first, and 48 per cent in the second unit. But although students liked Knowledge Map, they raised an important issue when asked to provide specific input into the design process early in the first unit. As Hewitt and Webb wrote:

One particularly telling proposal was to provide each group [of students] with its own map. As one student explained, “you don’t really need to see what the other groups are doing.” It became obvious ... that despite repeated reminders of the importance of cross-group collaboration, most students still focused on their own group’s work. ... Consequently, it was decided to take a more class-oriented approach in the “Heat” unit. Rather than have notes clusters represent group divisions, they instead represented different lines of inquiry. And instead of working in groups of 3 or 4, the entire class was instructed to work as a single research team. (Hewitt & Webb, 1995, p. 16)

In the second unit, John had the students contribute individually to discussion notes that the whole class worked on. That approach, however, led to a new problem at the end of the unit: some students complained that only a few students could contribute “good principles” to discussion notes John had created at the end of the unit for recording what the class had learned. A teacher problem introduced by Knowledge Map, at least for John, was that it became harder to see the progress of a particular student. For assessment, he resorted to printing notes out, and arranging them by student, but this was very time consuming.

The knowledge map of the second unit, taken on the last day of the unit, is shown in Figure 3.2. The three largest clusters are the discussion notes John created to begin the work on CSILE for this unit.
Figure 3.2: Knowledge Map at the end of John's "Heat" unit in 1994/95
Rectangles denote discussion notes, squares graphic notes, and circles text notes. The three largest clusters in the center of the window radiate out from three discussion notes John created to start the CSILE component of the unit. The isolated cluster of four notes on the left are the discussion notes he created to bring the unit to closure (chap. 4).

3.1.2.3 How students use the CSILE design features.
During the first decade of the CSILE Project, CSILE has been used mainly in Grades 1-8. In the earliest grades, the software is introduced through its graphic notes. Only after the children have become comfortable with the technology and can make some graphic notes, does the teacher introduce text notes. In Grade 1, many children cannot yet read, type, or write very well, and a teacher or another student may help by reading notes aloud, asking the children what they think about them, and then typing a response for them. In an inner-city school with a large Hispanic population, somewhat older children translated notes from Spanish to English, and conversely. In all these classrooms, there was evidence for knowledge building.

2In 1995, CSILE became one of four "beacon technologies" in the National Centres of Excellence Telelearning Research Network. Directed at Simon Fraser University by Linda Harasim, the mandate of this consortium of Canadian universities, government, and the private sector, is research and development of computer supported learning technologies and educational practices that make use of them in K-12 and post-secondary education, in the private sector, and in teacher education. As part of this mission, recent research with CSILE has been extended to Grades 9-12 classrooms and post-secondary education.
Woodruff and Brett (1993) concluded that Grade 1/2/3 children wrote comments that indicated intent to help others. In a paper-and-pencil task, students in Linda’s Grades 1/2/3 class wrote more “constructive” and “substantive” comments than a control class that had never used CSILE. The teacher who used CSILE (Linda) revealed a focus on goals when she explained how she had introduced specific features of CSILE to her class:

Cut-and-paste was not introduced until the students were accustomed to the notion of looking at other’s work, not in order to “take” something from them, but as part of the “exchange” that goes on where, “one day I may have something that helps you,” and tomorrow the situation may be reversed. Even after the functions had been introduced, discussions were always meaning-centered, the focus being on what was interesting about the note or picture. Consequently, students tended to focus on issues of content and understanding. (Woodruff & Brett, 1993, p. 91).

Evidence of knowledge-building, as part of extended projects, begins to appear in Grade 3 or 4. For instance, one Grade 4 class investigated cockroaches in 1996/97, over a period of three months (Lamon, Caswell, Scardamalia, & Chandra, 1997). Marie-Ann and her Grade 4/5 class sustained an extended unit on the science of flight, as well as shorter units on astronomy.

In Grades 5/6, most teachers introduce CSILE in the context of a problem on which the class has already done other work. Teachers use the features of CSILE according to their own preferences. John, for instance, preferred discussion notes because “text notes are just for people working on their own, and we are working in groups” (CSILE Video Archive, 1994). The instructions he gave to his class during an introductory lesson about CSILE at the beginning of the 1994/95 school year may be summarized as follows:

- Students were to have one group member start a discussion note for their group, on the topic the class had been studying. John explained carefully why, although only one group member did this, everyone in the group — and the class — could add to this note.

- All students in the group were to enter the ideas they already had about the topics, using the ‘My Theory’ (MT) thinking type. John also explained that students should read the theories of other groups to learn what the whole class knew about the topic.

- Each student was then to think what they needed to find out “to decide if their theory was correct,” using the ‘I Need To Understand’ (INTU) thinking type. The group was to divide the things to find out in some way, “do research” to find

---

3This preference indicates an unfortunate misunderstanding of the purpose of text notes.
answers and, add these to the group's discussion note (using the 'New Information' [NI] thinking type).

Activity around these thinking types (MT, INTU, and NI) is typically sustained for 4-8 weeks in John's classes. At the end of the unit, there is an effort to bring things to closure. This can be done with notes of the 'What We Have Learned' (WWHL) thinking type or published notes.

3.2 Teaching, Learning and Research With CSILE

3.2.1 Models of CSILE use

Bereiter and Scardamalia (1987b) have proposed three models of teachers: Model A teachers focus on exercises; Model B teachers are more goal-oriented than Model A teachers, but themselves take on responsibility for metacognitive strategies such as deciding when and how to review or to plan; and Model C teachers, also goal-oriented, put this responsibility in the hands of the students. Clearly, Model C teacher behavior is the most consistent with CSILE design aims.

Lamon, Lee, and Scardamalia (1993) examined the interaction between teacher model and the amount and nature of reflection students did in the CSILE database, for John, Paul, and Ann-Marie. They classified John and Ann-Marie as Model C, and Paul as Model B. They based their decisions on the terminology the teachers used to describe their classrooms in structured interviews. For example, asked about "the greatest difficulty in a classroom faced by a teacher," the Model B teacher (Paul) used terms like "control and discipline," whereas the Model C teachers were concerned about "staying on top of what they're doing" (Table 2, Lamon, Lee, & Scardamalia, 1993). Asked what students would need to become successful learners, the Model B teacher mentioned "social goals" and "self esteem," while the Model C teachers mentioned "learning goals."

In a prior study, Lamon, Abeygunawardena, Cohen, Lee, & Wasson (1992) examined CSILE and non-CSILE students' reflections on their own and others' work as represented in portfolios for writing, mathematics, and science. That study showed CSILE students to be more reflective than non-CSILE students. Lamon, Lee,
and Scardamalia (1993) also used portfolios, but examined students' comments on their portfolio entries, rather than the entries themselves. Most of the analyses suggested to them that Paul's students were less reflective than John's, who were in turn less reflective than the second Ann-Marie's students.²

All three teachers said they had never used portfolios in such a reflective way before. Lamon, Lee, and Scardamalia asked to what extent they had incorporated the reflection strategy into their teaching. John had appropriated the portfolios, but rather than using them for reflection, he saw them as a repository of finished products — print-outs of published database notes (cf. Note Status in section 3.1.2.1). He noted: "If you look ... you read what they're producing ... if it's work on CSILE, you read what they're producing, and you get a general sense that the quality is improving, and the depth of thinking" (pp. 24-25). Paul, on the other hand, did not appropriate the portfolios at all:

... It was an opportunity for the kids to reflect on their work, and sometimes it was work that they hadn't had the opportunity to do that with, and I think that the portfolios helped that. But otherwise, it didn't have too much to do with my teaching. ... What they put into the portfolios was material that they had been taught without the portfolios in mind. (p. 25)

In another study of John and Paul's classes, Bereiter and Scardamalia (1991) extracted two models of database use (pp. 3-4):

- **The Independent Research Model.** Students work relatively independently, raising their own questions, seeking answers to them, and reporting what they have learned. (Paul's class)

- **The Collaborative Knowledge-Building Model.** Students jointly plan investigations, assign subtopics or tasks, and comment constructively on one another's contributions. (John's class)

They claimed that while the Independent Research Model is in principle "viable," its activity structure is conducive to questions that are answerable from a text. As they put it, "because students are expected to seek answers to the questions they generate, there is an incentive to ask questions that they know can be found in the available reference material" (p. 5). Such an approach could be task-oriented.

---

²For writing peer commentaries and mathematics peer commentaries, students in Ann-Marie's class were found to be more reflective than were those in John and Paul's classes, who did not differ from each other.
Bereiter and Scardamalia used a step-wise regression analysis of measures of constructive activity, knowledge quality, and language achievement gain, using five predictors: (a) initial status (age, math, language), (b) productivity, (c) exploring work of others, (d) collaborating, and (e) advanced processes. Separate analyses were done for the two teacher models. The dependent variable singled out for discussion was knowledge quality, because it was an overall significant predictor for both classes. However, the way the independent variables contributed to the variance was very different:

- In Paul’s class, productivity and advanced processes were significant predictors.
- In John’s class, initial status, exploring work of others, and collaborating were significant predictors.

Paul’s class also had superior language gains (vocabulary), for which the only significant predictor was initial status.

Brett and Woodruff (1993) examined how teachers come to realize a classroom culture in which students are “authentically in control of their own learning” (p. 3). They used 11 experienced CSILE teachers from schools in three regions of the US and Canada as the experimental group, and 11 pre-service teachers, randomly selected from a pre-service program that stressed constructivist approaches to teaching, as the control group. They used a questionnaire to probe these teachers’ beliefs about teaching and learning, examined CSILE teachers entries in a public portfolio, and examined the private portfolios of the four CSILE teachers at one school, who had “already worked in a team teaching situation, so had an existing collaborative peer support for one another as they began to work with CSILE” (pp. 4-5).

General conclusions from the electronic portfolios were (Brett & Woodruff, 1993, pp. 6-8):

- CSILE teachers revealed willingness to reflect and analyze their own level of knowledge and understanding.
- CSILE teachers actively created “a culture of the classroom in which learning and knowledge advance is primary and tasks and activities are subordinated to these goals.”
- They continually monitored the level of students thinking — where they were stuck or moving ahead, and where the class as a whole was in its thinking.
- They monitored and continually adjusted their interventions based on a balance between avoiding intervening too much and leaving a student floundering.
Moreover, as teachers took on the challenge of subordinating tasks to goals, the superficial level of students' skills became apparent to them. For example, teachers commented that students were not used to searching for information in different texts to answer a particular question. And, CSILE teachers said, students needed much help in generating questions which some teachers have described as "the big questions" (p. 9). Finally, according to the CSILE teachers, students were not taught "study skills ..., they [were] being shown how to take control of their learning."

One teacher commented:

I agree direct instruction in metacognition is not the way. It works when children experience the need for problem solving, induce their own, discuss them ... and practice them. I see this practice as being the daily work of kids on MacCSILE. (p. 10)

### 3.2.2 Assessment studies

Scardamalia et al. (1992) have reported a number of assessments obtained from John and Paul's classes. Where a control class was used, it was taken from a Grade 5/6 class at a nearby school of similar socio-economic status.

- **Quantity of writing.** Students in the CSILE classes produced about 1 word of text per minute of time signed on to CSILE. Of course, that time also included reading other notes and making graphic notes. Averaging 30 minutes per day, and not counting non-CSILE writing, that is more than the average American high school student produces (cf. Mullis, Owen, & Phillips, 1990, cited in Scardamalia et al., 1992).

- **Depth of explanation.** Hand-written answers to the question "What have I learned from this unit?" were obtained, using a Depth of Explanation scale. CSILE students were found to be significantly higher than the control group (\(M = 2.89\), compared with 1.20, \(p < .001\)). A replication study in a subsequent year, using a constructive activity scale previously used by Scardamalia & Bereiter, 1991) revealed a significant effect that favored the CSILE condition for John's, but not Paul's, class.

- **Knowledge quality.** The results for this measure were similar to those for the depth of explanation measure at \(p < .001\).

- **Question asking.** Students were asked to write, on paper, what they wondered about the topic they were about to study (selected by the teacher). The questions were rated on a scale ranging from asking for isolated facts to asking for explanations. The collaborative knowledge-building class (John's) was significantly higher than the control class \((p < .01)\), but Paul's class was not.

- **Canadian Test of Basic Skills (CTBS).** Five of the subtests of this test battery of basic skills were administered in October and (an alternate form) in June of the following year. Overall, using the aggregate of all experimental- and all control group scores, the CSILE groups outperformed the control group \((p < .01)\). Examining gains on subtests, the CSILE classes were found to have gained more than the control group.
on the vocabulary subtest (approximately four times as much, cf. Fig. 8 in Scardamalia et al., 1992; \( p < .01 \)). Dividing the classes into above and below average on the basis of incoming language composite scores, it was found that the advantage of CSILE to gains in the composite language score was not limited to above average students (\( p < .001 \)). Indeed, below average students gained more than above average students in all three classes.

### 3.2.3 Students' beliefs about learning: high level goals

Lamon (1992) conducted a study using a Jasper Woodbury problem (Cognition and Technology Group at Vanderbilt, 1990, 1992; see also Lamon et al. 1996), which requires solving a large number of hierarchically related subproblems. After watching a video tape that described the problem, students worked either in small face-to-face (cooperative learning) groups or used CSILE. A problem is often that children lose sight of the higher level goal when solving the sub-problems (Scardamalia, Bereiter, & Lamon, 1994, pp. 220-221). In the CSILE condition there was a significantly higher number of references to high level goals. Lamon also administered pre- and post-tests of word problems similar to the content of the super-problem, and found CSILE students outperformed the control group.

Lamon, Chan, Scardamalia, Burtis, & Brett (1993) tested the following hypothesis: If collaborative knowledge building with CSILE changes students’ activity from task-oriented to goal-oriented, does it also change their approach to learning to a more active one? Students in John, Paul, and Ann-Marie’s classes and two non-CSILE control classes were given pre- and post-test questionnaires about their beliefs concerning learning, as well as “difficult texts,” about which they were asked to answer some questions. The questions required them to extract underlying concepts that had not been taught. Two texts (on evolution and on photosynthesis) were used in counter-balanced fashion. In the Fall there were no differences between the conditions, and the children’s responses showed evidence of misconceptions. In the Spring, the CSILE classes outperformed the control group, \( M = 4.68 \), compared with 4.25, \( p < .05 \). Lamon et al. also found a significant but modest correlation between the two measures at \( p < .05 \).

### 3.2.4 Electronic- and face-to-face discourse

To conclude this review of research with CSILE, I describe two studies that have compared aspects of electronic discourse with CSILE with face-to-face classroom
discourse. In the first study, Earl Woodruff (1995) observed triads of pre-service teachers as they helped each other to understand two articles related to a psychology unit of a post-graduate B.Ed. program. In the CSILE condition, the students worked at adjacent computer stations, but were not allowed to communicate with each other verbally. Woodruff concluded that: (a) there were 6.4 times as many objectification episodes in face-to-face discourse, compared with CSILE (p. 84), (b) almost twice as many requests went unfulfilled in CSILE (p. 94), and (c) that over 91 per cent of “the discourse in the face-to-face setting was not relevant to promoting an understanding of the text and was therefore noise in the correlation statistic” (p. 144). The version of CSILE used was different from that in the studies of this dissertation, and the student population was also different. (IBM computers RT/PC computer workstations running System 5 UNIX under X-Windows, version 11, release 3, were used, see Woodruff, 1995, p. 61 for details.) Nevertheless, Woodruff attributed effects he observed to the computer medium, arguing that similar effects would be obtained in other networked computer systems.

Andrew Cohen examined three aspects of learning science in CSILE and face-to-face conditions: (a) meta-processes, (b) equitable collaboration, and (c) problem-based actions (1995; Cohen & Scardamalia, in press). Meta-processes referred to the monitoring of the students’ own ideas, monitoring the ideas of others, and coordinating the ideas of all participants to create a more integrated framework for their work. Cohen used an enhanced face-to-face condition, in which triads of students in Paul’s Grade 5/6 class used computers in groups of three to develop Interactive Physics™ simulations to work through a unit on gravity and the solar system, a unit Paul had done on numerous occasions before. Students were also given a paper-and-pencil “proforma” to record their scientific discourse. In the CSILE condition, students also worked in groups of three, but had individual access to computers, and used an electronic proforma, accessible from CSILE. Cohen concluded that: (a) in the CSILE condition there were more meta-procedural exchanges at $p < .03$; and (b) although there were similar amounts of self-monitoring in both conditions, there were more instances in the CSILE condition where one student reflected on the ideas of another student at $p < .01$. 
3.3 Paul's Classroom During Study 2

3.3.1 Summary of prior research on John and Paul's teaching

This subsection compares John and Paul's approaches to teaching with CSILE, as evidenced by the research that has been reviewed in the previous sections. The purpose of this contrast is to present John and Paul as examples of teachers who used different approaches to teaching with CSILE. Clearly, successful knowledge building can be expected to depend critically on the extent to which a "culture of understanding" (Cohen & Scardamalia, in press) is present in the classroom, which in turn depends in the teacher's theoretical commitments. The contrast drawn here goes a considerable distance to understanding differential success at knowledge building in studies 1 and 2.

The contrast may be drawn along the following three issues.

- **Teacher model.** Lamon, Lee, and Scardamalia (1993) interpreted John as a Model C teacher and Paul a Model B teacher, the major distinction being that John put responsibility for meta-procedures in the hands of students.

- **Modeling intentional learning.** John modeled intentional learning for his class. He incorporated the portfolios into his teaching and played a crucial role in the development of discussion notes, threaded discussion, and Knowledge Map. When he saw that despite his careful instructions, students still primarily saw their responsibility to share knowledge within their own groups, he did the next unit without groups. There was no evidence in the studies cited that Paul contributed at this level to dialogue about CSILE and classroom practice. Nor was there evidence that, if he valued knowledge building, he was aware of research findings drawn from his classes, such as the models of database use study (cf. Bereiter & Scardamalia, 1991).

- **Task-orientedness.** The productivity and advanced processes variables were significant predictors of knowledge quality for Paul's class studied by Bereiter and Scardamalia (1991). Scardamalia et al. (1992) found superior vocabulary gains for Paul's class over John's, and interpreted this as evidence for a task-oriented approach to teaching. (That is, basic literacy, as measured by these tests, rather than progressive problem solving, was a goal.)

The remainder of this chapter probes these three issues with data collected as part of study 2, in order to determine if the above characterization of Paul as a Model B teacher who used the independent research model applied at the time of this study. The data sources are: (a) semi-structured interviews with Paul and with his students, (b) video recordings of two "class meetings" (i.e., class discussions) in Paul's classroom, and (c) additional field notes of other such meetings. I visited
Paul's class once every 5-8 school days; many visits were organized to coincide with on the rug sessions that had to do with the science unit. Each of these sessions lasted 25-30 minutes. Six of eight groups of students who had worked together on CSILE during the unit were interviewed near the end of the unit (April, 17-18, 1997).

3.3.2 Teacher model

The following is an excerpt of an class discussion Paul's class had about making “good questions for CSILE” (Dec. 5, 1996). This session followed two weeks of work by the students (in small groups) on science kits, and served as an introduction to the work on CSILE that was to begin. The students had already been introduced to CSILE earlier in the year. Paul used questions provided by former classes, on a variety of topics, to illustrate a scheme of four levels of questions, ranging from open-and-shut to open-ended questions.5

1. Paul: Well, already we have some differing opinions, don’t we? ... Can you find a level 4 question in any of these questions? Stephen?
3. Paul: What is gravity. OK. What — how do the rest of you feel about that? (Students call out number 2, number 1). We have a difference of opinion here. ... Does anyone want to [go?] with that for a second? Nathan what do you think?
4. Nathan: Well, I think that would be a level 4, but there is also for some of these questions, there’s a whole new level. It’s sort of a mix between 1 and 4, or there is one answer but, as in four, there are so many different ways to think about it. ... I think there should be a whole new category.
5. Paul: Yes. I think Nathan has raised a very interesting point here, and that is, we can talk about level 4 questions as having a lot of theories. We can also talk about questions that generate other questions, which is what you are saying. Ah, the one Stephen has picked out, you have the basic question, What is gravity?, but it also generates a whole lot of other questions as well. I agree with Stephen, it is a level 4 when you talk about it that way, but there’s a sense in which it is a level 1, yes.
6. Paul: Susan?
7. Susan: At the very bottom, with the star. ‘Has gravity remained the same throughout history?’
8. Paul: Aha. ... Has gravity remained the same throughout history. Obviously it does generate a lot of theory about that one, and you’d have to ask a lot of other kinds of questions too, to have a discussion on that. (pause) Darren, have you got one?

Paul used the following four examples to explain the four levels. Level 1: How many students study in [this classroom]? Level 2: What do students in [this classroom] study? Level 3: How do students in [this classroom] study? Level 4: What is a good education?
Paul appeared to understand the open-ended nature of knowledge building. His scheme of questions was a progression toward open-ended questions, he liked questions that could generate other questions. During our interview he remarked on this issue:

I guess it's a particular interest of mine that a discussion does not necessarily have to stay, you know, rightly on topic. ... Simply starting a discussion and going in various directions. And I think that one of the things that I was happy to allow was that if they got talking about force and energy and leading to a black hole, that was fine, they could carry on with black holes and do as much as they wanted to on black holes, or whatever they find themselves into .... (Feb. 20, 1997).

It is also clear from the above excerpt that Paul was in control of the discussion. He called on students (1, 3, 6, 8); he validated statements (5, 8); and he summarized (1, 3, 5, 8). Four of the eight speeches were Paul's. All this is typical teacher behavior. Fisher (1997) talks about a "necessarily asymmetric relationship" between teacher and students in face-to-face discourse, which means that "when seeking out or formulating new ideas, there is a tendency for pupils (and perhaps teachers too) to see the teachers' ideas as the ones which should be accepted" (p. 22). Of course, it is not problematic in Phase 1 of the LKB Model that the teacher dominates talk — even a didactic lesson could have been fine. What is problematic for intentional learning, however, is that almost all of the meta-procedures were carried out by Paul, not the students. That is Model A or Model B teacher behavior.

The purpose of a later, February 5, 1997, discussion was to evaluate where the class was with the unit, and to establish a way of finishing it. This sort of event is located well into Phase 3, and should be a democratic discussion. The transcript, however, revealed that the teacher-student asymmetry described above was still strong; there was also some evidence of a gender difference in participation. Table 3.1 shows the number of students who participated actively in the discussion, the number of speeches, and the number of words per speech. The gender effect is with respect to the number of participants only; on the other measures boys and girls were similar. (The speeches made by girls were quite important, so additional data could show that girls make contributions that were more central to the purpose of the discourse.) Previous research (e.g., Gallas, 1995; Roth, 1998) has shown that
dealing with domination of science class discussions by boys is a challenging pedagogical problem.6

Table 3.1: Participation of boys and girls in class meeting (Feb. 5, 1997)

<table>
<thead>
<tr>
<th>Issue</th>
<th>Boys</th>
<th></th>
<th>Girls</th>
<th></th>
<th>Teacher</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Students actively participating</td>
<td>9</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Speeches per participant</td>
<td>27</td>
<td>1.7</td>
<td>3.0</td>
<td>0.0</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Total speeches</td>
<td>24</td>
<td>-</td>
<td>6</td>
<td>-</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Words per speech</td>
<td>24.0</td>
<td>17.7</td>
<td>20.8</td>
<td>14.4</td>
<td>42.5</td>
<td>26.1</td>
</tr>
</tbody>
</table>

Note: 15 boys and 11 girls were present.

3.3.3 Modeling intentional learning and task-orientation

As asked if any important issues had been missed in the interview, Paul said:

With these classrooms, this [is a] grade 5/6 classroom, so my grade 5 students, whom I've always kept as grade-6es, they become valuable technicians and helpers, you know, to the kids that are new to the classroom. That's an important thing. And I've used the kids a lot, in terms of doing things on CSILE. When I'm busy, I would say ... (points at a student desk), "Would you go and ...?" (points to the computers). And the kids know lots about it that I don't know. (Feb. 20, 1997)

Paul's point is important. All teachers at this school depend on assistance from students. For example, in a lesson introducing CSILE to her Grade 1/2/3 class, Linda used experienced students to show the new students how to create a note, save it, and so on (CSILE Video Archive, 1994). Having experienced students help less experienced ones while working at the computers is also a good teaching strategy. With "the kids know lots about it that I don't know," Paul indicated that they told him what was going on in the database. That is beneficial as well. What is troublesome is the degree to which Paul depended on others for knowledge of the database. In this unit, most of Paul's knowledge of the database resulted from reports from students and researchers, and perhaps looking at computer screens while the students were at work. It was not that he did not bring his expertise to bear on the database discourse, he did that through face-to-face discourse, but that he accepted uncritically others' interpretations of the class's limits of understanding. His view might have been quite different. If his efforts to support the class's inquiry

6For further discussion of gender differences arising from teacher-student interaction, see Sjoberg and Imsen (1988); according to Kahle (188), such gender effects become more pronounced in high school.
were misdirected, he would have no way of knowing it. In informal conversation, Paul said that he did not have time to read the database. There was also evidence that the class was working with rather limited knowledge of CSILE functionality. During group interviews, students in two groups said they had had difficulty accessing notes because they had not known how to create a note that could be edited by all group members.

Finally, there was evidence of a task-oriented approach to teaching and learning with CSILE in Paul’s interview:

One of the great strengths is that it promotes self-direction, self-learning. What they do on CSILE they have to do themselves, they have to, uh, work on their own. If they work in groups. ... Um, the uh great strength is in the knowledge building. It’s not possible in a classroom situation, say a classroom with 30 computers without CSILE, it’s not possible to get that depth of knowledge.

... I’d say it works better on CSILE. I think that CSILE — well, first of all, CSILE organizes it better. Uh, if I were to do something like that here in class, these mounts of paper.

3.3.4 Conclusion

The data discussed in sections 3.3.2 and 3.3.3 support the characterization of Paul as a Model B teacher who used the independent research model made earlier in the chapter. This is in contrast with John, who was characterized as a Model C teacher who used the collaborative knowledge building model (Bereiter & Scardamalia, 1991). It is worth mentioning that even in learning environments with constructivist underpinnings, Model B teachers are far more prevalent than Model C teachers. Several of the instructional programs described in section 2.3.2 reflect Model B teaching. While Model B teaching is philosophically preferable to Model A teaching (it pays more attention to metacognition), it undermines intentional learning and knowledge building because it assigns responsibility for metacognitive strategies to teachers, not students; of the three models, Model C teaching is most consistent with knowledge building.

---

7One of the most striking examples of knowledge building occurred in an earlier class of Paul’s. In one note, a student asked why sponges have three ways of reproducing while humans have only one. This resulted in a series of 12 notes on this issue (cf. Bereiter & Scardamalia, 1991, 1993; Scardamalia et al., 1992). That Paul was completely unaware of this episode, until a researcher told him about it months later is another indication that he tends not to know where the limits of understanding are in his databases.
In the next two chapters, I present case studies of two science units in which CSILE was used. Neither of the teachers used the LKB model (I had not yet developed it at the time they planned their units); not surprisingly, both units fall short of the kind of educational experience the LKB model is intended to guide. But, as we shall see, the units fall short of this goal in different ways. The first unit was planned in accord with knowledge building epistemology; it raises important questions about the role of the teacher in knowledge building — when, and how, should a teacher participate actively in the discourse? The second unit, the discussion of section 3.3.1 and chapter 5 suggest, was not conducted in accord with knowledge building epistemology; this fact can, in part, account for differential knowledge building and concept articulation achievements between studies 1 and 2.
CHAPTER 4

STUDY 1: A VIEW FROM SUBJECT MATTER KNOWLEDGE

The study reported in this chapter examines the scientific merit of students' scientific explanations in the context of an inquiry into the interaction of heat and matter by John’s Grade 5/6 class of 1994/95. The research questions are (from chapter 1):

1. To what extent do children, using CSILE and having available a teacher committed to knowledge building values, develop scientific understanding, as judged from CSILE transcripts and knowledge of the literature on children’s ideas in science?

2. What, if any, is the relationship between the knowledge building strategies students used in the database and the depth of understanding they developed?

4.1 Heat and Matter

This study takes the point of view of a subject matter specialist in the field of the student inquiry, and asks to what extent the claims made by students would be acceptable to a physicist. Such an evaluation must be made keeping the purpose of the claim in mind (Lemke, 1995). For, instance, while physicists know that the kinetic theory of gases taught in high school physics and chemistry does not tell the whole story of the behavior of gases, it is adequate in many situations. Physicists and chemists do, in fact, use the theory in those situations. In the context of the unit analyzed in this study, it is assumed that John’s global goal for the discourse in CSILE was progress toward scientific understanding of the subject matter. In addition, students may have had local goals on particular occasions, such as lending support to a line of reasoning introduced by another student.

The study examines student work as (partial) scientific theories, that can be elaborated further with subsequent education and experience. There is, however, considerable debate among physicists around the teaching of heat. This section,
therefore, describes my stance in this debate and the "scientific understanding" it entails. Figure 4.1 shows the conceptualization of the subject matter that is used to frame the study. It is far from a complete conceptualization, but an adequate one for the present purpose; it stresses a dialectic in which theory development and elaborating what the theory can explain are mutually constituted. The major components of the conceptualization are briefly described below.

The first law of thermodynamics: Physicists solve problems involving heat by specifying a system, the state of that system, and interactions of this system with its surroundings. The system might be an open container of ice, described by its energy. Understanding the system involves understanding the nature of the substance in question (e.g., the phases, molecular structure, etc.) and ways in which energy can be stored by that substance (e.g., random motion of molecules, lattice vibrations, or chemical bonds). Thus one aspect of learning the domain of heat and matter is learning the differences between solids, liquids, and gases. The second idea attached to the first law in Figure 4.1 — interactions — describes ways to change the energy of the system. Heating a substance constitutes one way, doing work on the system (e.g., compressing it) is another.

A number of science education researchers advocate abolishing the word 'heat' in favor of 'heating' (cf. Harris, 1981; Heath, 1974; Sommers, 1987). In my own teaching I have stressed the role heat plays in the first law — not a form of energy but a source of energy — and have addressed the non-material character of heat by comparing it to other ways of transporting energy.1 In my opinion, talk like 'heat flowed from A to B', 'heat sink', and 'heat capacity' is unavoidable because it is well established in everyday and scientific discourse (see also section 2.4.3). It would be better to be clear about what we mean by such statements than to avoid them altogether. The results of this study indicate that children in grades 5/6 can grasp the distinction between a material system and interactions with that system. That distinction turns out to be a deep one in physics, that children will not have to give up later. That is not true of the notion of heat as a form of energy. And using only the verb 'heating' misrepresents the way scientists talk about thermal phenomena.2

---

1 A wave is an example of energy transport without material transport.
2 In the first law, I use the convention that heat is been positive if it increases the energy of the system. This is contrary to practice by most physicists in North America, but as Arons (1990) argues the this sign convention can be introduced in university; the convention I use avoids unnecessary complication in the present context of school science.
The framework emphasizes a dialectic between scientific theory and what such a theory can explain. The links are indicates the hierarchical nature of the framework. The shaded boxes indicate the parts of the framework the students worked on in this unit.
The second law of thermodynamics: Two aspects of this law are important to this study. First, heat flows spontaneously from a high temperature to a low one, and ceases when the temperatures have become equal. Second, temperature is conceptually distinct from energy — it is a statistical parameter and does not, like energy, depend on the quantity of substance present (see also Wiser & Carey, 1983).

What concepts can explain: The lower part of Figure 4.1 shows my conceptualization of the phenomena that the children in John’s class discussed most. Although some students discussed convection, radiation, and conduction as examples of heat transport, no students wrote about making glass and fire as involving chemical change. Thus, in the conceptual framework ‘Chemical Change’ represents a super-ordinate concept that was missing from the discussion, but that I considered valuable in its interpretation. In the figure, concepts that the students did discuss are represented by shaded boxes.

It is tempting to see the model as presenting a distinction between theory and application, but this is not the intended distinction. The intellectual activity is similar in the two components — both depend on theoretical as well as empirical work. Consider, for example, a study of thermal expansion via the equation \( y = \alpha x \), where \( y \) is the percentage change of the length of a bar of some material under a temperature change \( x \), and \( \alpha \) the linear coefficient of thermal expansion. Such a study could give students experience with the idea of a mathematical function and develop in them a degree of numeracy, that is, an appreciation of the range of values that \( \alpha \) has for different materials.\(^3\) It would give students knowledge of the importance of the phenomenon of thermal expansion in specific situations, including knowledge of its effect size, but it would also be a phenomenological study. The known range of \( \alpha \) values is something that a theory about heat and matter must be able to explain, and this puts a constraint on theory development. As we shall see, students’ theories do not have this level of explanatory power.\(^4\) A second point

---

\(^3\)The idea of numeracy, as used here, is due to Case (1996b) and his co-workers, who have done considerable work on the teachability of the idea of mathematical function with children of the ages studied in this chapter.

\(^4\)Some justification of the term central concept is needed. The component near the top of Figure 4.1 demarcates thermodynamics from other domains of physics. It involves concepts from single-particle mechanics — work, energy, force, etc. — and concepts that are needed for treating the class of problems (multi-particle systems) that thermodynamics addresses, such as entropy and temperature. This model takes the particular point of view of situating heat in physics, but concepts not regarded as central here — e.g., manufacturing process — may become central in conceptualizations that take
about the conceptual model is that it does not accurately describe students’ at this grade level.

4.2 Method

4.2.1 Subjects

The 31 students of John’s combined Grade 5/6 class of 1994/95 participated in this study; 9 of these students were in Grade 5 and 22 in Grade 6; 12 students had used CSILE for at least one school year, the remainder having had exposure to it in earlier grades. The school is a middle class, urban elementary school.

4.2.2 Materials and procedure

Science kits: Prior to working on CSILE, the class worked through a science kit on heat and matter that consisted of the following “experiments”: 5 (a) examining thermal expansion by using a ring and a ball; (b) examining heat conduction in different objects; (c) heating bimetallic strips to see them bend; (d) heating different types of liquids; and (e) heating or cooling air captured in a container to see how the air pushes the water level in another container up or down. Students divided themselves into small groups to work on these experiments for a period of two weeks.

CSILE: For this unit, CSILE 1.5 — with threaded discussion notes — was used in conjunction with Knowledge Map. (Thus, when students logged on to the database, they were presented with a graphic display that revealed the connectedness of the notes, similar to that shown in Figure 3.2.) John began the work with CSILE by creating three whole-class discussion notes that asked how heat affects solids, liquids, and gases, respectively. Students added their own theories to these discussion notes, but they also created other notes on their own initiative. They improved their theories by discussion and research. Students spent approximately

another viewpoint (e.g., engineering). As always, what counts as a central concept depends on the purpose of the conceptualization.

5The students used the word ‘experiment’ to refer to these activities; however, they would more properly be referred to as hands-on laboratory activities.
45 minutes daily on science, including time to do library research, engage in class discussion, and work on CSILE. After approximately seven weeks, John created four new discussion notes, on which the class worked for an additional three weeks. Three of these asked what the class had learned about solids, liquids, and gases, respectively; the fourth question, ‘What general principles has the class learned about how heat affects matter?’, spanned solids, liquids, and gases, and reveals John’s desire for the class to reach a high level of conceptual knowledge.

4.2.3 Data and measures

4.2.3.1 Overview.
Since the purpose of the study is primarily to inform model-building, rather than demonstrating causal relationships, I chose a qualitative methodology, within which some of the data are quantified and analyzed statistically. Thus, the scales used are not based on a theory or model that is to be tested, but are grounded in the data and my understanding of teaching and learning in environments with constructivist underpinnings. The statistical estimators such as p-values and effect sizes serve as indices that help to describe the data. I interpret a statistically significant effect as one that should be taken into account in model building. To show causal relationships and generalizability across diverse educational settings, an experimental study would still be required. Such a study would have to include a careful analysis of the face-to-face discourse (see, for instance, Lampert, Rittenhouse, & Crumbaugh, 1996; Roschelle, 1996; Roth, 1998; Roth & Roychoudhury, 1994).

The present data corpus, moreover, does not allow inferences about conceptual change at the level of individuals. To see this, it is important to understand how a note comes to be written. Typically, a group of students may read a note, and engage in face-to-face discourse in an attempt to understand it and/or compose a response. Based on this discourse, a student may enter a response for the group. Each note, at best, represents a "best explanation" that is based on face-to-face discourse and reading of the database. In addition, if a student introduces a concept late in the unit, this does not mean that he/she could not have used the concept earlier — there may not have been a need to do this earlier. Therefore, what the analyses can show is how the conceptual content of the domain becomes more fully
articulated by a group of learners as the database discourse develops. Hence I use the phrase concept articulation rather than conceptual change.

The study consist of six components. I first provide an overview to show how the components fit together; Table 4.1 then provides an overview of the analyses of each component, the scales used, and reliability measures used; finally, I discuss these in some detail.

- **Basic literacy skills.** Writing is a central activity in Learning with CSILE, so one would expect basic literacy skills to influence success at knowledge building in this medium. Therefore, the Canadian Test of Basic Skills was administered near the beginning and end of each school year at this school. For this study, incoming scores (obtained in the Fall) and change scores during the year are examined for vocabulary, reading comprehension, and spelling. In addition, general features of the writing in the CSILE database are examined: the number of words and the number of notes written, and the Flesh-Kincaid reading level (determined from all of a student's writing, taken as a single sample, and using word processing software). These data are used to contextualize the case studies (see below).

- **Holistic evaluation of scientific merit of writing.** The writing of individual students is examined in the aggregate (i.e., all the writing is taken as a single sample) to evaluate if it passes the test of "acceptability to a subject matter specialist." These evaluations are used to construct a grouping variable, learner achievement, which is used in the remaining components of the study.

- **Case studies.** Four case studies document a variety of learning experiences in this unit. They examine three variables relevant to the LKB model: (a) the extent to which students develop a research agenda, (b) how they work this agenda to develop knowledge, and (c) sharing knowledge with the class to advance collective understanding. The case studies also examine the end-of-unit knowledge represented by the students' CSILE discourse.

- **Concept articulation.** The conceptual framework of Figure 4.1 is used with a vocabulary based on a glossary taken from a physics textbook (Hewitt, 1987) to retrieve notes from the CSILE transcript. While this glossary was not used in teaching, it allows us to examine to what extent students used a scientific "register" (see Lemke, 1995) in their writing. Two aspects of the use of terms from the glossary are examined for the retrieved notes: (a) the number of notes in which each term is used, and (b) how the terms are used to convey ideas.

- **Knowledge building activities.** Three kinds of activities central to knowledge building — discussing, questioning, and commenting — are examined for the purpose of understanding plausible causes of disparities that the case studies and concept acquisition analyses reveal. Discussion notes are examined in the greatest detail, using a diagrammatic technique to show relationships between note entries, as well as a checklist.

- **Note ratings.** All 630 notes in the database are rated using a qualitative scale that reflects the kind of scientific contribution (as an explanation) the note makes to the
database. The distributions of the notes over the categories of this scale are analyzed.

Table 4.1: Overview of component analyses, variables, and reliability measures, study 1

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Issues probed/Method</th>
<th>Variables</th>
<th>Reliability measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic literacy skills</td>
<td>- Contextual factors that may influence success with the unit</td>
<td>- Canadian Test of Basic Skills (CTBS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- No. of words</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- No. of notes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Flesh-Kincaid reading level</td>
<td></td>
</tr>
<tr>
<td>Holistic evaluation of scientific merit of writing</td>
<td>- “Scientific acceptability” of a student’s writing, taking into account evidence supporting as well as refuting this hypothesis, presented together in a “writing portfolio.”</td>
<td>- Learner achievement</td>
<td>- Blind second rater</td>
</tr>
<tr>
<td>Case Studies</td>
<td>- Phase 3 (LKB model) knowledge building characteristics</td>
<td>- Research agenda</td>
<td>- Narrative account</td>
</tr>
<tr>
<td></td>
<td>- Theory building</td>
<td>- Research process</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Collaborative effort</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- End-of-unit knowledge claims</td>
<td></td>
</tr>
<tr>
<td>Concept articulation</td>
<td>- Differences in progress between learner achievement levels</td>
<td>- Term frequency</td>
<td>- Random selection of notes from CSILE transcript</td>
</tr>
<tr>
<td></td>
<td>- Use of scientific vocabulary</td>
<td>- Sentences</td>
<td>- Narrative account</td>
</tr>
<tr>
<td></td>
<td>- Articulation of conceptions that use these terms</td>
<td>- Term density</td>
<td></td>
</tr>
<tr>
<td>Knowledge building activities</td>
<td><em>Discussion notes:</em></td>
<td>- Term working</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Role of “mixed framework” notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Collaboration (Phase 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Length of discussion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Questions:</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Differences in the quality and quantity between learner achievement levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Comments:</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Differences in the quality and quantity between learner achievement levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Note ratings: scientific contribution to database</td>
<td>- Distribution of notes over five types of contributions, and differences between learner achievement levels</td>
<td>- Scientific contribution</td>
<td>- Time-delayed second rating by author</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A number of these analyses are described below.
4.2.3.2 Holistic evaluation of scientific merit of writing.

The database transcript was divided into two periods: T1 consists of the first 1/3 of the note entries, and T2 of the remaining 2/3. In terms of the LKB model, T1 begins with the work on CSILE late in Phase 1, and should extend to the consolidated research program, that is, early in Phase 3; T2 then would account for the remainder of Phase 3. But it must be stressed that the present study preceded the LKB model by more than a year, and forcing such a fit may not be warranted — the teacher did not plan the unit with the LKB model in mind. The idea behind the holistic evaluation of writing was to present examples of writing that would probably pass a test of “acceptability to a subject matter specialist” together with examples that would not, to arrive at an overall judgment of the student’s writing during that period. Thus, following Miles and Huberman (1994), the writing portfolios present inferences about the scientific merit of writing, together with examples of evidence for them. In all but two outlier cases,6 all of the student’s notes in a given period were used, although text that seemed to be extraneous to the claims being made (e.g., personal commentary) was removed.

Rating criteria: A writing portfolio was judged to be scientific if a learner provided one explanation that invoked a scientific model — molecular or macroscopic — with a sufficient amount of surrounding textual material to indicate that the student understood this model (e.g., an elaboration or example), but without compelling evidence for commonsense notions. An example of an explanation that provided a scientific fact, without surrounding textual material, was not deemed sufficient. Nor was a single example with evidence for animism or materialism interpretations of heat sufficient to make a portfolio unscientific. The reasoning required by this criterion is similar to, but not as demanding as, reasoning that Driver, Leach, Millar, and Scott (1996) have called “model-based reasoning” in their framework for characterizing features of students’ epistemological reasoning in science. Somewhat weaker evidence was allowed if several propositions conveyed the same meaning, that is, if they cohered. (The propositions could come from one entry or from several.)

6Two students, Matt and Sean wrote 5527 and 5119 words respectively ($M = 1598$ for HFC learners, see Table 4.3); for them, the writing portfolios were limited to the first two pages to ensure that arriving at an overall judgment of the writing would not be more difficult than for other students.
Reliability: All 62 of the portfolios were rated by two independent raters. For T1, 27 portfolios were judged unscientific and 4 scientific; for T2, 18 were found to be unscientific and 13 scientific. The raters agreed on 92% of the 62 portfolio ratings.

Example: Allan’s T2 portfolio is shown in Table 4.2. The first three entries were judged to be supporting the hypothesis of acceptability to a subject matter specialist, and the last one refuting it. Moreover, the supporting evidence was in two subdomains — solids and gases. This portfolio was therefore rated scientific.

Table 4.2: Allan’s writing profile for T2, showing three pieces of evidence supporting and one refuting the hypothesis that Allan’s writing passed the test of acceptability to a subject matter specialist. This portfolio was rated scientific.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Text sample</th>
<th>Type of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>solid</td>
<td>&quot;...the water molecules which make up an ice cube have less kinetic energy (a kind of energy that is in everything that moves, the faster the object is moving the more kinetic energy it has) than those in hot water. (Hot water molecules have more kinetic energy and move much faster.) Sometimes the hot water molecules bump into the edge of the ice cube and transfer some kinetic energy to it. The transfer of kinetic energy continues until the energy is shared between the mixture.&quot;</td>
<td>supporting</td>
</tr>
<tr>
<td>gas</td>
<td>&quot;When the air rises up it makes room for itself and makes the cold air go down. This process will repeat itself till the flame is taken away from under the container.&quot;</td>
<td>supporting</td>
</tr>
<tr>
<td>gas</td>
<td>&quot;My theory is that gas is one of the states of matter and is most of the time invisible. The molecules in gas are more farther than in liquids or solids.&quot;</td>
<td>supporting</td>
</tr>
<tr>
<td>fire</td>
<td>&quot;My theory is that when two sticks are rubbed together they make fire because the molecules in the stick are quite dry and when you rub them they get very hot and create fire.&quot;</td>
<td>refuting</td>
</tr>
</tbody>
</table>

Grouping variable: The T2 portfolio rating was used as a grouping variable for the other analyses. The 18 student achievements with unscientific T2 ratings were labeled Low Focus on Concepts (LFC), the 13 with scientific ratings High Focus on Concepts (HFC). The proportion of Grade 6 students whose achievement was HFC was higher than that of Grade 5 students (45% compared with 33%, z(31) = 3.15, p < .001).

7Throughout this study and study 2, pseudonyms are used for all participants. In quoted writing by students, spelling has not been corrected unless otherwise noted, and any italics have been added (CSILE uses only ASCII text).
4.2.3.3 Concept articulation.

Terms like 'force', 'energy', 'impulse' are used in everyday situations, but in the context of scientific discourse such terms have specific meanings; this allows for a degree of exactitude in scientific discourse. One would not expect the writing of students who do not use such terms at all to resemble scientific discourse. A list of terms potentially useful in scientific discourse about heat and matter was compiled from the glossary of Paul Hewitt's *Conceptual physics: A high school physics program* (1987), adding some terms relevant to understanding fire and several other topics. After correcting the CSILE transcript for spelling mistakes, word processing software was used to search it for each of the terms in the list, using stems in the searches, and also including words that might be treated as synonyms by students: e.g., 'atom' and 'molecule'. Of course, a student's intention in writing such a term as 'temperature' may not be scientific at all, but the absence of this term from the discourse should be a red flag that the discourse was not scientific. All occurrences of the terms were marked in the transcript.

To probe semantic aspects of how the terms were used by students to convey ideas, 20 notes were randomly selected (10 for each learner achievement level), and the following variables examined: term frequency, sentences, term density, and term working. Each of these variables is described briefly below.

Term frequency: The number of terms per note (counting each occurrence of a given term). This variable is expected to have a global effect on the scientific merit of a note: notes judged to be of high scientific merit should have more terms than other notes. However, this measure confounds scientific merit with writing output — longer notes may contain more terms than short notes. Threshold and ceiling effects are conceivable: notes with fewer than a certain number of these terms may be incapable of communicating a scientific conception, while above a certain other number, additional terms may not improve such effectiveness.

Sentences: The number of sentences per note. A well expressed note provides some context, a thesis statement, and additional statements that elaborate on the

---

8 This is an important problem with lexical searches. An alternative method, Latent Semantic Indexing, is more sensitive to the problem of synonymy and also to the related problem — that a term is not used in a way that suggests a particular type of context (e.g., scientific) — of polysemy (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997; Storey, 1996).
thesis (e.g., examples or explanations of its limitations). Many notes consist of only 25-30 words; such notes tend to be less valuable than longer notes, say of 60 words, because one or more of these rhetorical components is missing.

**Term density:** The number of terms (counting each occurrence of a given term) divided by the number of words in the note. One would expect the more articulate notes to have a higher term density than other notes. If a thesis statement introduces the term 'kinetic energy', this term is likely to be used several times again, as the argument of the note proceeds. In particular, one would expect sentences with several distinct terms to be important — sentences that connect concepts to each other.

**Term working:** The number of terms divided by the number of unique terms in the note. In an articulate note, the meaning of the term can become more elaborated with each sentence in which it occurs — first in the thesis statement, and then in subsequent sentences that build up the note's argument. In some notes, however, high term working may indicate immature writing skills: successive occurrences of the same term do not enlarge what the term represents. If substantial differences are found between learner achievement levels on term working, it remains to determine if this is more easily attributed to the first or the second of these causes.

The transcript was divided into four segments (two arising from the levels of the time variable, and two from the learner achievement levels), and the number of notes in which each of the terms was used examined. The conceptions that the use of the terms suggested was then examined for T1 and T2 in order to describe concept articulation.

**4.2.3.4 Question form.**

**Question form scale:** A total of 207 questions were analyzed for form; 54 of these were the problem statements of discussion notes, and the remaining 153 'I Need to Understand' entries. For responding to questions, Harlen (1993) classified questions into four categories: (a) questions which are comments or expressions of interest, (b) questions asking for information, (c) questions which are investigable by children, and (d) questions which require complex answers beyond the comprehension of the children. Concerning the last kind, she wrote on a specific example, 'Why is the soil brown?':
These require some clarification in discussion with the child. They may be philosophical questions, if for example, the intention was 'Who made the soil brown?' These kinds cannot be answered and are best treated as questions which are an expression of interest. Taken at face value, however, it is possible to answer the question ... in terms of how we see the colour, the composition of white light, and what happens to white light reflected from objects that we see. (p. 91)

She argued that such answers would be too complex for elementary school children, also noting that the questions should not be left unheeded, but shifted to others that can be investigated empirically.

I modified Harlen's scheme in two ways for the current investigation. Her first category — questions that are statements of interest — was eliminated. Distinguishing reliably between such questions and questions that require answers "too complex" was thought to be too difficult with electronic data, without additional clues to the children's intentions. In addition, the notion that some explanations are too complex needed modification. I concur with Bruner that "almost any idea can be explained to a child of any age in an intellectually honest form" (paraphrased). Certainly, as we shall see, some children in this study could provide and understand advanced explanations. A distinction thought to be useful was one between questions that were suitable for hypothesis testing, involving a single variable and mainly logico-mathematical deduction, and questions requiring more complex interpretation involving multiple variables and/or deductions of a more complex nature, including, for instance, ethical questions. Frequently, a task to be accomplished by children is to turn a question involving several variables into questions involving a single variable. The following (categorical) question form scale was used:

- **Level 1.** Questions that may require a complex explanation, but need clarification before they can be investigated. Often such questions begin with 'Why is ...?', 'How is ...?' or similar expressions.
- **Level 2.** Questions that require factual information. Information need not be interpreted, but merely reported.
- **Level 3.** Questions amenable to hypothesis testing. These are questions like 'Is A the case, or B?', 'What would happen if we change A?', and so on.

Clearly the distinction drawn between levels 1 and 3 need not be one between concrete-operational and formal thought: both types of questions can require formal thought.
Reliability: All 207 questions were rated by two independent raters; there was 86% agreement between the raters.

4.2.3.5 Examples of ratings with the question form scale

Level 1: “Why in the project #1 (c) the bymectoagic stirp [sic] bended?” was rated a level 1 question. This question does not ask for information without interpretation, nor is it of a form that could be tested. It lacks clarity. It could be answered by saying that the strip bent because it was heated, but it could also be interpreted as one that takes that for granted but aims for an understanding of underlying mechanisms. Or it could take those for granted as well, and aim for a philosophical understanding of why the world is that way — could it have been different? Staying with the second interpretation, one can see that the question requires a complex answer. It takes answers to many questions before one can go from the idea that molecules occupy more space when heated to that of materials “bending” when heated.

A second example of a level 1 question is: “Why does heat have different effects on different solids? For instance: If fire burns wood why does it melt wax and why does it expand metal?” Clearly this is a complex question; indeed, the way the student has framed it makes it clear that there are several questions here.

Level 2: “What happens to the molecules in fabric to make the fabric lose colour?” I interpreted this question as a request for a fact (i.e., not a level 1 or level 3 question). Whether or not it will be a successful question depends on the research skills of the student and the available resources. A search probing the properties of fabrics may yield information of color fastness, although the explanations found may not involve molecular models. At any rate, the student who posed the question did not ask for interpretation but merely for facts that could be interpreted later.

Level 3: “Are ice crystals symmetric, or are they arranged in a completely random form?” This question can be answered by formulating a hypothesis and alternate hypothesis; these can then be researched. (“Research” can take the form of library research.)
4.2.3.6 Collaborative quality of comment notes.

Three dimensions of collaborative quality: The 62 Comment notes were also examined. Brett and Woodruff (1993) identified three dimensions of collaborative quality: (a) intent to collaborate, (b) constructive feedback, and (c) substantive contributions. Constructive feedback could be criticism or praise. While Brett and Woodruff used scales for all three dimensions, in the present study comments were examined only for evidence that each of the dimensions was present in the comment.

Example 1: “I think you should write that on your entry [sic], and you still did not answer my other question.” This comment implies collaborative intent, and it provides criticism, that is, constructive feedback; however, it does not contribute content that could add to the scientific knowledge of the class.

Example 2: “I don't think that the burner is the flask. P.S. I like your drawing.” All three dimensions were judged to be present.

Example 3: “[student name], I think that you should add some information. I think that you should show what happened to the water when we did this experiment. I think that you should explain how the water reacted to different temperature changes during the experiment.” This comment has collaborative intent because the author’s concern is with making the note referred to a better note. The comment does this by providing criticism. However, while it indicates the kind of information to be added, it does not reveal its exact nature.

4.2.3.7 Scientific contribution scale.

All 630 note entries were rated twice by the author, using a qualitative scientific contribution scale, with a two-week interval between the ratings. The first time, the notes were dealt with one learner at a time, with the names of the learners masked; the second time, one classification category was rated at a time, with its name masked in addition to the learners’ names. There was 88% agreement between the ratings. The levels of the scale used were intended to convey the sort of scientific contribution the note entry made to the database.

A potential criticism of the note ratings is that no blind second rater was used. This choice, given the volume of data to be coded, was judged the best that could be
made. It would have been too expensive to have a second rater, with a similar background in the subject matter and relevant teacher knowledge as the author, rate all 630 notes. Of course, a second rater could have rated a small subset of the notes (i.e., 100-150 notes), but this would have introduced sampling problems, especially due to the large variation in writing output with respect to: (a) learner achievement, and (b) topic.

The design of the scale was emergent. After examining approximately 100 note entries for content, a provisional scheme was formulated that characterized in what ways an entry fell short of being scientifically accurate. Trying the scheme out on additional notes led to improvements, until I was satisfied that the ratings could be done reliably. The final scale was as follows. (The names of levels are in parentheses for future reference.)

- **Commonsense knowledge (CS):** No evidence for scientific explanation.
- **Mixed framework (MF):** Either a scientific concept was used in addition to commonsense ideas, or the conception used went some distance towards a scientific idea.
- **Factual knowledge (F):** A correct scientific fact, without surrounding text to suggest understanding, such as examples or elaborations.
- **Scientific knowledge (S):** There were three types of notes that were rated scientific. (a) notes that described correctly the application of a scientific concept; (b) notes that suggested the learner understood a scientific concept, but did not realize that there were limitations to its applicability or validity (e.g., when the learners failed to recognize that water can boil at less than 100°C); and (c) notes with assertions a scientist would probably accept.

- **Question (Q)**

**Triangulation:** The note ratings and portfolio ratings are mutually dependent because they both depend on the criterion of “scientific acceptability.” The note ratings are based on small text units, and they provide more information: the scientific contribution scale has five levels, whereas the portfolio ratings are dichotomous. To check the consistency of the two types of ratings, a discriminant analysis was conducted. Such an analysis computes group membership from a linear combination of a set of dependent variables. The variables used were the proportions of a student’s notes that were in a given category. Thus if a student wrote 12 notes, and 4 of these were in category 2, then $X^2$ was .33. For T1, 83.9% of
the 31 portfolio ratings were predicted correctly from the proportion of T1 notes rated scientific (eigenvalue .273, Wilk’s lambda .785), and for T2 90.3% from the proportion of T2 notes rated scientific (eigenvalue = .527, Wilk’s lambda = .655). In both cases, a step-wise method was used; variables were entered if $p < .05$ and removed if $p > .10$. The results confirm that scientific note- and portfolio ratings measure similar features of the students’ writing. This is the desired reliability check. But the high values of Wilk’s lambda indicate that the note ratings account for only a relatively small part of the variation in the portfolio ratings — 34.5% in T2.

4.2.3.8 Examples of note ratings.
Several notes are quoted completely, and their ratings discussed.

I think when you burn somethings like metal it start to turn black because when you put the wood and metal into the flame I think that the molecules and liquids splits out. If the flame still keep heating the metal until part of the molecules and liquids completely split and gone and that part well turn to black and that means molecules and liquids in the metal are dead. (Josha, LFC, T2)

Animism is a conception often suggested by children’s utterances, but frequently it can be assumed to be metaphorical (Driver, Squires, Rushworth, Wood-Robinson, 1994). The animism in this note entry did not appear to be metaphorical. That is why the note entry was rated commonsense knowledge, despite its mention of ‘molecule’ and the presence of ideas that could be useful in conceptual change (e.g., that molecules spread out under heating). In cases like this, where different parts of the note suggested opposite ratings, the coherence of each such part with the rest of the note text was used as a deciding factor.

The next note, by Cindy, was rated mixed framework:

I think that when the water gets realy hot the monecoulss transform in vapor. The vapor goes in the air and the monecoulss transform in air. (Cindy, LFC, T2)
It appears that Cindy saw the water transforming into air, but this passage might also be interpreted as saying that the water molecules mixed with the air. In any event, there is at least one idea here that can be used in getting to a more scientific explanation: during boiling, molecules leave the liquid.

9Using her theory of conceptual change across ontological categories, Chi might conclude from “the molecules ... are dead” that Josha made an ontological category error. If so, I would agree with her in this case.
An example of a note judged to be scientific is
When the heat from the sun warms up the (puddle) it makes the molecules in the water move faster and faster the warmer the water gets. At the surface of the water some molecules escape into the air forming water vapor. This process continues until the puddle is dry. (David, LFC, T2)

This recognition that evaporation can occur at low temperatures was evident in the writing of only one other student.

Finally, Nina’s explanation of cooking in a microwave oven was rated scientific:

Food has molecules of water which all point in the same direction. For instance if you put a piece of meat in the microwave the microwaves will pass through the meat, making the molecules line up with the rays. The microwaves will pull the molecules making them twist back and forth. The molecules move over 2,500 billion times in one second which heats the food. If you cook something over an ordinary stove the heat rays will cook the surface of the food and then the middle which will take more time then if you cook in a microwave. (Nina, HFC, T2)

This information was probably looked up and miscopied — the dipole moments of the water molecules are initially not aligned — but Nina’s explanation of cooking on a conventional stove seemed to be in her own words, suggesting that she could adapt what she had read to the present context.

4.3 Basic Literacy Skills and Writing Output

4.3.1 CTBS scores

Grade equivalent scores of vocabulary, reading comprehension, and spelling CTBS subtests were compared with pooled data from 22 other classes at the same school obtained between 1991 and 1996. With 556 cases, even minor differences between mean scores are statistically significant, and a one-way MANOVA showed main effects for all three of the subtests at $p < .001$.\(^\text{10}\) Examining the means, it could be seen that there were in some cases substantial differences between classes — even between successive classes of the same teacher. However, there were only moderate differences between the means of John’s 1994/95 class, and the means of the pooled data. The largest differences were for the Grade 6 spelling scores: the incoming

\(^{10}\text{For incoming scores on the three subtests, Wilk’s lambda was } .033, F(69, 1587) = 49.4, p < .001; \text{ for the change scores, Wilks’ lambda was } .391, F(60, 1224) = 7.54, p < .001.\)
grade equivalent scores ($M = 5.86$, $SD = 2.01$, $n = 21$) were .31 standard deviation below its pooled mean, and the gains of the spelling scores during the school year ($M = .78$, $SD = .66$, $n = 18$) .37 standard deviation above its pooled mean. Thus the literacy skills of this class, as measured by CTBS tests, were not unusual when compared with classes over a period of five years at this school. There was no significant difference between the incoming CTBS language composite scores of students with LFC and HFC achievements in Grade 5, but for Grade 6 students with HFC achievement were substantially higher ($M = 7.2$, $SD = 2.0$; Mann-Whitney test, $U(21) = 24.0$, $p < .03$, effect size 1.07). However, among the Grade 6 students there were five students designated ESL by the school board, several of whom had very low CTBS language scores, while there were only 2 ESL students in Grade 5.

4.3.2 Writing output and quality

There was a considerable disparity between the writing output of students with LFC and HFC achievement, respectively. In T1, nine of the 18 students with LFC achievement wrote less than one note in two of the note categories, compared with only three of the students with HFC achievement. Collectively, students with LFC achievement wrote 39.1% of the notes and 27.1% of the words in T1, while they constituted 58.1% of the class; in T2 this improved, students with LFC achievement now writing 46.1% of the notes and 37.8% of the words. The number of words and the number of notes, for the whole unit, were not normally distributed. There was much dispersion, for LFC as well as HFC achievement levels, as Table 4.3 indicates. A Mann-Whitney (Rank-Sum) test showed the mean ranks of the number of words written to be significantly higher for students with HFC achievement than for LFC achievement, $U (N = 31) = 63.0$, $p = .03$, effect size = .78; the difference in the number of notes was marginally significant, $U(N = 30), 69.5, p = .057$, effect size = .80. Flesh-Kincaid reading levels were obtained from the total writing output of each student using word processing software. Five students with LFC achievement produced an insufficient amount of writing for the procedure to be carried out. No significant differences were found between LFC and HFC achievement levels for the remaining 26 students, $U(N = 26) = 64.0$, $p = .29$.

11These statistics are based on the Grade 6 scores only. A two-way Grade by Learner Achievement ANOVA of the composite scores did not reveal any significant main effects or even a Grade by Learner Achievement interaction. For the effect size reported the root-mean-square of the standard deviations for each grade were used.
Table 4.3: Characteristics of notes written by Grade 5/6 students with Low Focus on Concepts and High Focus on Concepts achievement levels

<table>
<thead>
<tr>
<th>Measure</th>
<th>LFC (n = 18)</th>
<th>HFC (n = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Total No. of words</td>
<td>626</td>
<td>447</td>
</tr>
<tr>
<td>Total No. of notes</td>
<td>13.2</td>
<td>9.69</td>
</tr>
<tr>
<td>Flesch-Kincaid</td>
<td>7.03$^a$</td>
<td>1.43$^a$</td>
</tr>
</tbody>
</table>

$^an = 13.$

4.4 Case Studies

In this section, four case studies are presented that examine student work and achievement that can lend support to several aspects of the LKB model:

- Establishing a personal research topic.
- Researching that topic.
- Helping others advance their understanding.

I selected the cases to show a range of knowledge building performances. The first case, Matt, was an outlier in terms of writing output and the quality of the knowledge building activities he employed; I view Matt’s performance as a benchmark, that practice in other settings should aim to replicate for students with similar achievement (e.g., CTBS scores). The next two cases, Jerry and Emma, represent the high end of HFC achievement in this class, excluding the two outliers. I view Jerry’s performance as an example of the next level of achievement below that of Matt; Emma’s performance was near the LFC/HFC border. The fourth case, Sandy, is an example of a student with LFC achievement who was quite articulate. (Many students with LFC achievement wrote so little that cases could not be built from their writing.)

4.4.1 Matt

Matt was in Grade 6, and one of two outliers, writing 5527 words; his incoming grade-equivalent CTBS language comprehensive score was 7.7; his achievement was rated HFC.
4.4.1.1 Establishing a research question.

Matt began with the following My Theory entry in the 'How does heat affect liquids?' discussion note John had created to begin the work on CSILE.

My theory is that when liquids get to certain temperatures they will boil and eventually evaporate. I think that water, when it is boiled, is 100 degrees Celsius. It is hard to believe that water can get so hot. My theory is that if you took a different liquid such as milk, it would have a different boiling point and eventually would evaporate too. I think that the boiling point of milk would be much higher than the boiling point of water because milk is a much denser liquid. I think that the particles in the milk would take a long time to expand and start boiling. On the other hand I think that the particles of water will not take that long to expand and boil. For the same reason, milk is a denser liquid. I think that since milk is a denser liquid, it has more particles. The reason that I think this is that if a liquid has more particles it will take a longer time for the particles to expand and make the liquid boil. I think that there are fewer particles in water. Then the water will be faster to boil and evaporate. I think that if you took a liquid such as honey, there would be so many particles that it would take a very long time to get at a high temperature and maybe boil. I think that because honey is a very, very thick liquid. I think that if mercury in a thermometer was exposed to very hot temperatures, it would leak and maybe cause some damage. I think that people storing mercury in big amounts would have to be very careful, because it could cause great danger. (Nov. 25, 1994)

Italics in this note indicate examples of features of this text a scientist might argue with. For instance, the density of milk is only a few per cent higher than that of water. (Students tend to overestimate effect size.) Density does increase with the number of particles per unit volume, but so does it with the mass of the individual particles — a point missed by Matt this early in the unit. What is perhaps most interesting about this note entry is that it introduces scientific concepts (particle number and density), and that the note refers to the behavior of several substances. This note was a strong start, and an excellent example of a mixed framework note; it reveals the beginnings of scientific understanding.

Three days later, another student with HFC achievement wrote an entry in this discussion note that made reference to Matt’s paragraph about boiling honey. Matt then posted an INTU entry, which he subsequently cut-and-pasted it to a new discussion note, making it the problem statement. That is a knowledge building act. It says to others in the class, ‘I have found an important problem. Please all think about it and add your theories.’ The question Matt posed was:
I need to understand what would happen if honey and water were boiled? Would there be any difference in the boiling point of the two? Maybe we can try an experiment. (Matt, HFC, Nov. 28)

From a knowledge building perspective, this was a good opening because it already suggested a particular research activity. (If an experiment was ever done, it was never reported in the database.) A number of students would contribute to this discussion note, and it may be argued that with its creation Matt had formulated a preliminary research agenda. He would pursue the influence of density on a number of phenomena.

4.4.1.2 Helping others advance their understanding.

It turns out that Matt did relatively little further work on liquids, except in the context of melting and boiling, relatively late in the unit. He left his questions about the boiling point of honey for others to consider for the time being, turning his attention to other things. One of the first activities he turned his attention to was reading the database and commenting on other students' note entries. The following excerpts show that he (a) helped students improve their notes, (b) contributed theories, and (c) contributed scientific facts.

• ... If the wire expands when it is heated then why does the wire snap if it is heated too much? From your explanation it sounds as if it gets bigger and bigger when it is heated? Why would it snap then? Please explain. (Nov. 30, 1994)
• My theory is that fire needs oxygen to continue burning. I think that this is because the molecules are moving around and need oxygen to energize them. When you take oxygen away there is no place for the molecules to move around so the fire is put out. (Dec. 7, 1994)
• ... I found out ... that the boiling point of water is 100 degrees Celsius. (Dec. 7, 1994).

This commenting appeared to be peripheral to his main interest. He did little further work on fire and did not develop a new line of research from a comment on “turning sand into glass” in January. In other cases, he did do that. For example, having made a contribution to the note ‘What is heat?’, in which he stated that the sun is our main source of heat, he began a new thread in that note to discuss how light travels. (This discussion note is discussed in detail in section 4.5.2.)
4.4.1.3 Researching the role of density in the interaction between heat and matter.

Apart from the commenting activities just described, Matt pursued his interest in density in the domain of solids such as wax, wood, and metals, as the italicized phrases in the following examples illustrate.

- My theory is that the difference in reaction of the solids depends on the molecules inside. I think that wax has molecules which are extremely crammed ... It melts, because there are so many molecules and that makes them hard to expand. ... (Dec. 14, 1994)
- I think that some solids bend because the molecules are dense .... The bimetallic strip will bend because the molecules are dense . (Dec. 15, 1994)
- I need to understand how dense the molecules are in different types of solids. (Dec. 15, 1994)
- My theory is that a wooden spoon takes longer to heat because the molecules are very spread out and take a long time to heat. I think that if you take a metal spoon, it will get heated sooner because the molecules are denser and take a short time to heat. (Jan. 12, 1995).
- I need to understand how dense the molecules of different metals are? Therefore how fast do they get heated? Are there big differences? (Jan. 17, 1995)

He also became interested in the effects of gravity on matter, particularly gases, and created a discussion note on that topic. He wanted to know if it had anything to do with the density of the molecules.

Two of Matt’s note entries from this body of work are discussed to give a flavor of the scientific quality of his writing while his research was in progress.

My theory is that metals are better conductors of heat than wood because there are more molecules in metals and that results in the metal getting heated a lot faster. I think that wood does not have as many molecules as metals therefore it will not get heated as fast. I think that different metals would have different points when they get extremely hot. For instance I think that a metal such as copper will not need to be heated for as long a time as brass because copper has more molecules and will get heated a lot faster than brass. I think that gases have very few molecules, therefore it takes a very long time for the gas to get very hot. (Jan. 17, 1995)

I think that the denser a solid the more molecules it has, because the solids are much lighter than gases. The molecules in the solid are very dense therefore the object is very heavy. If you take gas for an example, the molecules are few and far between and it is very light. Gas has some mass because it is affected by gravity. I think that gas is affected by gravity because if it was not, it would just zoom up into the air and be caught in another planet’s gravity. Since gases rise slowly, they are affected by gravity. The molecules in liquids are in medium density comparing them with solids and gases. The mass of a liquid such as milk is much lighter than a solid, but much, much heavier than a gas. (Jan. 2, 1995)
There are more molecules in metals ...” and “the denser a solid the more molecules it has” suggests that Matt continued to consider particle number (per unit volume?) as a variable influencing density, but not particle mass. The density of copper at 0°C and 1 atm. is only 2.5% more than that of brass (cf. Wilson, 1990) and probably does not need explaining at this grade level. “Gas has some mass because it is affected by gravity” by itself would suggest faulty logic — a physicist would say that gas is affected by gravity because it has mass, as Sean, another student with HFC achievement, did in another note. But Matt went on to assert that gas is affected by gravity on empirical grounds, drawing the conclusion stated in the first sentence of the second paragraph (in italics). An interesting argument! He followed this entry up with an I Need To Understand (INTU) entry, a strategy he employed frequently. Both of these notes are MF notes: they illustrate that Matt’s concept density required further elaboration, that he did not pay much attention to effect size; at the same time, they provided information that can be retained in scientifically more valid explanations.

4.4.1.4 End-of-unit knowledge claims.

Finally, two entries Matt wrote as part of conveying what the class had learned are presented. (Some condensing of the notes was done in the interest of brevity; problematic phrases are italicized.)

... When an object is being heated, the atoms move around faster and they take up more space. Therefore the object expands. I found that solids and liquids expand only a little, gases expand a lot more.

... If you have an ice cube and heat it, the molecules start moving around and cause heat. That will cause the ice cube to melt. The form of matter changed from a solid to a liquid. If water is heated even more it turns into a gas. All substances boil if they are heated at a high temperature.

... Heat can be caused by friction, two objects rubbing together and causing friction. This movement would eventually make heat and maybe a flame. This movement is called kinetic energy. Matches also need to [be] struck against something to cause friction in order to light.

Heat travels quickly through metals. This is proved by the fact that if you leave a metal object beside something hot, the metal object will start getting hot. ... Wood does not carry heat as well as metals. This means that metal is a better conductor of heat. ... On the other hand if you leave wood in extreme heat for a long time it will catch on fire. If you heated a metal it would get very hot and melt. Brass lets heat pass through itself. Wood gathers all the heat it gets and eventually it will catch on fire. This is because wood cannot let the heat escape. (Feb. 2, 1995)

You cannot boil water any higher than 100 degrees Celsius. The water will start to evaporate. After some time goes by, the water will be completely gone. If the
temperature of the water was high then it would evaporate faster, if the water was
colder, it would not evaporate as fast.
Water is made up of many little moving molecules. This is how water flows from place
to place. The warmer the water, the faster it moves, because the molecules are moving
closer. In the winter time, when the temperature is cold, the molecules move slower. When the
temperature gets to 0 degrees Celsius, they move so slowly that the water freezes and
turns to ice.
One way heat travels, is by convection. Convection happens when you are heating a pan
on a stove. The pan takes the heat generated by the stove and uses it to heat the matter
inside. We use metal pans because they absorb the most heat, they are one of the best heat
conductors. (Feb. 2, 1995)

Matt did get his qualitative description of the expansion of solids, liquids, and
gases correct here, which is an improvement over earlier notes. His description of
heat in the second paragraph suggests a confusion between random motion and the
transfer of random motion. In the second note, he did not distinguish clearly
between macroscopic and microscopic motion. At high temperature, the molecules
move faster in their random motion, but this does not imply a higher flow velocity.
His statement that molecules move slower in winter than in summer, while true, is
again an over-estimate of an effect size. On the Kelvin scale, a winter temperature
corresponding to -15°C is 88% of that corresponding to a summer temperature of
+20°C. That translates into a change of average molecular speed of 6%.

4.4.1.5 Summary.

Matt began with notes that revealed quite a bit of evidence for commonsense
(and scientific) ideas, but he made much progress toward scientific understanding
on several issues. In many CSILE classes, the best performances are not as good as
this. And, of course, Matt had well developed basic skills, which Oshima (1995)
found to influence knowledge building in a CSILE database significantly. This case,
therefore, is to be taken as an idealized case of a best performance. In my opinion,
consistent performance at this level by the highest achieving students across classes
and curriculum topics is a realistic goal to be worked toward.

How did Matt work the database? First, he established that he would work on
explaining differences in the behavior of several materials (e.g., honey, metals,
wood, gases) on the basis of several scientific ideas. He did this in all three of the
subdomains. Second, in at least two cases, he wrote INTU entries, raised the status
of these to that of a new discussion note, but then largely left the discussion notes to
be developed by others, turning his attention to other things. (Whether or not used a
specific plan to do this is not clear.) Third, he appropriated ideas for new lines of
inquiry from discussions peripheral to his main interest — e.g., 'How does light travel?'. That appears to be a benefit of the intention to help others, arising from thinking deeply about notes written by others. Fourth, there was evidence of reflective writing: there are examples where a My Theory entry led to an INTU note.

4.4.2 Emma

Emma was a more typical student with HFC achievement, writing 811 words, or slightly less than the class mean of 1034. Her CTBS incoming language composite score was 8.5, more than a standard deviation above the pooled mean for Grade 6.

4.4.2.1 Early contributions (T1).

Emma began with a My Theory entry in the 'How does heat affect solids?' note:

I think that heat affects solids by burning, melting, and expansion. When the heat is exposed to these different solids one of these things usually happen to it. For example metal expands when applied to heat, wax melts, and wood burns. But I think there are other ways that heat affects solids which I would like to find out. (Nov. 28, 1994)

Several days later she made the following entry in the 'How does heat affect liquids?' discussion note.

I think that when heat is applied to water it transmits the heat to it, and makes the water hot itself. When that heat gets to a certain temperature it boils, this is when the bubbles which have formed in the water begin to pop out onto the surface, and steam is produced off the top of the water. In our experiment today we were testing heat's effect on water, and when we finished steam was coming off the top, and bubbles were forming at the bottom of the beaker, but they were not popping, at that point the temperature was 90 c. I think that the boiling temperature of water is 100 c. (Nov. 30, 1994)

Following this, Emma did not enter any notes until December 12, and then only a few. Following up on her entry of November 30, she wrote:

I think that when heat is applied to metals the molecules which form the metal expand or bend. Depending on the material, the existing molecules I think get larger and produce more, or soften and bend with the effect of gravity. (Dec. 12, 1994)

Expansion of individual molecules (rather than the space occupied by a collection of them) is an example of what Driver, Squires, Rushworth, and Wood-Robinson (1994) have identified as macroscopic properties attributed to molecules. Emma’s other
entry that day was in the same discussion note (on solids), an INTU entry following an entry by a student with LFC achievement in which its author reported on the bimetallic strip experiment all the students had done before the work on CSILE. She wrote no additional notes until January.

To sum up the weeks from the beginning of the work on CSILE (Nov. 25) to Christmas vacation, there were several substantial periods in which Emma did not write at all and her notes were focused on surface features of the domain (observations) rather than scientific concepts. There is too little writing to know if she had a sense of what she wanted to accomplish, that is, if she had converged on a research topic.

4.4.2.2 Convection, conduction, and radiation.

After Christmas, Emma began to write more and focus on heat transfer. She began with liquids:

- When heat is applied to a liquid the molecules in the liquid start to move faster causing the liquid to conduct the heat and make the water warm. ... (Jan. 16, 1995)
- Convection is more frequent in liquids and gasses because it involves heat only reaching one point in matter then moving around causing the whole thing to be hot, this would be very hard for most solids because they do not change form very easily. For example, heat transfer happens when a pan of water is heated on a stove. ... [The] water caries the heat around the pan untill the whole pan is hot, that is the process of convection. (Jan. 16, 1995)

Both of these entries were rated scientific.

There is some evidence that Emma interacted in the database with Sean on the topic of convection. On January 20, Sean wrote a long entry on convection, heat conduction, and radiation in a discussion note Matt had created, ‘How does heat travel though space?’ It followed Emma’s note, in which the word ‘convection’ was used for the first time, by only a few days, so it is likely that Sean was, in part, inspired by Emma’s note, which did not explain what she meant by convection. Emma restated Sean’s conclusion later on the same day.

4.4.2.3 Helping others advance their understanding.

In T2, Emma wrote more comments than in T1 that were clearly intended to help others along in their inquiries, including Matt’s on the boiling point of honey and on
the effect of gravity on gases, and other discussions on the nature of heat and fire. I give one example:

I think that some gases are affected by gravity and some aren't. Good examples would be carbon dioxide which has no gravity since it is almost like air, or helium which does the opposite, by rising higher instead of falling with gravity's pull. (Jan. 6, 1995)

This note indicates that Emma was trying to help Matt in his attempt to understand how gravity affects gases. The note is interesting because it introduces, as a possible topic for discussion, apparent weight. 

4.4.2.4 End-of-unit knowledge claims and summary.

The final note entered by Emma was the following:

Certain solids are insulators, they are very bad conductors, such as wood, rubber and plastic. Many liquids and gases also insulate, even though heat travels through them using convection. For instance wool clothing is used in winter because the air between strands of its wool get trapped, since the air cannot move convection won't happen. This air keeps out the cold, and keeps warm air near the body. This also occurs with animals like sheep which have curly hair, they do not need other sources of warmth in cold climates because of this. Houses also use this with layers of special materials like polyurethane, foam and fiberglass which is put between the outside walls and the roof to stop heat from escaping. (Jan. 30, 1995)

Drawing from several everyday phenomena, this note reveals fairly deep understanding of convection. The comment that molecular structure affects how materials respond to heat is an advanced insight. There is evidence of incorrect understandings (e.g., that gases are not affected by gravity), but not of animism and material interpretations of heat.

In conclusion, Emma did not appear to work a well defined problem in different contexts, as Matt did, but she did contribute some high level ideas to the class's work. She was certainly very articulate, and may have benefited from a close collaboration with Sean and Matt, the most prolific contributors to the database. The value of her own contributions to the database, it appears, was limited by a relatively small writing output.

\[12\] In the case of a helium balloon, there is an upward buoyant force which is greater than the force of gravity, resulting in a net upward force (and hence acceleration).
4.4.3 Jerry

Jerry, a Grade 5 student, is the last of the HFC cases. He wrote almost twice as much as Emma (1482 words), and ranked third in the class on writing output. His incoming grade equivalent CTBS composite language score was 7.0, well above the mean for his grade.

4.4.3.1 Establishing a research question.

Like Matt, Jerry began with a substantial amount of writing. On November 25, he made entries in all three of the discussion notes created by John — a total of 334 words. A condensed form of his notes follows (problematic phrases are italicized):

... When heat comes in contact with some solids, the solids will melt. When cold comes in contact with most solids the molecules become closer together because the object is trying not to freeze. ...

On some fireplaces ... the fire has a temperature monitor showing what degree that you can put the fire temperature at and what the safe heat is and what temperature is not safe. ... Ceramic for instance, does not burn under a temperature of 900F immediately, but over a period of time the ceramic will be damaged. I think that what happens is that the molecules are affected by the long term heat and cannot withstand it any longer.

Another theory I have is that the molecules in the fake log are moving so fast and expanding that the object cannot stand the molecules moving so fast and then the log breaks down into the fire. (Nov. 25)

In most cases heat will evaporate the liquid. But it might not evaporate till a very long time. It depends on what the temperature the heat is at, and how much water there is. ...

When heat gets near a liquid at 100 Celsius it will boil. Gasoline is also a form of liquid, when heat gets near gasoline, the gasoline will catch on fire. The reason this happens is because gasoline has some sort of explosives in it. (Nov. 25)

I think that when heat gets anywhere close to a gas it will make a sort of spark. Some gas has a bigger reaction to heat than other gases .... (Nov. 25, 1994)

Like Matt’s, Jerry’s writing includes a number of unscientific claims (e.g., that objects try not to freeze, that objects “can’t stand the molecules moving so fast,” and that gasoline has “some sort of explosive”). It also makes reference to scientific concepts and a variety of phenomena (e.g., “heat will evaporate a liquid”), and provides scientific facts (e.g., “when heat gets near a liquid [water?] at 100 Celcius, it will boil”).

How did Jerry follow these notes up? He took up his point about fire from his first entry:

- Why does heat burn wood and evaporates water? (Nov. 28, 1994)
- What is there in water which causes a fire to be put out? (Dec. 1, 1994)
I think that the molecules in water have some sort of reaction to a fire so when water hits the fire the molecules react to the fire and somehow put the fire out. (Dec. 1, 1994)

Why does lack of oxygen put a fire out? (Dec. 1, 1994)

The last entry became the problem statement of a new discussion note he created on December 1. It appears, then, that he established a preliminary research agenda within a week of work on CSILE. Several students, including some who had LFC achievement, had by that time already mentioned fire in their notes, so he had converged on a problem that the class was interested in.

4.4.3.2 Researching fire.

He continued with a note that provided a theory that fire requires oxygen.

My theory is that fire, like humans, for some reason needs oxygen to breathe. I mean breathe in the sense that if it does not have oxygen the fire cannot survive. What happens is that the fire needs the air because the air keeps the fire going. ... If the fire can't get oxygen it will disappear. Air must have a special sort of molecule that fire needs, like humans. (Dec. 8, 1994)

"I mean breathe in the sense that if it does not have oxygen the fire cannot survive," more than previous entries by Jerry, gives a strong indication that animist statements can be metaphorical (Driver, Squires, Rushworth, & Wood-Robinson, 1994). He followed it up by creating a discussion note, and got it started with a theory. However, he would not return to this topic for some time, spending some time making comments on other students' notes.

On December 12, he also began a note sequence toward developing a whole-class inquiry about gases and gravity. This is another example of writing with the knowledge transforming model, through an interplay between My Theory and I Need to Understand entries. Entries 4-6 below are nearest neighbor entries in the same discussion note.

1. Are there atoms in gas, and if there are would they be further apart than a solid? (Dec. 12, 1994)
2. My theory is that atoms in gas. (If there are any) are further apart than a solid because gas isn't affected by gravity as much as a solid. Also gas can hardly be felt and you can't pick up gas and you can pick up a solid. Solid just fall to the ground when dropped, but gas floats. (Dec. 13, 1994)
5. I think that gas is affected by gravity because gas is matter and anything that is matter is affected by gravity. (Dec. 13, 1994)
6. Does gas weigh anything? What would happen if you had put some sort of gas in a box and the box could contain it. What would be the effect? Would the gas weigh anything inside the box because the gas is contained? (Jan. 4, 1995)

Before returning to the study of fire, he touched on several other topics, such as the expansion of solids, liquids, and heat conduction. His interest in the conduction of heat took him back to fire. His notes on fire did not suggest understanding of that topic, but they did include ideas useful in scientific explanations, such as the idea that "heat [temperature] builds and builds" if there is no way for heat to be taken away from its source:

... I think the reason why brass does not burn is because the heat is escaping the brass and by letting the heat go all through the brass and then going away. With wood, the heat stays inside it and the heat builds and builds .... (Jan. 16, 1995)

4.4.3.3 End-of-unit knowledge claims.

Jerry was a prolific writer at the end of the unit as well. He wrote about solids, liquids, gases, fire, and the nature of heat — a total of 503 words. He wrote longer than most other students, making his last entry on February 13. The notes he wrote were not copy-and-pasted versions of his own or other students' notes. He did not now, or earlier in the unit, write at all about sand and glass. I focus on notes that touch on fire, because he was probably one of the most articulate writers on the subject.

- The class knows that solids will also burn. ... Some types of friction will start a fire. The fire needs solids to create friction and once the friction has the fire created the fire does not need solids to make more friction to make it bigger instead the fire grows on itself. The one thing that the fire does need though is oxygen to keep the fire going. (Feb. 6, 1995)
- ... When a gas gets anywhere near to a flame there might be a short ignition period when the gas makes contact with the flame. There are different kinds of gases some have a bigger reaction to a flame then other gases. ... (Feb. 9, 1995)
- ... When heat gets near to a gas the gas in the molecules will expand intill the gas cannot stand it any longer at that point the gas will explode. (Feb. 9, 1995)

There is nothing here beyond surface features, although Jerry wrote many notes of scientific quality in other domains.
4.4.3.4 Summary.

Jerry’s performance was more like Matt’s than Emma’s was: he established a research agenda early and worked it. And he sometimes left discussion notes he had started for others to contribute to, while he focused on reading and commenting on other students’ notes. He was also a prolific writer, and collaborated with the same group of students as Emma. His performance is a more realistic example that future classroom work should attempt to replicate than Matt’s was — his writing output was higher than the class average, but it was only a few hundred words higher. His incoming CTBS score, was only about half a standard deviation above the Grade 6 mean, so the replication would probably be more readily obtained with Grade 6 rather than Grade 5 students.

4.4.4 Sandy

Sandy’s writing output, at 990 words, was slightly below the class mean, but higher than Emma’s. His incoming grade equivalent CTBS composite language score was 6.6, well above the mean of 5.7 for Grade 5. His achievement was rated LFC.

4.4.4.1 Early writing.

Sandy began with several entries in the ‘How does heat affect gases?’ note, between December 12-15, that were all concerned with the visibility of exhaled breath in winter. On December 12, he wrote:

I think that when we breath in the winter the smoke we see is a form of gas. I think that it’s our warm carbon dioxide and the cold oxygen from outside combinding. So it makes us see it. (Dec. 12, 1995)

Two days later he created two discussion notes on “How does heat affect gases?,” apparently not remembering that the teacher had already created a note with this problem (and announced this to the class). On December 15, he created yet a third discussion note and added the following theory:

I think that when we breathe in the winter the steam we see is a form of gas. I think that it’s the cold oxygen in the air and the warm carbon dioxide from our body combining. I think that when that happens we see smoke because when hot water and cold water combine we see steam. So I think it’s the same with carbon dioxide and oxygen. I think it happens because when the warm carbon dioxide goes into the air it hits the cold air making the gasses curl together and they pass through each other making us see the gas.
I think that smoke and gas aren't the same thing because when you put your hand through steam your hand becomes wet and warm but when you put your hand through smoke you can't feel anything. (Dec. 15, 1995)

The last sentence provided a promising theory, but it did not yet provide a resolution of the problem of why we see our breath. In any case, Sandy had established a research agenda quickly once he began writing.

4.4.4.2 Later writing.
Matt offered one theory to follow up on Sandy's, but, unfortunately, that is where things were left. Sandy did not write on this problem again. Instead, he wrote a fairly large number of comment notes and questions. None of these notes reached the quality and scope of his theory of December 15. Essentially all of them were concerned with observations and topics on which the class largely remained stagnant. Some examples are:

- I think that when you put a piece of metal over a fire, the metal will absorb the heat and amplify it by about 12%. I think because the metal is a form of rock and rock can make fire, that metal would be attracted to fire for sure. (Jan. 5, 1995)
- I need to understand why when you make when your doing [sic] the ball and ring experiment when you make the ball hot and the ring cold they don't go in but when you make them both hot they slip through easily. (Jan. 5, 1995)
- I think that solids melt because when the molecules that in the solids get really hot and they they pop and desolve. (Jan. 14, 1995)
- I don't think that all sun rays travel at the same temperature because we have different temperatures at different times of the year. (Jan. 16, 1995)

The idea that metal is a form of rock has several possibilities for improvement. As we will see in the next section, another student concluded that metal can be obtained from some rock by smelting. Another possibility for knowledge building would have been the idea of hardness. The class could have debated the adequacy of this idea for classifying solids, probably finding out that hardness is not what distinguishes metals from insulators. The second entry, concerning the ball and ring experiment, was a good one, but was dealt with relatively quickly by another student; it did not stimulate much discussion. The last note, about sun rays, could also have led to knowledge building. The claim about the sun rays not traveling at the same speed might have been refuted, and the problem it was expected to explain resolved another way. All these notes, except the third one, were rated mixed framework for their potential along these lines to drive knowledge building and the acquisition of scientific knowledge. Thus Sandy made important contributions to
the database. However, he never again established a substantial theory, and never converged on the central concepts of the unit (section 4.1).

4.4.4.3 End-of-unit knowledge claims and summary.

Sandy contributed several entries at the end of the unit, including to the 'What has the class learned about ...?' notes. Some of these were INTU entries; others still had some of the problems that disappeared to some extent in the other cases, such as teleological explanation and the inappropriate application of molecular models. Some examples are:

- I think that fake heat is made from different chemicals combining.\(^{13}\) (Feb. 1, 1995)
- When a mammoth gets hot, it's molecules shake and it gets tired. (Jan 30, 1995)
- We have learned that gases spread out so they are harder to see to the naked eye, and that we can't touch it because the molecules. (Feb. 13, 1995)
- A basic principle of heat is that heat always gets bigger by spreading, for example fire spreads and burns harder. This happens because fire burns things and becomes stronger, and by collecting oxygen it gets stronger too. (Feb. 13, 1995)

In sum, Sandy was a reasonably productive writer; the notes that have been quoted indicate that he was articulate as well — his notes were well written. He did start out with some focus, but quickly lost it, and then went from topic to topic without making substantial conceptual gains in of them. His writing was largely in observational, rather than conceptual, terms.

4.4.5 Summary

All of the students discussed were highly articulate, they all had CTBS scores well above the pooled average for their grades, and three of four of them had been designated "gifted" by their school board. Not all students with HFC achievement were this élite, but it was difficult to develop cases from the relatively small volume of writing of other students with HFC achievement. Even for Emma it was difficult to make inferences about how she worked to advance her knowledge. Only two of the cases (Matt and Jerry) suggested they established personal research programs. In the LKB model these research programs would still have to be consolidated into an overall research program for the class. Developing a research program may be a good way to ensure that all students contribute to the database from the beginning.

\(^{13}\) Some students referred to heat produced by human activity, in contrast to heat produced in nature, as "fake heat."
A reasonable goal might be to attempt to get the average number of words written up to Jerry's level — around 1400, up 35% from the current level for the class (although more than double the current level for students with LFC achievement).

The cases indicate two factors in addition to writing output that may influence successful knowledge building: (a) the subject matter topic a student researched, and (b) who the student collaborated with. Certainly, insight into the nature of fire was not evident in the last two cases. And Matt, Emma, and Jerry were part of a cohesive group that read each other's notes and commented on them (Sean was another important member of this group). Sandy, on the other hand, was not part of this group; indeed, he (and several other students with LFC achievement whose writing was examined in the process of developing the cases) collaborated with students with similar achievement levels. There was an interesting knowledge building mechanism at play in two of the cases. Both Matt and Jerry provided examples that suggest they reflected on their theory. In each case, the student wrote a My Theory (MT) entry, and followed this with an I Need to Understand (INTU) entry. In a further step, this INTU entry, which was usually embedded in a discussion note, then became the problem statement of a new discussion note. (See the sequence 1-6 in section 4.4.3.2 for an example.) Although there was evidence for individual commonsense notions in these cases (there are utterances by Matt and Jerry that could be interpreted as animist, for instance), there was also evidence for commonsense notions of a more general and interesting sort. These have to do with (a) effect size, and (b) the over-application of a theoretical idea. For instance, the temperature differences one observes between summer and winter lead to only small changes in the speeds of molecules. Moreover, Matt's ideas about density did not include the mass of the particles, only the number of particles. This inability to fail to see the limitations of applicability of an idea or algorithm has been noted before in several domains (e.g., Perkins and Simmons, 1988).

4.5 Concept Articulation

4.5.1 Scientific terms usage

The case studies have provided an account of the specific knowledge students developed, and how they went about doing this. In this section the focus is on the
overall scientific content worked on by the class, using a list of terms developed from Hewitt's (1987) glossary. Terms in the glossary, but not used by any students were: Brownian motion, caloric, conserved, element, equilibrium, free fall, fusion, hypothesis, joule, law, pascal, rate, saturated, scientific method, specific heat, sulfur, thermostat, work, and watt. Terms used by John to formulate the problems to be discussed in the 'How does heat affect ...?' discussion notes were excluded (heat, solid, liquid, gas, principle, matter).

4.5.1.1 Semantic aspects of term usage.

Despite the disparities in writing output already reported in section 4.3.1, the writing of students with LFC and HFC achievements, respectively, was sampled evenly by the searches for terms (see section 4.2.3.3). For T1, notes by 10 learners with LFC achievement were retrieved, and for T2 15, amounting to 62.5% and 93.8% of students, respectively; for students with HFC achievement these numbers were 69.2% and 100%. Of the variables listed in Table 4.1 for this analysis (sentences, term frequency, term density, and term working), only term frequency showed a statistically significant effect. Notes written by students with HFC achievement contained, on average, 1.45 terms (SD = .16), and notes written by students with LFC achievement these numbers were 1.28 (SD = .20), Mann-Whitney test, U (N = 20) = 21.0, p < .03, effect size .93. For all variables the HFC tended to be higher than the LFC means; although there were no statistically significant differences, some of the effect sizes were relatively large (means and standard deviations for students with HFC achievement are shown in brackets):

- **Sentences**: the number of sentences per note entry, U(N = 20) = 30.0, p = .12, effect size = .80. (M = 4.90, SD = 3.70)
- **Term density**: the number of terms divided by the number of words in the note. U(N = 74) = 494.0, p = .17, effect size = .45. (M = .086, SD = .064)
- **Term working**: the number of terms for every unique term in the note, U(N = 20) = 36.5, p = .77, effect size = .80. (M = 2.36, SD = 1.23)

The lack of significant effects for these variables is entirely expected with a sample of 20 notes. However, some tentative conclusions may still be drawn about the usefulness of the variables. Examining the p values in increasing order, it may be seen that the variables with effects closest to significance (p < .05) are term frequency and sentences (these variables are correlated, r = .95, p = .001). This suggests that writing longer notes, that contain more terms, may positively influence the quality
of notes, but that *precision of expression is of secondary importance*. This conclusion is consistent with the lack of a large difference between groups for the Flesch-Kincaid reading levels. It suggests that basic literacy skills influence the scientific quality of notes through (the quantity of) writing, and not, in addition, through the local features of the notes probed by these variables.

4.5.1.2 Use of terms in notes.

An important point to consider in the interpretation of the usage of these terms is that in a communal database, once one student has made a statement, another student may not need to make another statement to the same effect for quite some time. Thus, if a student with HFC achievement asserted in T1 that the boiling point of water is 100 °C, and a student with LFC achievement stated this much later, say in T2, this *does not mean* that the latter student did not have this knowledge earlier. This problem, inferring what students know from what they write about is essentially intractable. The best we can do is examine what they write about and examine if there is progress in the coherence of statements made, e.g. a "deepening and widening" (Oshima, 1995) of understanding. For instance, do students leave a topic quickly, or stay with it to examine limitations of their theories or the effects of variables not considered before — do their theories explain more with time?

General patterns of the use of terms were examined, see Table 4.4. Several terms that were used only a few times, with no difference between learner achievement levels and time, are not shown. It should also be noted that the categories are not independent: a note about boiling may also use such terms as 'kinetic energy' and 'molecule'. The main features of the data are discussed briefly below.

*Newly introduced terms:* These are terms that were used in several notes in T2, but hardly at all in T1. In some cases — e.g., 'kinetic energy' and 'absolute zero' — the simplest explanation is that these terms were learned during the unit. In other cases, a term may not have been needed earlier in the discourse. It is difficult to accept, for instance, that students could formulate theories about melting and evaporation in T1, but not about freezing. It is more likely that 'freezing' was introduced to underscore that it is a phenomenon in the same class as other phenomena that had been discussed earlier. The terms used most in this category are 'friction', and several terms that have been combined under energy transport (convection, radiation, and conduction). The use of 'friction' was limited (by all learners) to
recognition that friction is caused by two objects rubbing against each other, and that it can be a source of heat.

Table 4.4: The number of note entries with terms that could be used in scientific discourse, for students with Low- and High Focus on Concepts achievement levels, early and late in the unit

<table>
<thead>
<tr>
<th>Learning dynamic</th>
<th>Term</th>
<th>T1 LFC</th>
<th>T1 HFC</th>
<th>T2 LFC</th>
<th>T2 HFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newly introduced terms</td>
<td>absolute zero(^a)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>freezing</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>kinetic energy</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>internal energy(^a)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>friction</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>energy transport</td>
<td>2</td>
<td>0</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Substantial usage gains by both groups</td>
<td>boiling</td>
<td>5</td>
<td>15</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>evaporation</td>
<td>3</td>
<td>9</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>flame</td>
<td>9</td>
<td>6</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>melting</td>
<td>15</td>
<td>27</td>
<td>27</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>molecules</td>
<td>20</td>
<td>55</td>
<td>52</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>oxygen</td>
<td>3</td>
<td>9</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>wood</td>
<td>9</td>
<td>9</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>Used primarily by HFC learners</td>
<td>density</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>gravity</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>expansion</td>
<td>13</td>
<td>25</td>
<td>7</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>experiment</td>
<td>7</td>
<td>17</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>glass</td>
<td>16</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>

\(^a\)These terms were used only by Sean and Matt.

Note: Based on the aggregate of all searches with the terms.

Substantial usage gains by both groups: For ‘boiling’, although students with HFC achievement used the term more than students with LFC achievement, both groups revealed knowledge: they distinguished between boiling (the forming of bubbles) and evaporation, which, according to the students, follows it. In T1, students with HFC achievement identified 100°C as the temperature at which the bubble formation occurs (students with LFC achievement did not use the term). In T2, students with LFC achievement stressed that the quantity of water present diminishes as a result of evaporation and that the highest temperature water can be at is 100°C. These are important contributions, but students with HFC achievement, in contrast, did more to extend the conceptual base used for discussing boiling. They observed that the boiling point (a) depends on the pressure of the surrounding air, (b) is considerably lower at high altitudes, and (c) depends on the density of the liquid. (They did not identify a connection between [a] and [b].) They also noted that evaporation can occur at low temperatures.
Three of the terms in this category — 'flame', 'oxygen', and 'wood' — were included in the list of terms to conceptualize fire. ('Smoke' was used on only a few occasions, and more in the context of seeing one's own breath in winter than in that of fire, and is not included in the table). Most notes that used flame focused on observation, e.g., that one can move a hand through a flame without injuring it. There was no mention of a flame as a combustion zone. The situation was similar for oxygen. Students wrote in many notes that oxygen is required for fire, but do not seem to have realized that oxygen participates in a reaction. This finding is in agreement with previous studies of 11- and 12-year-old students. Driver et al. (1994) commented that most students know that air (or oxygen) is needed for burning, "although the function of air is not generally understood" (p. 87). And according to Meheut, Saltiel, and Tiberghien (1985) some students regard oxygen as necessary for combustion, but they do not think of it participating in the reaction. Finally, 'wood' was discussed extensively, especially by students with HFC achievement in T2. "Wood gathers all the heat and it eventually will catch on fire. This is because wood cannot let the heat escape" (HFC, T2). "Solids with looser molecules like wood can get hot faster and that other solids like metals take longer to get hot because the molecules in the metal are tighter together so the metal takes a longer time to catch fire" (HFC, T2).

Terms used primarily by students with HFC achievement: The topic addressed by these notes has already been addressed to some extent in the case studies (Matt) and is discussed further in the section on discussion notes (section 4.6.2.1).

Other terms: In T1, 'experiment' was used more by students with HFC than with LFC achievement levels, while in T2 the converse was the case. Only two notes, however, referred to doing an experiment at a juncture where it might have settled a question. All other uses of the term were in the context of describing what had happened in the experiments at the beginning of the unit. Other epistemological terms such as 'hypothesis' and 'evidence' were not used at all. In T1, students with LFC achievement wrote more about glass than students with HFC achievement, but this equalized in T2.
4.5.2 Concept articulation by students with LFC achievement

The conceptions written by students with LFC achievement are presented in Table 4.5. The wording used in Table 4.5 is for the most part that of the students, but duplication has been removed and in some cases two utterances have been combined into one; the organization of the table follows the lowest layer of the model of Figure 4.1. An analogous presentation for students with HFC achievement is given in Appendix A. Each subtopic is discussed briefly below.

Particless and making glass: There was little elaboration/articulation, although the notion that a molecule entails atoms (5) is important; that of glass as a form of sand (18) interesting. There was no shift away from the notion that molecules move away from the heat (3, 4).

Fire: Students with LFC achievement did make a number of important observations. Claims 21-24, while unsubstantiated, all have elements of scientific understanding, but they seem to me to fall short of insight. More than "mixing" (22) is involved: atoms from wood (mainly carbon) and oxygen participate in a chemical reaction. Smoke is the gaseous product of that reaction, mixed with small solid particles.

Energy Transport: There was little writing, but the last two claims in this topic (29, 30) could have made important contributions to the knowledge of the class if they had become widely known; there is no evidence that they did.

Thermal Expansion: The learners made several commonsense claims, but there was no writing that explores them further; nor did students with HFC achievement go much beyond saying that the molecules expand (Appendix A). It is not clear if students meant whether the molecules themselves become larger or if they did so in the aggregate (35). Students with HFC achievement offered better explanations, even in T1: e.g., "Objects expand because the molecules move quicker."

Phase Changes: There were some factual claims (e.g., 42-44), but most attempts at explanation were in terms of commonsense notions. An exception is a description of the water cycle (49).
### Table 4.5: Concept-Articulation by students with Low Focus on Concepts achievement

<table>
<thead>
<tr>
<th></th>
<th>T1 (n = 10)</th>
<th>T2 (n = 15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Although gas is matter it has no weight whatsoever.</td>
<td>Molecules move as far away from heat.</td>
</tr>
<tr>
<td>2</td>
<td>If gases were not affected by gravity all the air would rush to space.</td>
<td>An object consist of millions of clusters of more than one kind of atom, molecules.</td>
</tr>
<tr>
<td>3</td>
<td>Molecules move away from the hot spot.</td>
<td>If molecules get too hot they start to get soft, then melt.</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Molecules expand because they get hot.</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Water is the thinnest of all liquids.</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>HEAT AND ENERGY</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Friction makes heat, and is caused by pieces of the same substance rubbing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>together.</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Heat energy can only be transferred by vibration and collision of neighboring</td>
<td></td>
</tr>
<tr>
<td></td>
<td>atoms.</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Making Glass</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Sand turns into glass by it getting to hot.</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>There is something in sand that helps it turn into glass.</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Only extreme heat is needed to transform sand into glass.</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>The heat is too unbearable for the sand.</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>FIRE</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Wood burns because it has different molecules than other objects like metal.</td>
<td>Two things are needed before a fire can start: oxygen and heat.</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>When wood burns the oxygen molecules are splitting apart the wood molecules.</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>Burning is caused by molecules in a very hot object mixing (linking) with the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>molecules in the oxygen.</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>Iron molecules will not combine with oxygen molecules so iron does not burn.</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>While burning, a block of wood produces hot gases which heat the surrounding air.</td>
</tr>
</tbody>
</table>
Table 4.5 (continued): Concept Articulation by students with Low Focus on Concepts achievement

<table>
<thead>
<tr>
<th>T1 (n = 10)</th>
<th>T2 (n = 15)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ENERGY TRANSPORT</strong></td>
<td></td>
</tr>
<tr>
<td>25. Metals are good <em>conductors</em> of heat.</td>
<td>26. A <em>conduction</em> is a flow formed by an object to another.</td>
</tr>
<tr>
<td>27. Liquids and steam <em>conduct</em> heat.</td>
<td></td>
</tr>
<tr>
<td>28. Wood is a poor <em>conductor</em> (is an insulator) of heat.</td>
<td></td>
</tr>
<tr>
<td>29. Poor <em>conductors</em> of heat don’t have free electrons.</td>
<td></td>
</tr>
<tr>
<td>30. Air is more insulating than wood.</td>
<td></td>
</tr>
</tbody>
</table>

| **THERMAL EXPANSION** |
| 31. Heat *expands* and bends metals. |
| 32. Heat makes gases (helium) *expand*. |
| 33. Cold makes things shrink |
| 34. Heat can make fabrics lose their color. |
| 35. *Expansion*: the *molecules* in the metal become larger. |
| 36. The *molecules* and atoms in the piece of metal split farther apart than they are supposed to, making the metal *expand*. |
| 37. Metal can shrink, bend, and *expand*. |

| **PHASE CHANGES** |
| 38. During boiling the *atoms* pop. |
| 39. German silver is a metal that will not *melt*. |
| 40. Heat can *melt* glass. |
| 41. When water *boils*, steam rises into the air to create vapor. |
| 42. Report of some *melting* points. |
| 43. When the water *freezes* it expands. |
| 44. If you put salt in the water, it will not freeze. |
| 45. Water *boils* at 100°C or 212°F, and cannot be hotter. |
| 46. Thicker liquids take longer to *boil*. |
| 47. In *boiling* water, bubbles form underwater and steam rises under pressure. |
| 48. In *boiling*, water disappears from an open pot. |
| 49. When the humidity goes up the water on the ground *vaporizes* (& water cycle). |
| 50. Metal will *melt* because the heat breaks down the *molecules*. |
| 51. Things *melt* because the heat becomes too intense for the object to handle, and the object starts *melting* to cool itself off. |
| 52. Solids will *expand* until they can’t *expand* and then the molecules start turning into a liquid or they *melt*. |

*Note:* Terms in Table 4.4 are shown in italics.
In sum, there were a number of interesting claims that would appear to be advanced for the students' grade level (e.g., 2, 10, 11, 26, 29, 43-45, and 49), some examples that had elements of scientific explanations (e.g., 22-24), and a tendency to report factual information, in observational terms (e.g., 8, 28, 30, 39, and 45). In contrast, while students with HFC achievement also made unsubstantiated claims and reported facts, they wrote more articulate notes that provided explanations or provided evidence for thinking about factors not considered before that might influence an effect. Only a small number of instances is shown below; the full set can be found in Appendix A.

- At high altitudes the boiling point of water will drop considerably.
- A way to change the boiling point and the melting point is to alter the pressure of the surrounding air.
- When a solid melts and becomes liquid, the network of bonds that held its molecules together breaks apart, and the molecules wander all over the place. That is why liquids flow into any shape.
- Convection is more frequent in liquids and gases, because it involves heat only reaching one point in matter then moving around causing the whole thing to be hot.

4.5.3 Summary

The analyses showed that, apart from writing more notes, students with HFC achievement also wrote notes with more terms; but this effect appeared to be related to writing longer notes, rather than superior ability to write articulate sentences. With respect to concept articulation, there was evidence that students with HFC achievement went into the subject matter more deeply, while students with LFC achievement remained at the surface. This finding agrees with earlier studies of CSILE databases (Hakkarainen, 1998; Oshima, 1995). For instance, students with HFC achievement thought of variables not initially considered as relevant to boiling. An important issue for further research is: Why did students with LFC achievement write shorter notes, that used fewer scientific terms and remained more at the surface, than students with HFC achievement did? Two possible explanations are:

1. Students with HFC achievement began with highly superior subject matter knowledge, which allowed them to write longer notes early and to benefit more from transformative writing, which in turn allowed them to learn more subject matter knowledge. This phenomenon has been referred to as “the rich getting richer” in reading research (Stanovich, 1986; Walberg & Tsai, 1983).
2. Students with HFC achievement used intentional learning, of which transformative writing is an example (see chap. 1), and this led to greater learning for them, even if they did not initially have superior subject matter knowledge.

There is evidence that students who have used CSILE longer can provide deeper explanations (Scardamalia et al., 1992), but further work is needed that can show that students who begin with a disadvantage (writing output, subject matter knowledge), can within a reasonable time (e.g., several units) advance from this predicament.

4.6 Discussion notes

As explained in section 4.3, creating a discussion note can be a knowledge building act. In this subsection the discussion notes created by students are examined. The students created 54 discussion notes, with a total of 334 entries (53% of the database). As was the case with the case studies, the first discussion note to be examined, Is gas affected by gravity?, indicates an ideal, to be aimed for in other learning contexts. The next two discussion notes, What is heat? and How is heat made? were selected to explore the effect of two variables on the success of the discussion: (a) the mix of learner achievement levels represented by the students who participated in the discussion, and (b) the distribution of the note entries over note categories (from the note ratings). Chan, Burtis, and Bereiter (1997) proposed maximizing cognitive conflict, followed by sustained knowledge building as an effective strategy for conceptual change. Thus one might expect that discussions in which students with LFC and HFC achievement levels participated similarly (writing a roughly equal number of entries, with contributions from both groups that are decisive in the development of collective understanding) would be most successful. The last discussion note, What is the chemical to turn sand into glass?, is an example of a discussion that was stifled by several of the more powerful students in the class, and points to an important social issue to be addressed in collaborative learning. Following these examples of discussion notes, checklist data of all 54 discussion notes are presented to examine several additional features of the discussion notes.
4.6.1 'Is gas affected by gravity?'

This discussion note was created by Jerry (cf. section 4.3.4). It was active from December 13, 1994 until January 27, 1995. Two students with LFC and 6 with HFC achievement contributed to 11 entries after the problem statement. Since the discussion is quite short, every entry is shown below.

1. **Sean (HFC):** I think that gas is affected by gravity, why wouldn't it be? I think that gas just doesn't seem to be affected by gravity. Because of its extremely light weight, I think that gravity would not be able to keep gas from rising. I think that when the gas comes in contact with heat the heat would make the gas even lighter and it would rise faster.

2. **Albert (LFC):** I think that although gas is matter it has no weight whatsoever. Since there is no weight, there is no way that gravity can pull gas to the ground. Some gases can be turned into liquid, gasoline is turned into liquid so it can be used to power a car, liquids can be turned into gas. Water, when exposed to heat, vaporizes.

3. **Albert (LFC):** I need to understand why the matter of one substance can be changed into another type of matter, I would also like to know why only some forms of matter can be changed and others can't.

4. **Julius (HFC):** My theory is that gas is affected by gravity. I think that gas has no weight so you can't really see the effect.

5. **Matt (HFC):** My theory is that gas is affected by gravity. I think that everything that has matter is affected by gravity. I think this because everything that has matter has weight. Even air has weight, and that is why we can't just float in air forever. I think that gas, being very light, would be very weakly affected by gravity. I think that gas is affected, and therefore does not rush up into space that fast. I think that if it was not affected, the gas would zoom very fast straight into the sky, because it would be free to travel wherever it wanted.

6. **Andrew (LFC):** I think that gas is affected by gravity. Why not. I think that if it was not affected by gravity there would be no life on earth because all the air would rush to space.

7. **Jerry (HFC):** I think that gas is affected by gravity because gas is matter and anything that is matter is affected by gravity.

8. **Jerry (HFC):** Does gas weigh anything? What would happen if you had put some sort of gas in a box and the box could contain it. What would be the effect? Would the gas weigh anything inside the box because the gas is contained?

9. **Derrrik (HFC):** I think that gas is affected by gravity, everything is. But since gas is so light gravity cannot keep it on the ground. If gas wasn't affected by gravity then all the oxygen that we need to survive would float out into space and we would die.

10. **Emma (HFC):** I think that some gases are affected by gravity and some aren't. Good examples would be carbon dioxide which has no gravity since it is almost like air, or helium which does the opposite, by rising higher instead of falling with gravity's pull.

11. **Matt (HFC):** I need to understand what would happen if different gases had different weight? Would the affect of gravity be stronger? Weaker?
Sean (1) opened the discussion with a scientific theory (S); he did not justify this theory with argument. Albert challenged Sean’s theory in the first part of his entry (2). Implicit in his challenge is the idea that something can be matter, yet have no weight. What makes this part of his entry mixed framework is that his conclusion was correct (but circular) — if gas has no weight, then gravity has no effect on it. The second part of his note, beginning with the third sentence, is irrelevant to the problem at hand. In (3), Albert continued the discussion, letting the class know that he considered the problem resolved, and raised the second part of (2) to a higher level, that is, he made an INTU entry about it. Then Julius did something very fortunate. His entry (4) had the effect of bringing the discussion back to gravity. But he also put Sean’s correct thesis, that gas is affected by gravity, together with Albert’s incorrect one (2). This prepared the way for improving Sean’s theory; Matt rose to this task (5). Andrew (6) restated Matt’s thesis, but added an important insight. Jerry (7) added no content, but his note may have lent support to the theory developed so far. Why did he write entry 7? A plausible explanation is that he had learned something — that gas has weight. In Entry 8 he posed a very interesting ‘What if ...?’ question. His question was not taken up directly, although contributions by Derrik (9) and Emma (10) are relevant to it.

This is a very good discussion for several reasons. First, conceptually, it goes to the heart of the matter. Whether gases are matter and whether they have weight is a subtle question. Reflecting on Minstrell’s classroom work with high school physics students on gravity and air pressure (e.g., Minstrell 1992; Hunt & Minstrell, 1994), this discussion struck me as of a high level. Second, it provides a clear example of the important role mixed framework contributions can make to the discourse, most notably in entries 4 and 10. Third, the discussion shows several important features of knowledge building — especially stating opposition to a thesis, returning the discussion to its purpose when it has wandered off, and extending theories.

4.6.2 The role of learner achievement and note type

Two discussion notes are examined for the role the various note categories — commonsense, mixed framework, factual, scientific, and question — and learner achievement play in the success of discussions. The two notes examined are:
- **What is heat?** Created by a student with LFC achievement, whereafter 8 students with LFC and 5 with HFC achievement contributed 33 entries.
- **How is heat made?** Created by a student with HFC achievement, whereafter 3 students with LFC and 3 with HFC achievement levels contributed 16 entries.

Figures 4.2 and 4.3 show models of the relationships between the note entries for each discussion note. Shaded rectangles represent entries by students with HFC, and unshaded rectangles with LFC achievement levels, respectively. The arrows indicate proposed semiotic relationships between the note entries. In some cases, there were explicit clues in the text, such as ‘X, about your theory in (7) ...’. In other cases, for instance if an entry mentioned ‘molecule’ in a theory, and there had not been a molecular explanation in several entries, I looked for the last occurrence of ‘molecule’ in earlier entries. Sometimes one question was suggested by another. For simplicity, no transitive relationships are shown. In other cases, two or more entries were made by the same person in a short time interval (e.g., ten minutes), and it was difficult to decide if one entry had led to another; such entries are drawn very close to each other, without an arrow between them.

One reason the choice was made to examine the first note, **What is heat?**, was that it has a large number of entries and participants. But after a close reading of the note, it became apparent that it could be interpreted as four more or less semiotically independent lines of inquiry: (A) ‘What heat is’, (B) ‘How heat travels’, (C) ‘Ways of making heat’, and (D) ‘Heat and fire’ (Figure 4.2).

Of these, I judged (A) to be most successful. Its first theory was given by Matt: “The earth’s main supply of heat is the sun. Even on cloudy days the sun is scorching the earth with heat ...” (2). He followed this theory up himself with entry 3, thus beginning the second line of inquiry. His theory (2) was rated mixed framework, although it could have been commonsense. Entry 10 was rated factual (F), but it played an important role in the discussion because it introduced ‘energy’. Entry 16 served a mediating role, in which Emma made the following observation: “heat is heat is a type of energy, because it is neither a solid, liquid or gas. It can effect one of the three but it is not actually matter” (22).

---

14For example, in the case of the discussion note discussed in the previous section, entry 11 would have been judged to be suggested by entries 8 and 10.
Figure 4.2: Note entries of 'What is heat?' discussion note
Shaded rectangles represent entries by students with HFC and unshaded with LFC achievement levels, respectively. Closely-spaced rectangles without an arrow between them represent entries written by a single student, in the same session. Arrows indicate links between note entries, either physical or suggested by the content of the entries. A letter A or B next to a note entry indicates such a link with cluster A or B, respectively.
Another student with HFC achievement, also a girl, finally formulated a theory (23) that is essentially heat as energy of random motion (cf. section 4.1). Entries 7 and 13 were both written by Sean. The first of these introduced heat as molecular motion, but with an inappropriate interpretation: "I think that goosebumps are molecules that need more space to move faster." In note entry 13, he seems to have attempted to use authority to find more support for his theory (7):

... Francis Bacon (1561-1626) had a theory. His theory was that heat was motion, nothing else. He thought that the faster the particles in an object moved, the hotter the object was. He was also supported by several other leading scientists in his time. Bacon's theory is also the same as my theory in entry #7. Bacon's theory is now the definition of heat. (Sean, HFC)

As Figure 4.2 indicates, no one responded to this note.

In sum, the 'What is heat?' subdiscussion was effective: it did lead to substantial movement toward scientific explanations. But according to the above analysis, there were only two entries by students with LFC achievement that contributed to this success: the creation of the discussion note itself, and entry 16.

Subdiscussion B, 'How heat travels', involved only Matt and Sean. There were several good questions (e.g., "I need to understand if all rays from the sun travel at the same speed," entry 4 by Matt), but evidently these questions stumped Matt and Sean, and no one else touched them.

In C, ‘Ways of making heat’, things were worse. A student with LFC achievement provided factual information (11) and followed that up with the question, ‘How is energy converted to heat?’ That question certainly had possibilities for concept articulation, but this student does not appear to have read Sean’s notes (7, 13).

Finally, in D, ‘Heat and fire’, there were many questions by students with LFC achievement. Some examples are:

- I need to understand how a natural forest fire starts. (19)
- I need to understand how heat burns wood. (25)
- Does fire have weight? (26)
These appear to be questions that could have triggered concept articulation. Why was ‘Does gas have weight?’ so successful, compared with question 26? A hypothesis is: because there was insufficient working knowledge in the domain of fire — there were not enough ideas that could be useful as part of scientific explanations.

Two preliminary conclusions may be drawn from this discussion note:

- Access to adequate working knowledge was crucial, for both types of learners (compare subdiscussions B and D).
- There do not appear to be important differences between the learner achievement levels in the amount of questioning done.

Turning to the second discussion note (Figure 4.3), How is heat made?, it should be noted that the creator of the note was also its most dominant contributor; he wrote entries 1, 6-7, 11-14, and 16 — 47% of all entries. Moreover, Matt, Sean, Jerry, and Emma did not contribute to this discussion at all.

![Figure 4.3: Entries of 'How is heat made?' discussion note](image)

Figure 4.3: Entries of ‘How is heat made?’ discussion note

Shaded rectangles represent entries by students with HFC and unshaded with LFC achievement levels, respectively. Closely-spaced rectangles without an arrow between them represent entries written by a single student, in the same session. Arrows indicate links between note entries, either physical or suggested by the content of the entries. A letter A or B next to a note entry indicates such a link with cluster A or B, respectively. Note the dominance of entries by students with HFC achievement. Entries 11-14 were written by the creator of the discussion note, in the same session on CSILE.
One problem with this discussion note is that it contains many commonsense (CS) and factual (F) contributions: it is a discussion mainly about observations and surface features. For example:

- I think one way to make heat is by getting a magnifying glass and some wood. Now put the wood under some sun light then put the magnifying glass in the light and aim the light so that the light burns the wood. (CS, 6)
- Dark colours takes in the most heat and lite colours reflect heat and only take in a little heat. (F, 12)

The following analysis focuses on the branch leading up to entries 8 and 10: a series of four mixed frame-work notes. In entry 2, Derrik introduced friction as a force:

Heat comes from fire and one way to get fire is friction. If you strike a match against a rock or other rough, edged object it will cause friction which is the force caused when two objects rub against each other. When the match is struck it causes a spark and the spark catches fire providing heat. (HFC)

This entry was rated mixed framework because while it introduced friction, it left much to be desired as a scientific explanation. Neither the idea of a force of friction, nor how a fire starts from a spark was made explicit. The next entry (5) in this branch was written by an LFC learner. It was a salient entry in this discussion note, as it was the only one to which more than one entry subsequently made reference.

My theory is that friction makes heat. When atoms move they cause friction which makes heat. Heat is when atoms move very quickly. When you strike a match the atoms move quickly causing heat, matches have a special chemical called "phosphorus" which makes the atoms move faster to cause fire. ... (LFC)

This entry continued to discuss friction, relating it to the motion of atoms. The explanation offered seemed more effective for heat than for fire. Finally, one of the entries linked to this one is entry 8, also written by a students with LFC achievement:

... Friction is caused by pieces of the same substance {like wood or rock} rubbing together, wood will eventually burn, rock will eventually spark. The rubbing is movement, and any movement is kinetic energy, so therefore heat is made out of kinetic energy. Matches also prove this, a match needs to be struck on a rough surface to be lit, the spark is caused by friction. (LFC)
"Heat is made out of kinetic energy" indicates more the source of heat — the energy of the macroscopic motion of the match is transformed into heat — than an association of heat with a change in the kinetic energy of the random motion of the molecules.

In my opinion, although the students did not develop scientific understanding (at least by the standard employed in this study for judging their writing) there was in these three entries substantial concept articulation. Both the ideas of friction and heat became elaborated. This depended on students being willing to put up for debate ideas that had elements of scientific explanation, but still needed improvement. In this particular case, students with LFC achievement contributed to the collaborative process as well as students with HFC achievement.

4.6.3 The role of power

It is likely that the class saw Matt and Sean as authoritative figures. They wrote much, and their explanations tended to be the best available. Moreover, their teacher co-authored a few notes with each of them, which he did not do with many other students. The power this status gave to Matt and Sean was, on at least one occasion, quite destructive to knowledge building, as the following discussion note, What is the chemical to turn sand to glass?, illustrates.

1. Matt (HFC):
   Stephen,
   I do not think that you need to make this note because I do not think that you need a chemical to turn sand into glass.

2. Sean (HFC):
   Stephen,
   Why do you think that there is a chemical that is involved in turning sand into glass? I don’t think that you should make an INTU if you are not sure that there is such a problem. If you could find some proof that there was a chemical in sand that affects its turning into glass it would be justified.

3. Albert (LFC):
   Stephen
   As Matt said, you do not need a chemical to turn sand into glass. The only thing needed in turning sand to glass is extreme heat. I think that you should delete this note. ...

4. John (teacher):
   Stephen,
   This is an interesting problem. I’m not sure whether you need a chemical or not but the whole question of chemical change which may happen when you add heat to matter is worth studying in more depth. Perhaps you should change the problem into one such as, "How does the chemical composition of sand change when heat is added?" ...

If you knew that you did not need a chemical for making sand, then why did you make this note in the first place.

Sean's comment (2), while critical, left open a possibility for debate about whether or not Stephen's question was relevant to the class's inquiry. Stephen could have used Sean's contribution as a cue to refining his question. Matt and Albert's comments (1, 3), on the other hand, were less tentative, and suggest stonewalling to me (cf. Wegerif & Scrimshaw, 1997). John came to Stephen's rescue, endorsing the idea of chemical change as a topic "worth studying in more depth." However, his suggestion included a problemshift away from Stephen's conjecture that a chemical must be added to sand to make glass, a different issue.

This discussion note may have been a missed opportunity at knowledge building. One cannot help wonder what the outcome would have been if Sean or Matt, or even Jerry or Emma had decided to research this question. The class, especially students with LFC achievement, had a considerable interest in glass. The question could have provided a better balance between microscopic and phenomenological explanation — and it could have brought the subject of technology into the unit.

4.6.4 Checklist data
The checklist probed a number of knowledge building issues:

- The learner achievement level of the student who created the discussion note.
- The mix of contributions by students with LFC and HFC achievement levels, respectively.
- The number of contributions per discussion note.

Students with LFC and HFC achievement levels created approximately the same number of discussion notes: 29, compared with 25. Nor was the number of entries per note different between the learner achievement levels (Mann-Whitney test, \( U(N = 54) = 282.5, p = .16 \)). However, contribution mix, defined as:

\[
CM = \frac{n(HFC) - n(LFC)}{n(HFC) + n(LFC)}
\]
revealed an interesting effect. In this equation, \( n \) represents the number of note entries of a particular learner achievement level. A \( CM = -1 \) indicates that all the note entries were made by students with LFC achievement, and a value of +1 the converse. There was no significant difference between the absolute values of \( CM \) between learner achievement levels (\( p = .064 \)). While students with both levels of achievement contributed roughly an equal numbers of entries to notes created by students with LFC achievement, students with HFC achievement dominated notes created by another students with HFC achievement, see Figure 4.4. An examination of the frequency distribution of the number of notes with particular \( CM \) values revealed that there were a large number of discussion notes in which participation was limited to the learner achievement level of its creator. For six notes created by students with LFC and ten by students with HFC achievement levels this was the case.

![Figure 4.4: Percentage of entries contributed to 29 discussion notes created by students with Low Focus on Concepts achievement, and 25 created by students with High Focus on Concepts' achievement](image)

Notes were divided into groups with small and larger values of \( |CM| \) to explore the effect of this variable on the number of note entries. Although notes with \( |CM| < .25 \) were, on average, longer than other notes (\( M = 8.75 \) entries, versus 5.45, SDs 8.60 and 4.06); this difference was not significant (\( U(N = 54) = 187.5, p = .18 \), and
only had an effect size of .60. However, the 11 notes with $|CM| < .25$ included the first two notes discussed earlier in this section, and another two notes that also appeared to be successful. This suggests $CM$ may be an important variable in combination with another variable, such as writing output or subject matter knowledge, but not by itself, apart from the result shown in Figure 4.4.

Another result obtained from the checklist data was the following: for both groups, there was a large number of discussion notes with only a few entries, $M = 6.19$, $SD = 5.48$; 48.1% of discussion notes had fewer than 5 entries after the problem statement.

4.6.5 Summary

The note discussed in section 4.5.1, Is gas affected by gravity?, is striking because three knowledge building activities found in it were not found as often in other discussion notes: stating opposition to a thesis, re-focusing the discussion, and extending theories. Many discussions, not only in this database (e.g., Hewitt, 1997, 1998; van Aalst, 1997; van Aalst & de Jong, 1997), produce predominantly linear chains of entries, at the expense of “rise above” entries that consolidate two or more points of view about the same topic and introducing new lines of thought. The frequency of questioning did not appear to affect the success of discussion notes, but contribution mix did tend to. Is gas affected by gravity? is a clear example of a note with nearly equal participation rates by the two learner achievement levels, as is subdiscussion A in What is heat? In the last discussion, two other subdiscussions did not take off, despite many questions; both had strongly unequal participation (Figure 4.2), although lack of subject matter knowledge was a factor in both as well. One of the discussion notes studied indicated that discussions are sometimes stifled by overt actions by dominant contributors who discourage further work on the discussion note.

4.7 Questioning and Commenting

4.7.1 Questions

The findings for question form are shown in Figure 4.5. Thirteen students with LFC and eleven with HFC achievement wrote questions. An analysis of
homogeneity showed that the differences in the distributions of questions over the levels of the scale were significant, $\chi^2(2, N = 207) = 17.0, p < .001$; the largest effect was that students with HFC achievement asked more level 3 questions — 2.82 per student, compared with .54. Note the relatively large number of level 1 questions (37.7% of all questions); many of these questions, as explained in the method section, require further elaboration before they can be tested empirically.

The content of the questions using the following categories: (a) particles, heat and energy, (b) chemical change, and (c) physical change. These categories correspond to aspects of the conceptual model that was presented in Figure 4.1: particles, heat and energy are part of the theoretical framework in the top of the diagram, and chemical and physical change part of what the this framework should explain.

The distribution of questions over these categories was significantly different between the learner achievement levels, also by analysis of homogeneity, $\chi^2(2, N = 194) = 10.5, p < .01$. Students with HFC achievement wrote more questions about particles, heat and energy than students with LFC achievement, (50.4% and 28.1%, respectively), and fewer questions about chemical change (21.0% and 34.8%) and physical change (28.6% and 37.1%, respectively).
4.7.2 Comment notes

The class wrote 62 comment notes, equally divided between 10 students with LFC and 9 with HFC achievement levels. Analyses of proportions of the frequencies for the three dimensions of collaborative quality — collaborative intent, constructive feedback, and substantive contributions — revealed no significant differences between learning types. For the last dimension, students with HFC achievement were marginally higher, $z(62) = 1.615$, $p = .053$. Combining the data, 87.1% of comments showed evidence of collaborative intent, 58.1% constructive feedback, but only 33.8% substantive contributions. The distributions of comments over the categories of the content scale that was also used for the questions were nearly identical for both learner achievement levels ($\chi^2(2, N = 57) = .49$); combining the data here as well, 31.6% were about particles and energy, 28.1% about chemical change, and 40.4% about physical change. There was proportionally more commenting in T1 than there was in T2.

4.8 Quantitative Analysis of Note Ratings

Statistical analysis of the note ratings was difficult for several reasons: a substantial number of students wrote very little — 12 wrote in fewer than two note categories; Matt and Sean, on the other hand, each wrote as much as five times as many words as the class average. The analyses of the note classification into commonsense, mixed framework, factual, scientific, and question notes can be approached in two ways. In the first approach, one asks how students’ notes are distributed over the note categories — e.g., are a greater fraction of notes written by students with HFC achievement questions than is the case for students with LFC achievement? A difficulty with this approach is that the obtained measures would be invalid for students who wrote very little. One would put little stock in knowing that all of a student’s writing was in one note category if the student wrote only a few notes. A second approach is to ask the same questions about students’ contribution to the aggregate of all notes — do students with HFC achievement contribute proportionally more mixed framework to the database? This approach has the advantage that the difficulty of the first approach is avoided: if a student increases on mixed framework notes with this measure, we know that the student ended up contributing a greater percentage of such notes. But what we do not know
is if this is because the student began writing more notes, or if there was another change that led to a shift in the student's activity.

The approach taken is a two-step approach: all the note categories are first examined with the first approach, and the effect of writing output is then examined for one note category: mixed framework notes.

4.8.1 Proportions relative to individual students' writing output

Students who wrote very little would, if included in the analysis, enhance any differences in means between students with LFC and HFC achievement levels, and therefore bias the analysis toward the hypothesis of differences in the distribution of note proportions between the learner achievement levels. For this reason, the 12 students who wrote less than one note in each of two categories were excluded, leaving 9 students with LFC and 10 with HFC achievement in the analysis for T1, and 16 with LFC and 13 with HFC achievement for T2. With only 19 subjects in T1, statistically significant effects involving T1 would not be expected at the $\alpha = .05$ level. Nevertheless, one could consider effects not quite statistically significant at $p = .05$ in model building. The question why some students wrote so little is important but intractable with the available data. Six of the students removed were absent when the CTBS tests were administered, so that uneven attendance, for whatever reason, may have been part of the problem.

The five variables $X_1, \ldots, X_5$ were the percentages of a student's notes that were in categories 1, ..., 5, respectively. Shapiro-Wilk tests revealed violations of the normalcy assumption for all the T1 variables and several of the T2 variables ($ps < .03$). For this reason, nonparametric Mann-Whitney tests were used to evaluate the significance of differences between LFC and HFC achievement levels, checking that $\alpha_{\text{total}} \leq .05$ was maintained.\(^{15}\)

The data are presented in Figure 4.6. Note that in T2 the means shown are based on more subjects than in T1; the graph contrasts the distribution of means for students active in T1 with that of means for students active in T2. The effect of the ten students who were included only in the T2 analysis was to raise the means for factual

\(^{15}\) For $n$ comparisons, $\alpha_{\text{total}} = 1 - (1 - p)^n$; see Lomax (1992).
notes by 4.5% and for scientific notes by 3.9%, over and above any changes due to the 19 original students, and to lower the means for questions by 7.5%. This indicates that some of the students who started to write late (in T2) contributed more factual and scientific notes, and fewer questions, than the rest of the class. (One example of such a student is Emma, discussed in section 4.4.2).

Wilcoxon Signed Rank tests revealed that there were fewer commonsense notes in T2 than there were in T1 (10 negative ranks, 4 positive ranks, 5 ties, \( p < .02 \)), and more mixed framework notes (5 negative ranks, 11 positive ranks, 3 ties, \( p < .05 \)). The improvement in the proportion of scientific notes in Figure 4.6, larger than that for factual notes, was not significant for the original 19 students.

![Graph showing note distribution](image)

**Figure 4.6:** Mean percentage of a student's notes in each of five note categories, early and late in the unit. For a given period (T1 or T2) the bars add to 100%, so that the T1-T2 shifts indicate shifts in the nature of students' contribution to the database.

For T1, Wilcoxon Signed Rank tests showed that the proportion of notes that were questions was significantly higher than that of all other note categories, except factual notes (in all tests there were 1-4 negative ranks, 14-16 positive ranks, 1-5 ties; \( ps < .038, \alpha_{total} = .042 \)). There were no significant differences between learner achievement levels.
For T2, the proportion of commonsense notes was smaller than that of all other note categories (ps < .004), and that for questions it was greater than for all other categories except factual notes (ps between .008 and .02). The number of negative ranks, positive ranks, and ties in these tests ranged from 4-6, 19-22, and 1-6, respectively. The family-wise $\alpha$ for these contrasts was 0.05. T-tests led to similar conclusions about the significance of the contrasts, with $\alpha_{mult} = 0.015$. Mann-Whitney tests for differences between learner achievement levels revealed significant differences for scientific notes, $U (N = 29) = 47.5$ all $p = .012$, effect size .91, Ms 11.5% ($SD = 13.4\%$) for LFC and 24.8% ($SD = 15.7\%$) for HFC achievement levels, respectively.

One effect of interest concerns mixed framework notes. As Figure 4.7 shows, the increase with time in such notes was different for the two learner achievement levels. The difference for T1 between the learner achievement levels for this sample was not statistically significant ($p = .40$), but one may ask if a similar effect exists that is.

![Figure 4.7: Mean percentage of a student's notes that were Mixed Framework, for students with Low- and High Focus on Concepts achievement levels, early and late in the unit. Based on 9 students with LFC and 10 with HFC achievement levels in T1, and 16 and 13, respectively, in T2](image-url)
4.8.2 Proportions relative to the total writing output of the class

In this subsection, I examine how this effect is modified when the note proportions are measured relative to the total writing output of the class, rather than that of individual students. After all, if a high percentage of the notes written by a very productive writer like Matt are scientific, this will be more salient in the database than if a high percentage of the notes by an unproductive learner are scientific. If \( f_{ij} \) is the number of notes of category \( j \) written by learner \( i \), \( n_i \) the total number of notes written by that learner, and \( N \) the total number of notes written by all learners, we can introduce note proportions \( x_{ij} \) (learner) and \( X_{ij} \) (whole-class), and derive the following relationship between their averages over learners:

\[
x_{ij} = \frac{f_{ij}}{n_i}, \text{ and } X_{ij} = \frac{1}{N} \sum_{i=1}^{N} x_{ij} = \frac{n_i x_{ij}}{N}.
\]

(The bars represent averages over learners; the \( i \) subscript is absorbed by this averaging process.)

Same data set as the previous analysis (\( n = 19 \) in T1, \( n = 29 \) in T2): For the \( X \) (whole-class) proportions the statistical results are quite different from before, see Figure 4.8. As a percentage of their own writing, both types of learners, on average, tended to write more mixed framework notes in T2 than in T1. Measured relative to the total writing output of the class, students with HFC achievement wrote fewer such notes in T2, although still slightly more than students with LFC achievement. This indicates that the most productive students with HFC achievement — Sean, Matt, and Jerry — wrote fewer mixed framework notes in T2 than more typical students with HFC achievement did. The effect of this interaction is that the time effect for mixed framework notes that was significant for the \( x \) (learner) proportions no longer is for the \( X \) (whole-class) proportions. The difference between the learner achievement levels for \( X \) in T1 is still not significant, although with \( p = .16 \) it is closer to significance; its effect size has increased to .73.

All students (\( n = 31 \)). If writing mixed framework notes early is important for ending up writing scientific notes in T2, then not writing them early should cause a loss of potential for concept articulation. In an analysis of the \( X \) (whole-class) proportions with all 31 subjects, the difference between learner achievement levels was significant (Mann-Whitney test, \( U (N = 31) = 69.0, p = .031 \), as was the time
effect (Wilcoxon Sign-Rank test, 6 negative ranks, 19 positive ranks, 6 ties, $p = .034$). The effect size in T1 was .85.

![Graph](image)

Figure 4.8: The mean contribution individual students made to the collective body of Mixed Framework notes, expressed as a percentage of the total number of notes in the database, early and late in the unit

4.8.3 Summary

The analysis of the note ratings revealed individual differences concerning how a student’s notes were distributed over the note types. Striking was a large number of questions, by both groups of learners, and a large number of factual notes. The time effect of most interest in Figure 4.6 is that for mixed framework notes. Analysis of the temporal effect for mixed framework notes indicated that a greater percentage of notes written by students with HFC achievement in T1 were mixed framework than was the case for students with LFC achievement. This effect was not significant at $p < .05$, although the impact of this effect on the collective writing of the class was significant.
4.9 Discussion

4.9.1 Results

This study has provided an account of a student inquiry using CSILE, Knowledge Map, and with a focus on whole-class collaboration. In the case studies, many examples of advanced ideas by the Grade 5/6 students were shown. Some of the explanations students offered will not require much modification in future learning (e.g., explanations of convection), while some others do (e.g., that heat is a form of energy). It should be clear, however, that the students only worked in a relatively small part of the subject matter domain — there is little evidence that they were concerned with laws, even in areas in which they had much knowledge. There was evidence in the discussion notes that the students were capable of effective collaboration, and support for the proposal by Chan, Burtis, and Bereiter (1997) to maximize cognitive conflict in conjunction with sustained knowledge building effort (cf. the contribution mix variable). There was a large amount of questioning by students with LFC as well as HFC achievement levels. In part, this is expected of children in Grades 5/6 when they study science, but many of the questions were sophisticated in that they could lead to substantial advancement by the class if answered.

I found less evidence that students used epistemological terms than previous studies by Hewitt (1996) and Hakkarainen (1998) would suggest. As discussed in section 4.5.1, the term ‘experiment’ arose in an epistemological context in only a few notes. (In most cases, the term named an activity the class had done, without pointing to the role an experiment has in developing scientific knowledge.) Other epistemological terms, such as ‘hypothesis’, ‘evidence’, or even ‘test’ were not retrieved in substantial numbers in the analysis of terms in section 4.5.1. Of course, the terms ‘theory’ and ‘principle’ were not included in the analysis: the former was a thinking type, and the latter was part of John’s instruction concerning the kind of information students should be entering into the What the Class has Learned notes toward the end of the unit. Both were assumed to be “stop words.” The relatively small number of level 3 questions by both learner achievement levels (although students with HFC achievement wrote more) may also indicate a limited understanding of epistemology, particularly the importance of controlling for confounding factors by asking questions suitable for hypothesis testing. As Hakkarainen concluded from studying all CSILE notes generated by John’s classes
between 1989 and 1992, the vast majority of questions asked by students were what he called "explanation-seeking:" 'Why ...?', 'How ...?' and so on. He argued that such questions must be followed up by "subordinate questions." The point to be made here, however, is that a sufficiently large number of subordinate questions must be level 3 questions. On the other hand, it may be argued that the discourse showed more evidence for other epistemological aspects of knowledge building, such as the idea that knowledge is validated by a community. Students in this study, and apparently earlier classes of John as well, compared their own theories with whatever scientific information they could find in books and other resources. That kind of activity is an important part of doing science, but students will probably develop a more authentic view of science if their investigations include more experimentation and model-building than was the case in this unit.

The CTBS language scores and measures of writing output and quality were intended to contextualize the data on concept acquisition and knowledge building, but they revealed an interesting and complex story. Although there was a significant difference between the composite language CTBS (incoming) scores of the two learner achievement levels in Grade 6, this may be attributed to a larger number of students designated ESL, compared with Grade 5 students. Other measures relating to language skills showed no significant differences, except the number of sentences in a note and the number of notes written. Although the question should be studied more systematically in future research, it appears that basic skills did not contribute substantially to the quality of the contributions made to the database. There were no differences between students with LFC and HFC achievement levels with respect to the number of questions and comments written, although the latter wrote more level 3 questions. There was evidence that students with LFC focused more than students with HFC achievement levels on surface, or observational, features of the phenomena. This is apparent in the conceptions of Table 4.5, but also in some discussion notes that failed to develop, particularly notes about fire. Students with LFC achievement tended to work extensively rather than intensively, where the latter indicates going deeper and deeper into the topic (Hakkarainen, 1998). Part of the problem is failure to write expansive notes, but as mentioned before, whether this is due to a lack of knowledge or of writing skill cannot be determined from these data. The somewhat smaller set of terms used by students with LFC achievement (cf. Table 4.4) may also be a consequence of, or
factor contributing to, this lack of expansiveness of notes written by students with LFC achievement.

The lack of subject matter knowledge can be addressed partially in Phase 1 of the LKB model. The teacher did provide science activities at the beginning of the unit, and there is no reason to doubt that he assisted students, when needed, to ensure that they were valuable learning experiences. But they may have been ineffective in eradicating any disparities there may have been in the class with respect to subject matter knowledge. A pre-test of subject matter knowledge was not done for this unit, but there can be little doubt that such disparities would have been substantial in a multicultural class like this. It is also possible that there were disparities with respect to the process skills necessary for completing the experiments and finding and digesting information from books. All of these issues require learning, as opposed to knowledge building. Indeed, knowledge building by itself may increase disparities: although learners share the personal expertise they develop with the class, it is far more robust in them than in the rest of the class. There must be a balance between specialized knowledge and knowledge that everyone in the class is expected to have, and that is one of the key justifications for Phase 1. Learning to write more expansive notes falls under the general category of improving on the last experience of knowledge building, along with improving on questioning and commenting. Gains on these aspects of knowledge building can result from reflective activity. In writing a new note, the learner should keep past performance in mind, and not only work on contributing new knowledge to the database, but also on the rhetorical problems associated with communicating new knowledge effectively.

There are two knowledge restructuring processes that were important in this study. Suppose student A states that the boiling point of water is 100 °C. Student B might question this information and ask where this information was obtained or how the result would be affected by lowering the pressure of the surrounding air. Such questions, if answered, would enlarge the original knowledge claim by adding authentication information (e.g., "according to a book by X") or a qualification. The original claim becomes elaborated, but its core remains intact. In principle, factual and scientific notes can act as hosts of such knowledge restructuring, which may be called weak restructuring. Many of the questions asked by students in this study fell into this kind of activity. Frequently, a weak student can ask a sophisticated
question which he or she cannot answer, but another student can (Hakkarainen, 1998). This is also the process in which an expert can learn from the questions of an elementary student, as in the flight study (van Aalst, 1997) described in chapter 1, and in telementoring relationships (O’Neill, 1997).

Mixed framework notes, however, can act as hosts for a different kind of restructuring: strong restructuring. With these notes the content itself is problematic because of a tension within the note between two or more ways of understanding the topic of the note. The "ways of understanding" might be Western commonsense and scientific ideas, science and religion, or a similar combination. Here, the internal tension of the note must be resolved in some way. In some cases that means articulating the two viewpoints in the note and enlarging them by specifying similarities and differences between them, as was discussed in chapter 2. In other cases, it might mean forming alternate hypotheses and testing each. The mixed framework note is a archetypal of all knowledge objects — notes, collections of notes, etc. — that require multiple perspectives to be understood fully. In contrast with other note types used in this study, it is a better arena for concept articulation. The distinction between weak and strong restructuring is not one based on meaning or ontology, as Chi’s (1992) theory is; rather it is based on the presence of internal conflict that can generate knowledge building activity.

In this study, students with HFC achievement started writing a substantial number of mixed framework notes earlier than students with LFC achievement (Figure 4.7). By doing this, they were able to transform their knowledge by writing notes that had an internal tension, which needed to be resolved. In some cases the writing process was reflective, in several there was debate. Students with HFC achievement continued doing this in T2, but then they also wrote a more substantial number of scientific notes. Thus explicating their commonsense notions did not appear to impede conceptual change. Students tend to use scientific ideas rather than commonsense ones in science contexts (such as contributing to a database) if these are available, plausible, and understandable (Hakkarainen, 1998; Science Council of Canada, 1984; Strike & Posner, 1982, 1992). Thus one way to attempt to improve discourse in the CSILE database is to attempt to get all students contributing more mixed framework notes early. (This would also make the process of developing a consolidated research agenda more effective and democratic.) One way to attend to this issue is by direct teaching, although this is likely to be difficult:
many students will resist typing notes that are, in a sense, unfinished. A teaching strategy worth exploring is to return to factual notes to develop them further, since students with LFC achievement wrote a substantial number of factual notes. Factual notes can be viewed as underdeveloped notes: they have scientific content, but lack surrounding material that indicates understanding. A strategy to expand on such notes, adding explanations, attempting to integrate its information with existing knowledge, and so on, should introduce internal tension and produce mixed framework notes and some questions and scientific notes. This strategy would address both enlarging the visibility of mixed framework notes in the database and the problem of learning to write more expressive notes.

Resolving conflict in a mixed framework note is prototypical of knowledge integration and differentiation at other levels, such as in a discussion note. Learning to resolve conflict arising in a single note should prepare students for the more difficult task of resolving conflict in larger note structures. In this study, discussion notes were analyzed at the level of entries, and mixed framework notes seemed to mediate concept articulation in some cases (cf. the note depicted in Figure 4.3). Future research with a smaller unit of analysis — e.g., propositions within a small number of notes clustered around a mixed framework note — should be able to shed additional light on the question how successful discussions get started. Such studies should be conceptualized in a socio-cultural framework (e.g., Cobb, 1994; Roth, 1998).

Finally, the analysis of discussion notes in this study revealed that there are important issues of power arising from the social nature of collaboration within such notes. One note analyzed indicated that promising lines of thought are stonewalled by dominant contributors. There also was evidence of domination by students with HFC achievement of discussion notes created by a student with the same achievement level.

4.9.2 Method
In retrospect, some remarks on aspects of the method used in this study are warranted. First, as already mentioned, the data corpus did not allow causal inferences about conceptual change — for individuals or for the class. A detailed ethnographic study in which the roles of CSILE as a mediational tool, group
interactions, material artifacts are considered, and teacher actions are given more consideration can provide a fuller account of learning in CSILE classrooms. However, the study did raise a number of issues for further study, and was capable of guiding model development, as I will discuss in more detail in chapter 6. In future studies, database transcripts covering more time (e.g., several units) or many students (e.g., several classes doing the same unit) may demonstrate important learning effects. The finding that some students began writing late in the unit can induce changes in classroom practices that are aimed at overcoming this problem, and continual monitoring of writing output over several units would demonstrate any resulting improvements. Similarly, the class could attempt to write more MF notes early and study any gains on depth-of-explanation scores associated with this, again over several units. Thus, analyses of the database transcript, I conjecture, can become useful in the hands of teachers and students who want to make incremental improvements to practice.

A second issue concerns the note ratings: notes usually contained several ideas. I chose to code at the note level for two reasons: (a) when exploratory ratings were done on the basis of separate ideas, a difficulty was a substantial variation in how ideas were expressed, some students using only a sentence, while others needed several sentences, some providing examples, and so on; and (b) presenting smaller units (i.e., ideas) would have been a step further toward decontextualizing the discourse element being rated. That choice, however, introduced the difficulty of choosing a note segment as “most salient.” In this regard, the holistic ratings were on more solid ground because they captured the scientific merit of a student’s writing, “in the balance.”

A third issue is that the reliability of the holistic ratings could have been improved using the idea of coherence as the consistency of meaning between sentences. To see how this might work, Allan’s portfolio (Table 4.2) is re-evaluated. The first entry contains five sentences:

S1. The water molecules which make up an ice cube have less kinetic energy than those in water.

S2. [kinetic energy] is a kind of energy that is in everything that moves.

S3. The faster the object is moving the more kinetic energy it has.

S4. Sometimes the hot water molecules bump into the edge of the ice cube and transfer some kinetic energy to it.
The transfer of kinetic energy continues until the energy is shared between the mixture.

S1-S5 are true; S3 elaborates on S2; S4 and S5 elaborate on the theory S1-S3; S5 elaborates on S4. Overall, then, there is considerable support (including elaboration) for the claim that the first entry supports the hypothesis of a scientific explanation. A similar analysis can be applied to the second and third entry in the portfolio. Thus, the first three entries make a strong case for rating this portfolio scientific, but we must look for refuting evidence. Let's therefore look at the last entry, which consists of two sentences:

S6. When two sticks are rubbed together they make fire because the molecules in the stick are quite dry.
S7. When you rub them [dry sticks] they get very hot and create fire.

The second part of S6, taken literally, is false: dryness is a "macroscopic property" (Driver, Squires, Rushworth, & Wood-Robinson, 1994), that cannot be applied to the idea molecule. I would count the last entry as evidence refuting the hypothesis that the portfolio should be rated scientific, but it should be noted that S7 is true, so the refuting evidence is quite weak. Considering S1-S7 together, the portfolio would have to be rated scientific. This rating is the same as the one reported earlier in the chapter, but now there are within portfolio degrees of freedom. (Previously, only a single rating was done for a given portfolio by each of the independent raters.) Taking this analysis a step further, a rating of the aggregate of a student's writing could be reported together with an estimate of the quality of the rating based on explanative coherence (Thagard, 1992; Thagard & Verbeurgt, 1998).

4.10 Concluding Remarks

This chapter has provided an account of one CSILE database. We have seen that the unit mapped relatively well onto the LKB model, although there was room for improvement, particularly in getting all students to write productively and improving the cohesiveness of the discourse. There are implications for the role of the teacher and a need for further research into social dimensions of knowledge building to be discussed. But first a second case study is presented in chapter 5.
Although that case has a degree of overlap with the one presented here, it is not a replication study, and should be seen as one intended to enlarge the empirical base of knowledge building issues to be discussed in the final chapter.
CHAPTER 5

STUDY 2: A VIEW FROM CURRICULUM COVERAGE

The idea of a subject matter specialist building views of the subject matter while the unit was in progress, seemed promising. I hoped that such activity could improve the quality of the knowledge advancements of the class. In addition, differences in how John and Paul approached teaching with CSILE, and disparities with respect to success at knowledge building by their classes (cf. chapter 3), were known by the CSILE group, so that Paul's class seemed to provide a good teaching situation within which to explore this idea. An opportunity to explore these ideas presented itself in December 1996, when Paul was ready to begin a unit on force and energy, which would last until the end of April, 1997.1 This chapter reports on a study of that unit.

The intervention was framed by the AAAS (1993) benchmarks for science (Grades 5 and 8), and the objectives of a science kit on force and energy prepared by the local school board. A curriculum-based approach could complement the one used in study 1. One could argue that providing scientifically correct explanations is not a high priority in Grade 5, that all that is required is some acceptable measure of progress toward such a standard. A study of students' performance at knowledge building framed by curriculum documents can provide an alternate point of view. The weight of the study falls on an analysis of the learners' end-of-unit knowledge claims, using instruments that do not specifically require CSILE. The database is considered a work-in-progress, although a number of analyses that are intended to forge overlap with study 1 make use of it. Latour (1987) has argued that to understand science requires that one studies how scientific artifacts are created and used, rather than the artifacts themselves, and Roth (1998) that we should therefore assess process in school science, rather than only end products. Although I agree with that position,

1Prior to the commencement of the unit, I explained to Paul that I was interested in developing a model (not yet the LKB model) that could support improvement of knowledge building practice; I explained that the study could become a negative account of knowledge building that revealed problems. Although Paul agreed to a study based on this premise, we did not discuss his level of commitment to knowledge building in detail.
what makes processes interesting is the knowledge advancements they produce, so that a study of processes must be contextualized by a study of these. Thus, the present study can serve to inform studies into specific knowledge building processes, and hence contribute to the goal of continual improvement. In addition, the analyses I use in this study can be applied to smaller time intervals (e.g., a week's work at a time), thereby providing a better trace of knowledge building processes. Finally, the analyses used provide a way to explore aspects of assessment that are independent of using computers to support knowledge building, which should therefore be applicable to such learning environments. Like study 1, this is a qualitative and exploratory study; and although I intervened in Paul's classroom practices, it is not an experimental study. Rather, the subject matter and the author's actions in the class provide a context within which to examine learning and knowledge building in a different context from study 1. To sum up, this study is concerned with:

- Obtaining first-hand experience in a classroom with a teacher who used a traditional learning model, rather than one based on intentional learning.
- Exploring additional probes of concept acquisition and knowledge building.
- To better understand the role of the teacher, and the potential role of outside specialists, to support the class's knowledge building activities.

### 5.1 Background

The purpose of this section is similar to that of section 4.1 — to develop the necessary conceptual background for the study. The teaching context of the study will be described in section 5.2.

#### 5.1.1 The LKB model, assessment, and evaluation

Although children are likely to have common learning experiences in Phase 1 to acquire a working knowledge from which to conceive and plan their research, in Phases 2 and 3 they specialize and are expected to share their knowledge with the rest of the class. This is quite unlike traditional schooling, where students with similar achievement tend to be engaged with the same
learning activities (Bereiter & Scardamalia, 1993; Scardamalia, Bereiter, & Lamon, 1994). Brown and Campione (1994) have called this form of specialization in knowledge building classrooms 'majoring'. A strength of majoring is that the student comes to hold a unique position in the class, because of the knowledge only he or she can bring to the collective goals the class is attempting to meet. It matters to the whole class that each student learns. An example we have already seen is that when students do not contribute mixed framework notes, the whole class loses opportunities to improve the ideas such notes might have described.

The LKB model proposes two qualitatively different educational experiences, learning and knowledge building, and each can be assessed and evaluated independently. The selection of topics for study, of course, should be driven by the locally enforced curriculum. The LKB model need not be used for every topic, but when it is, it must be ensured that all children have covered a common set of subtopics. This can be achieved in Phase 1, where traditional methods of standardized testing can be used for assessing learning. For knowledge building, however, a more holistic approach to assessment is needed.

Recently published national frameworks for curriculum development in science (AAAS, 1993; Canadian Council of Ministers of Education, 1997; DES/WO, 1995; NAEP, 1996) have been influenced by research and development in history, philosophy, and sociology of science (Duschl, 1990; Duschl, Hamilton, & Grandy, 1992; Matthews, 1991, 1994) and Science-Technology-Society education (International Journal of Science Education, 10 (4), 1988 [special issue]; Solomon & Aikenhead, 1994; Ziman, 1980). Standardized tests based on such curriculum frameworks, provided they also include adequate attention to intentional learning, could be used to compare learning resulting from Phase 1 activities across educational contexts. It is important to ensure, however, that such testing will not undermine the goals of the LKB model. The 1988 Education Act called for standard assessment at ages 7, 11, 14, to validate teacher assessment, but the tests became "more comprehensive, rather than sampling the curriculum thinly

---

2I will make a distinction between assessment, the measuring of student performance relative to an educational program delivered by an educational jurisdiction, and evaluation, which is concerned with the performance of such programs themselves, and thus is concerned with teachers, school administration, and so on.
to moderate teacher assessment" (Brown, McCallum, Taggart, & Gipps, 1997, p. 272). Standardized tests are used in evaluation to make statistically reliable comparisons between educational jurisdictions, and focus on what can be measured and analyzed this way. They often leave out sufficient information for interpreting the results, such as the knowledge the children brought to learning the content sampled and the nature of the learning experiences. As Gipps (1989) has pointed out, success on a standardized test often shows only that a group has learned to take tests of a particular kind, or the selection of students by the educational jurisdiction. Moreover, a clear connection between the test items and the ability to use scientific knowledge where it is required in everyday life is not usually demonstrated (de Caprariis, 1994).

With the LKB model, standardized tests may be used to obtain a reliable indicator of the breadth of knowledge acquired by students — at the level of familiarity. Such tests, however, should not be the kind that sample a large curriculum area thinly, as is often the case; they should sample the specific subject matter studied by the children, with an aim of ensuring that the children are familiar with the most central concepts, and are suitably prepared for the next unit or grade, even in areas of the unit which they have not researched themselves. All students, for instance, who have completed John’s unit on heat and matter, may be expected to know that the boiling point of water is 100°C at normal atmospheric pressure, that the temperature of water while it is boiling is constant, that adding salt to ice lowers its melting point, and so on. However, no collection of such declarative knowledge, no matter how comprehensive, can be a reliable indicator of understanding. Standardized tests should be complemented by holistic assessment, which, as Hodson has put it, should “attempt to construct a ‘map’ of students’ actual knowledge within a particular domain. This requires students to volunteer information about what they do not know, or are unsure about ...” (1993, p. 132).

In sum, having had educational experiences aimed at covering a science topic using the LKB model, students’ knowledge should have (at least) four dimensions:

1. **Familiarity** with a sufficiently large area of the content of science, with access to more detailed knowledge. **Access** means not only the capacity to find the necessary
information, but also the capacity to relate it to scientific knowledge the learner already has, and the capacity to apply it to the problems the learner is engaged with.

2. Appreciation of the nature of science. This includes the relationship of science to other ways of knowing, the purpose of experiments and theories in science, and the limits of scientific knowledge.

3. Appreciation of ways science benefits/harms the environment and society.

4. Detailed knowledge of a relatively small number of science topics.

The first of these is to be met primarily by Phase 1 and the last three by Phases 2 and 3 of the LKB model.

5.1.2 Three probes of students' knowledge

In the present study, three probes were used to examine the learners' knowledge about force and energy after their research. They could in principle be used in learning situations that do not depend on CSILE, and are intended to generate an account that can balance that of study 1, which was based entirely on the CSILE transcript. The first probe is a paper-and-pencil test of 10 free-response items. The test is based on the Grade 5 AAAS (1993) benchmarks and the locally prescribed curriculum for force and energy (Appendix B); it was designed in collaboration with John and Paul. It probes subject matter knowledge in the domain of the unit (force and energy) as well as aspects of knowledge building. Several items on the test were designed to probe specific conceptual issues such as the use of impetus explanations (McCloskey, 1983).

Two other probes are conclusion notes and semi-structured interviews. The idea for the conclusion notes came from the students themselves, and was an effort to write a book on force and energy by reviewing their best notes, and writing summaries of their best work. The idea behind the interviews was that one indicator of knowledge after an extended period of research should be that students could hold up their end of a critical conversation about the topic they had researched. And one effect of sustained intentional learning should be that students rework their knowledge in a situation that requires it. Thus, while the

---

3 AAAS (1993) does not include benchmarks for Grades 6 and 7. Therefore, Grade 6 students were also given a test based on the Grade 8 benchmarks, on the assumption that some Grade 6 students would be able to meet these benchmarks. However, this test turned out to be too difficult, and is not discussed in the dissertation.
interviews began with general questions about what the students had researched and what they thought they had learned, they quickly became conversations in which the students and I collaboratively attempted to confront an issue (raised by me) that was in the domain of what they had researched. That is, they probed students' ability to engage in the process of discourse about science (Roth & Bowen, 1995).

5.1.3 The subject matter

In study 1 we found that the children's investigation was at the level of individual concepts, not at the level of developing from such concepts an integrative framework. That was also the case with force and energy, the subject of the unit studied in this chapter. As we shall see, much of the student inquiry of this study was concerned what force and energy are, not with such things as the laws of motion.

I briefly describe two frameworks physicists use to analyze problems in classical mechanics. Further details than provided here are given by Arons (1990). The first approach is to develop a description of an object's motion in terms of its position, velocity, and acceleration; the last of these concepts is related to the mass of the object and the net force acting on it. Given position, velocity, and the acceleration (or the mass and net force) at a particular time, and the time-dependence of the net force, the future motion of the object is assumed to be fully determined. (This is not always a valid assumption, e.g., in the case of chaotic dynamics, but it works for most of the situations students typically encounter in school science.) Newton's three laws of motion play a central role in this framework. Newton's first law makes a distinction between accelerated and non-accelerated motion (measured relative to a non-accelerated reference frame), and embodies Galilean relativity — rest and motion at constant velocity are indistinguishable. It also makes the point that no net force is required to keep an object moving at constant velocity, in contrast with Aristotelian physics. Newton's second law provides a quantitative relationship between an object's acceleration, mass, and the net force acting on it. An important step in analyzing problems in this framework is to determine the net force, acting on an object. Newton's third law is important when interactions with other objects must be taken into account, for instance, when a person leans against a wall.
A considerable difficulty for students with this framework is that the position and velocity of an object change *continuously*, which means that the equations that explicate Newton's second law are, in general, difficult to solve (they require calculus as well as vector algebra). This difficulty is avoided to some extent by the second framework, which focuses on holistic changes in the energy of a particle as a result of its motion. Any changes in the energy are accounted for in terms of the work done *on* the object (or system) by external forces. For some types of force — like the force of gravity — work is independent of the path of the motion, and with such forces one associates *potential energy*. In this framework, then, one must solve a single equation of the form \( \Delta E = W \), where \( \Delta E \) is the change of the energy, which does *not* depend on the path of the motion, and \( W \) is the work done by those types of force for which the work does depend on the path (e.g., frictional forces). The utility of this approach is that \( W \) can often be assumed to be zero, in which case one speaks of the conservation of (mechanical) energy, or it can be modeled in some simple way (e.g., in the case of sliding friction). The role of \( W \) in \( \Delta E = W \) is analogous to that of \( Q \) in the first law of thermodynamics (chapter 4). The idea of a *conservation law* (i.e., with \( W = 0 \)) in this approach is prototypical of a deep principle in physics. A second conservation law, used in conjunction with the law of conservation of energy, is the law of conservation of momentum (mass times velocity): in the absence of an unbalanced external force, the total momentum is constant.

Both of these frameworks are complex, and one would not expect students in Grades 5/6 to elaborate either of them fully. One would expect the most important benefit of the student inquiry to be in laying the groundwork for both conceptual frameworks by building an empirical basis for it — building up what it can explain. Students might make progress on a number of issues. For example:

- Realizing that scientific discourse has a register of its own. Terms like force, energy, and work have distinct meanings in physics, although they are only weakly differentiated in commonsense usage.

- Motion as continually changing versus holistically changing. Children tend make statements like "it went that way!", without considering the details of how the motion is changing (Dykstra & Sweet, 1994).

- Inanimate objects like tables can exert forces (Klaassen & Lijnse, 1996; Minstrell, 1982).
5.2 Method

5.2.1 Subjects
The 30 students of Paul's combined Grade 5/6 class of 1996/97 participated in this study; 18 students (60%) had used CSILE for at least one school year, the remainder having had exposure to it in earlier grades. The school was a middle-class, urban elementary school.

5.2.2 Materials and procedure
Science kits: Students worked for approximately two weeks in eight small groups to complete a series of seven activities in a kit on force and energy provided by the local school board. The activities were:

1. Measuring the force needed to move an object.
2. Exploring friction with different masses but the same surface,
3. Exploring friction for the same mass but different surfaces,
4. Finding ways to change the force of friction between objects,
5. Exploring jet propulsion,
6. Making a type of rocket,
7. Making a vehicle that operates by wind force.

Several students did not complete Activity 5 due to problems with the balloons used. Instructions provided with the science kits included the "science vocabulary" that students were expected to have learned after completing the activities (Table 5.1).

CSILE implementation: Paul used the same version of CSILE as John did in study 1, but without Knowledge Map. (Thus, students could access a note only via a note list produced by a search, or via links from an already open note to another note.) Students worked in groups of 3-5, in contrast with study 1, where students made individual contributions to whole-class discussion notes.
Table 5.1: Science vocabulary based on science kit from local school board

<table>
<thead>
<tr>
<th>term</th>
<th>constant</th>
<th>newton</th>
<th>f. of gravity</th>
</tr>
</thead>
<tbody>
<tr>
<td>force</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friction</td>
<td>lever</td>
<td>resistance</td>
<td>fulcrum</td>
</tr>
<tr>
<td>gravity</td>
<td>simple machine</td>
<td>average</td>
<td>mass</td>
</tr>
<tr>
<td>pressure</td>
<td>fair test</td>
<td></td>
<td>prediction</td>
</tr>
<tr>
<td>rocket</td>
<td>variable</td>
<td></td>
<td>propulsion</td>
</tr>
</tbody>
</table>

Introduction to student work with CSILE: Following the work with the kits, students were introduced to the work on CSILE, which lasted 13 weeks, in a class meeting (Dec. 5, 1996). Paul had written on Bristol boards questions on force and energy produced by his students of a few years earlier, and used them to frame a discussion about questions that would be good for discussion in CSILE. Students argued that 'What is ...?' questions are not very good, that questions that have multiple answers are better. At the end of the class meeting the children were instructed by Paul to select, in their usual groups, one of the following six themes as their specialization:

A. Force.
B. Energy.
C. Friction.
D. Levers.
E. Rockets.
F. Energy production/generation.

The themes were matched to the experiments the students had already done (two groups did theme B and two did theme E). Of the eight groups, four (B1-B4) consisted of boys and four (G1-G4) of girls. Each group was given a sheet with some suggested questions, was asked to generate additional questions, and then to enter their questions in a group discussion note. Anticipating Phase 2 of the LKB model, we imagined that the children would spend several days entering their questions (i.e., statements of interest for research), whereafter other students in the class could look at them and comment on them. The goal was to generate a sufficiently large pool of ideas to allow each group to develop a research agenda. This plan was not followed closely. Although most groups did enter their questions, they did not peer review or prioritize them. Instead, most groups began entering their theories almost immediately.\(^4\) Moreover, the students initially allowed only members of their own group to be co-authors of

\(^4\)Two groups (G2, G3) had trouble with basic CSILE operations, such as creating a co-authored discussion note.
discussion notes. This meant that other students in the class could only collaborate with a group in CSILE by writing comment notes, a procedure with a high procedural overhead, see section 3.1.2.2.

Duration of student work on CSILE: The first notes appeared on December 16, 1996; the class continued to write for a total of 13 weeks. No notes were written between December 17 and approximately January 10, and there was a two-week vacation in March, 1997. In that time, Paul did not contribute any notes, although he said he had read some.

Monitoring of work on CSILE: I accessed the database remotely when not on site, using Tele-CSILE (a version of the Macintosh CSILE client with remote access capabilities) to contribute to the database and send weekly email reports to Paul based on what I saw going on in the database. One example of such a report is the following, sent on January 26, 1997.

- There is a very nice discussion note on 'what is energy?' by Nathan et al. that gets into the origin of energy (big bang, etc.). ... [B1]
- There is some progress w.r.t. 'How does a rocket work?', but I think it needs to be brought together. ... It might be worth to remind students, as they begin to enter theories, to keep the vocabulary list in mind as they write their theories. For example, the word 'force' is lacking from much of the discussion about rockets. ...
- Some claims are being made by students without justification. It might be worth pointing out that stating the source is important so that we can make sense of the claim. ...
- Based on author searches, the following students in the database have not entered anything since we began the unit: ...

Paul used such notes in his teaching between the (approximately) weekly classroom visits. For example, after reading the above note, he talked to the students about justifying claims; several students subsequently added references to their notes.

During the Christmas vacation, I used the vocabulary of Table 5.1 to search the AAAS (1993) benchmarks, and entered a series of notes with the search results, including an organizer note with links to them:

---

5This appears to have been established practice in this class and is related to Paul's dislike of whole-class discussion notes (interview, Feb. 20, 1997).
... I have [also] added the vocabulary the ... school board thinks you should be learning as KEY WORDS. When you make a new text note or of you add to a discussion note, you should add these keywords to your note.

You can do this by going to the TOOLS menu and go to SUGGEST KEYWORDS. CSILE will then check if any of the keywords that are already in the database are also in your note. If CSILE finds some, it will ask you if you want to make those words keywords in your note. ... (Jan. 6, 1997)

In a subsequent class meeting, I discussed this procedure with the class, while two students experienced with CSILE demonstrated it. Paul and I hoped that by doing keyword searches later in the unit, the students' notes could be brought together with the AAAS benchmarks, so that students could see how their work compared with what the AAAS and the school board expected them to know.

The benchmarks for the end of Grade 5 retrieved by the vocabulary suggested students' knowledge should include the following facts:

1. Moving air and water can be used to turn machines.

2. The sun is the main source of energy and people use it in various ways, and the energy from fossil fuels comes from the sun indirectly.

3. Some energy sources cost less than others and some cause less pollution than others.

4. People try to conserve energy in order to slow down the depletion of energy resources and/or to save money.

5. Changes in speed or direction are caused by forces. The greater the force, the greater the change in motion will be. The more massive the object is, the less effect a given force will have.

6. The earth's gravity pulls any object toward it without touching it.

7. Without touching them, a magnet pulls on all things made of iron and either pushes or pulls on other magnets.

Class meetings: Paul conducted frequent class meetings, as discussed in chapter 3. Two of them were turning points in the unit. On February 17, 1997, Paul and I discussed aspects of knowledge building with the class — learning in small groups, and sharing that knowledge with others, and attempting to advance more in the direction of the AAAS benchmarks. We suggested that students: (a) put keywords on their best notes, and (b) conduct searches to bring them together with the notes based on the AAAS benchmarks. We wanted the
students to move in the direction of making a 'View' of what they had learned.6 A month later (March 17, 1997), we talked about finishing the unit. Paul advised students to finish editing their notes. He met with a student from each group to discuss possible ways of finishing the unit. They decided that the class would write a book on what they had learned about force and energy.

5.2.3 Data and measures

The component analyses of the studies are designed after their counterparts in study 1, see Table 5.2. The most important differences between the two studies are that in the present study: (a) the unit of analysis is in most cases the group on CSILE rather than the individual student, and (b) the case studies and the analysis of concept articulation are based on non-CSILE texts obtained from interviews and a paper-and-pencil test, both administered at the end of the unit, rather than on the CSILE transcript. I chose to focus on groups, especially for the interviews, to mirror as closely as possible, the social conditions under which students had developed their knowledge. Group results for the interviews and the test are presented with individual results for the CTBS, variables describing the volume of writing output, and Flesch-Kincaid reading levels, to contextualize the case studies. The analysis of concepts now consists of two stages: (a) presentation of specific results of the written test that were not reported earlier in the chapter, and (b) an analysis based on a science vocabulary, as in study 1. The role and character of the discussion notes are quite different in this study, as a result of the way the work on CSILE was organized in this classroom; thus an analysis such as based on contribution mix in study 1 would not be appropriate.

---

6In Knowledge Forum™, a view is an object that holds a group of notes that have something in common (e.g., notes by the same author or on the same topic).
<table>
<thead>
<tr>
<th>Analysis</th>
<th>Issues probed/Method</th>
<th>Variables</th>
<th>Reliability measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic literacy skills</td>
<td>- Contextual factors that may influence success with the unit</td>
<td>- CTBS language scores</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- No. of words</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- No. of notes</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Flesh-Kincaid reading levels</td>
<td>-</td>
</tr>
<tr>
<td>Overall achievement</td>
<td><strong>Test:</strong></td>
<td><strong>Test:</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Conclusion note:</strong></td>
<td><strong>Structure</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Interview:</strong></td>
<td><strong>Knowledge claims</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td>Case studies</td>
<td>- Active participation in a critical conversation</td>
<td>- Topical specialization</td>
<td>- Narrative account</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Content</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Knowledge building</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Nature of science</td>
<td></td>
</tr>
<tr>
<td>Test item analysis</td>
<td>- Specific responses on test</td>
<td>- Knowledge claims</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Next discussion note entry</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Thinking-Type-in-Use</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Logical argumentation</td>
<td>-</td>
</tr>
<tr>
<td>Concept articulation</td>
<td>- Acquisition of scientific vocabulary</td>
<td>- Term frequency</td>
<td>- Random selection of notes from CSILE transcript</td>
</tr>
<tr>
<td></td>
<td>- Articulation of concepts that use these terms</td>
<td>- Sentences</td>
<td>- Narrative account</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Term density</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Term working</td>
<td></td>
</tr>
<tr>
<td>Knowledge building activities</td>
<td><strong>Discussion notes:</strong></td>
<td><strong>Discussion length</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td>- Role of “mixed framework” notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Collaboration</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Questions:</strong></td>
<td><strong>Frequency</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td>- Comparison with study 1</td>
<td><strong>Question form</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Topical content</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Comments:</strong></td>
<td><strong>Frequency</strong></td>
<td><strong>-</strong></td>
</tr>
<tr>
<td></td>
<td>- Comparison with study 1</td>
<td><strong>Collaborative quality</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Topical content</strong></td>
<td></td>
</tr>
</tbody>
</table>

The organization of the study is as follows:

- **Basic literacy skills.** This analysis is analogous to its counterpart in study 1; see section 4.2.3.1 for the details.
• **Overall achievement.** Quantitative measures based on the test, the conclusion notes, and the interviews are reported to provide a general impression of the class's achievement before discussing the case studies.

• **Case studies.** Two cases based mainly on the interview data are presented to illustrate two types of performance observed in this activity. The focus is not the same as in study 1, since the data comes from an interaction of no more than 35 minutes and cannot probe such things as the development of a research program. Thus the focus of the cases is the end state — science content, knowledge building, and the nature of science — of the group.

• **Concept articulation.** A list of terms based on a glossary taken from a physics textbook (Hewitt, 1987) and the vocabulary provided by the local school board are used to retrieve notes from the CSILE transcript. Two aspects of the use of terms from the glossary is examined for the retrieved notes: (a) the number of notes in which each term is used, and (b) how the terms are used to convey ideas.

• **Test item analysis.** A selection of responses to test questions are discussed in detail. The selection of issues is based on the need to illustrate claims made in chapter 2 regarding the interpretation of student statements, and the need to examine further the findings of study 1.

• **Knowledge building activities.** These are based on the CSILE transcript and are as analogous to their counterparts in study 1 as possible. They are intended to provide a basis for making a contrast between the two classes. See section 4.2.3.1.

The variables are discussed next (Table 5.2).

**5.2.3.1 Test scores**

The test consisted of 10 tasks requiring short written answers; it was designed by the author, with feedback from John and Paul. Students were expected to take approximately 30 minutes, but the teachers were asked to allow more time if students needed it; most students took 45 to 60 minutes. Student responses were rated by the author. The test, the scale used for scoring it, and examples of the scorings for all items, are given in Appendix B. For Question 1-C, relating to evidence students would accept that living forms at one time existed on the moon, a distinction had to be made between evidence for current life and evidence for life in the past. The aim of Question 4-A was to examine the extent to which students differentiated among force, energy, and power.

**5.2.3.2 Conclusion note structure.**

Scale description: (a) Were the original question and the reason why the group chose to study that particular question stated? (b) Were sources identified? (c) Did the group report what science they learned? (d) Did the group report
other learning, such as insights into the nature of science or into their learning processes. And (e) in what degree was the note a collaborative effort? Each criterion contributed 2 marks, resulting in a 1-10 scale. (See Appendix C for the scale.) The scale did not take into consideration the quality of the knowledge claims in (c), since this was examined separately with the scale described in the next subsection. (According to Paul, students had prior experience with the genre used, but this scale was developed during data analysis, and was not communicated to the students.)

Reliability: Two independent raters rated all eight conclusion notes and agreed on 90% of the 40 decisions.

Example: The note by B1 (Appendix C) was rated an 8 for structure. One mark was awarded for stating the question that had been investigated, but still failing to state why the group had chosen this particular question. Two marks were awarded for stating what the group had learned about the content of science, and one for reporting an additional insight. Finally, three marks were awarded for the collaborative structure of the note: the group wrote a single note, but within it there were separate paragraphs by the various authors.

5.2.3.3 Conclusion note knowledge claims.

Scale description: The scale awarded one point for each knowledge claim about the science topic the group had researched (which varied from group to group). In addition, two points were awarded for each claim that revealed "an advanced insight" into the content, the nature of science, or knowledge building.

Reliability: Two independent raters rated all eight of the conclusion notes. The number of ratings varied from group to group, depending on the number of knowledge claims in the note: both raters had to agree that a section of text counted as a knowledge claim, and whether or not it, in addition, revealed insight. The inter-rater agreement was 90%, based on all 29 decisions.

Example: The conclusion note by B4 (Appendix C) was rich in knowledge claims — it was judged to contain 10 — and contained a paragraph that struck both raters as demonstrating advanced knowledge building ideas. The authors of the note admitted their conclusions were provisional and implied they are
improvable, thus clearly delineating the limits of their own confidence in their conclusion. There still was not a clear indication that the group saw their note as taking on a life of its own in the database, so the note was judged to be level 6 — Knowledge as improvable personal artifacts — on Bereiter and Scardamalia's knowledge building scale (1996b; Appendix E):

I think that we have come to the conclusion that we really aren't sure (of the answer to "How was Energy Made?") at all! We know that the galaxy wasn't always as it is now, but we also know that there must have been something before the "Big Bang" (for something) to happen. Maybe there was nothing to begin with but the fact that there was nothing for so long meant that there had to be something. Like evolution. Gradually, after a very long period of time, things will change and other things will change as a result of those changes. I, personally, think that this is our best guess. For now anyway.

5.2.3.4 Interview knowledge claims.

Six of the eight groups on CSILE were interviewed in group interviews, beginning April 19, 1997. Two groups (B1, G1) were not interviewed, as a quorum could not be present at the time of the interviews. Each interview lasted between 25 and 35 minutes. The interview protocol is given in Appendix D.

Scale description: Each interview was first parsed into segments consisting of knowledge claims, and then classified using the following scheme: (a) science content (SC), (b) knowledge building (KB), (c) and nature of science (NS).

In study 1, part of the purpose of the scientific contribution scale was how the distribution of notes over the levels of the scale changed from T1 to T2; that is, concern was with concept development. A disadvantage of that scale is the difficulty of judging the acceptability of claims to a subject matter specialist, especially its dependence on a large store of subject matter knowledge. My intention with the present scale was to avoid this difficulty. The scale describes what the conversation is about. For instance, does a group only make claims that can be interpreted as being about scientific facts, or does it also make claims about the nature of science and about the nature of knowledge building?

To aid the identification of what a claim is about, it is necessary to have "exemplars" (Kuhn, 1970) of explanations that are about, or refer to, given things or situations. For example, one can recognize the utterance 'Scientists do
experiments to ...' as being about the nature of science, although the particular view the student expresses about the role of experiments in science may not be one shared widely by scientists. Similarly, 'One thing I liked was that everyone shared their theories, so you don't just have your own idea but a whole bunch ...' may be recognized as being about a knowledge building strategy. Determining if an utterance is about science, however, is less straightforward. An utterance was rated 'about science' if it made reference to an event, the purpose of which was to communicate scientific knowledge. For instance, the science kit included an activity in which the newton was introduced as a unit of force; when students said 'Force can be measured', it was likely to be in connection with (or as a result of) this activity. In other cases it was clear that students were talking about something they had read as part of the unit, or about something they had seen on a TV program about science, such as Nova.

Reliability: Segments of the interview transcript that dealt with science content, the nature of science, and knowledge building, were selected by the author. The author and an independent rater each identified knowledge claims in these segments: they decided which sub-units of the segments were to be counted as knowledge claims. An elaboration of a previous knowledge claim was not counted as a new knowledge claim. The author identified 74 knowledge claims, and the independent rater 87, amounting to 85% agreement. The smaller set of claims was then classified by the author and another independent rater, resulting in 94% agreement.

5.2.3.5 Thinking-Type-in-Use.

Question 2 of the test (Appendix B) presented students with the first two entries of a discussion note, and asked them to make a third entry. The idea of 'Thinking Type in Use' is based on Argyris's (1994) Theory-in-Use, and was used to develop a scale designed to compare the content of the students' responses with the thinking types they declared with them.

Description of scale: The following classification scheme was used. 'Opinions' were statements without supporting evidence; 'mixed framework theories' provided evidence from everyday life, and 'scientific theories' provided evidence from science; 'critical comments' disagreed directly with one of the two entries provided in the question; finally there were questions. All critical
comments were additionally rated as being opinion, mixed framework theory, or scientific theory, leading to a scale of 7 independent levels.

**Example 1:** The following entry was rated a critical comment and opinion: "MT: Well, if energy can't be created then how did it get here? I think that all resources came from the sun in a direct or indirect manner." "If energy can't be created" argued with the first entry; the second part of the response was not supported by evidence.

**Example 2:** "MT: I disagree, because we have energy and we use it in the day and gain it in the night. And soon the sun will be destroyed because it has used up all its energy but maybe it will get it back when it's sleeping, I think." I disagree is a critical comment; a reason is given for this disagreement, but the evidence provided is clearly animist and probably did not result from a science context.

**Reliability:** All of the 54 of the responses were rated by two independent raters, who agreed in 85% of cases.

**5.2.3.6 Logical argumentation.**

**Description of scale:** Each argument provided in response to the test question, 'Is gas affected by gravity?' was rated as being one of four types: Type I, invalid premise, invalid conclusion; Type II, valid premise, invalid conclusion; Type III, invalid premise, valid conclusion; and Type IV, valid premise, valid conclusion.

**Reliability.** All of the 41 arguments provided were rated by two independent, raters, who agreed in 86% of the cases.

**Example:** One type of argument given was, yes, gases are affected by gravity because "all things are affected by gravity" (Argument 5, Table B.5, Appendix B). Is, to use Lehrer's framework (section 2.2.5), a Grade 5/6 student likely to be personally justified in accepting that all substances he or she knows are affected by gravity? If the student knew about massless particles like photons and assumed that only substances that have mass are affected by gravity, he or she might conclude that not all substances are affected by gravity. But there is no evidence for this in this example. That makes the premise of the argument true
(in Lehrer’s framework). The conclusion of the argument is also justified, since it would be more difficult to accept the opposite, that an unfamiliar substance (gas, in this case) is not affected by gravity. Explanatory coherence (Thagard, 1992) would also appear to support this interpretation.

### 5.3 Basic Skills and Overall Achievement

In this section, results on the various quantitative measures are presented to provide context for the case studies and subsequent analyses. A summary of the means on the writing measures obtained from the CSILE transcript and the achievement measures is presented in Table 5.3; the last three measures in the table are means with respect to the groups on CSILE. Conclusion note structure correlated highly with the (total) number of words, $r = .97$, $p < .001$, and the number of notes, $r = .83$, $p < .01$. Figure 5.1 shows standard scores for a number of the measures, and gives an indication of the variation from the means.

<table>
<thead>
<tr>
<th>Measure</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. of words written$^a$</td>
<td>280</td>
<td>249</td>
</tr>
<tr>
<td>Total no. of notes$^a$</td>
<td>6.11</td>
<td>4.78</td>
</tr>
<tr>
<td>Flesch-Kincaid reading level$^b$</td>
<td>7.11</td>
<td>1.64</td>
</tr>
<tr>
<td>Test Score$^c$</td>
<td>44.2%</td>
<td>7.47%</td>
</tr>
<tr>
<td>Conclusion note structure$^c$</td>
<td>4.63</td>
<td>19</td>
</tr>
<tr>
<td>Conclusion Note Knowledge Claims$^c$</td>
<td>3.88</td>
<td>4.52</td>
</tr>
<tr>
<td>Interview knowledge claims$^c$</td>
<td>15.8</td>
<td>5.95</td>
</tr>
</tbody>
</table>

$^a$Measures of individuals, $n = 30$. $^b$n = 12. $^c$Measures of groups, $n = 8$.

**CTBS scores:** Incoming scores and changes in scores during the year were compared with the same pool of data as in study 1. Again, there were no substantial differences between the class means and the means of the pooled data. The largest differences were the change scores for Grade 6 reading comprehension, $M = .712$, $SD = 1.01$, which was .48 standard deviation above the pooled mean, and for Grade 6 spelling, $M = .883$, $SD = 1.44$, also .48 standard deviation above the pooled mean. When the data from this class was compared directly with that of John’s class of study 1 (i.e., excluding the remaining 22
classes of the pooled data), the change scores of Paul's class on vocabulary were significantly higher than those of John's class, $M = 1.02, SD = .94, U (N = 55) = 246.5, p < .03$, effect size $= .59$.7

**Writing output and quality:** The database consisted of 213 note entries, 1/3 of the database of study 1. The mean number of words and notes written and Flesch-Kincaid reading levels are shown in Table 5.3. Of the 30 students who contributed notes to the database, only 12 wrote more than the approximately 250 words needed for computing the Flesch-Kincaid reading level; only 5 students wrote more than 600 words (approximately the LFC mean in study 1). In this class (as was the case for students with LFC achievement in study 1) a challenge for improving knowledge building practices is to increase writing output. Comparing class means of CTBS scores and Flesch-Kincaid reading levels for this class with that of study 1 indicates that the lower writing output was not due to literacy levels reported for the class.

---

7Or using a one-way ANOVA, $F(1, 53) = 4.63, p < .04$. 

---

**Figure 5.1:** Standard group scores of No. of notes, test scores, conclusion note structure, conclusion note knowledge claims, and interview knowledge claims. G4 and B4 were not interviewed. Observe the discrepancy between measures that do not depend on contributions to the CSILE database and those that do, for B3.
Test scores: The test results were lower than expected: $M = 45.3\%$, $SD = 13.8\%$. John's class, which had completed the science kits and a brief unit on force and energy, but had not used CSILE for this unit, obtained nearly the same mean, $M = 46.3\%$, $SD = 17.0\%$. However, there was a considerable difference in the distribution of marks for the two classes. In Paul's class, a large number of students scored near 50\%, while this was not the case for John's class. (Four of eight groups in Figure 5.1 have small $z$ scores.) Examining the scores by grade, the Grade 6 students, $M = 48.3\%$, $SD = 14.7\%$, were only 0.41 standard deviation above the Grade 5 students, $M = 42.0\%$, $SD = 15.8\%$; this difference was not significant at $p < .05$.

Conclusion Notes: Only two notes were judged better than a 5 (out of 10) on structure, and only three notes were better than a 4 on the content scale (one 5 and two 8s). Two common shortcomings of the notes were that they provided inadequate context for the reader and did not show very well what the group knew. Many notes did not meet the assignment requirement that it be a group summary, that is, a collaborative effort. The conclusion notes of groups B1, G2, G3, and B4 have been collected into Appendix C.

Interviews: For groups G1-G3, B1, B2, and B4 — six of eight groups — the number of notes written, the test, and the conclusion notes tell essentially the same story about the success the groups had with the unit, disregarding small $z$ scores (Figure 5.1). A notable exception is B3, who scored well below average on all these measures, except the test, on which they scored well above average. As Figure 5.1, shows, this picture breaks down when the number of knowledge claims made during the interviews is added (cf. G1 and G3, and to a more limited extent, B3). G1 is one of the groups to be discussed as a case in the next section. According to Paul, the girls in G3 were quite "shy," and their behavior during the interview suggested this was the case: it was difficult to obtain more than single-sentence responses from these girls. For B3, the interview and the test told a consistent story. The inconsistency of that story with the other measures can be explained by assuming that these students had considerable knowledge, but that this knowledge did not result from their work with CSILE (consider, for instance the number of notes for them in Figure 5.1).
Figure 5.2 shows the distribution of the knowledge claims over the various levels of the interview knowledge claims variable. To understand this graph, it is important to recall that, in all interviews, content was discussed before the nature of science. In most cases, a large number of claims about the nature of science indicates difficulty to keep the conversation going when it concerned the content of the science the group had researched. This interpretation goes some distance toward explaining the overall performance of B3 — the ability of this group to hold up their part of the conversation about the content they were supposed to have researched was modest. The students said they liked science, and evidently had substantial knowledge about science. It is also worth noting that G3 (the shy group) did admirably on content, compared with most other groups. B2 gave a strong performance with respect to content, but the small number of claims in the other levels kept their total number of knowledge claims slightly below the class average.

![Knowledge claims made during end-of-unit group interviews.](image)

**Figure 5.2:** Knowledge claims made during end-of-unit group interviews. The categories are based on “what the claim is about,” not on acceptability to a scientist of the claim made.

**Gender differences:** Figure 5.1 reveals a possible gender difference: the z scores of measures directly connected with work on CSILE (i.e., excluding the test and the interview) tend to be more extreme in the right half of the graph, that is, for boys. This data was normally distributed (Shapiro-Wilk test, p = .31), and the
variance of the errors was equal across genders (Levene test, $F(1, 22) = 3.22, p = .087$). A t-test revealed that $|z|$ were larger for boys than for girls, $t(22) = 4.17, p < .001$, effect size = 1.30. A Mann-Whitney test led to the same conclusion, $U(N = 24) = 18.5, p = .002$. Of course, this effect is the result of both below- and above average performance. For boys, six of twelve measures (three variables $\times$ four groups) were below average ($z < 0$); for girls that number was nine. Thus, although girls performed below average on more measures than boys, the degree in which they did this was less extreme than when boys did it (i.e., B2 and B3). On the other hand, girls were not extreme performers above average either. It was not simply a matter of writing less when the unit was in progress, because the form as well as the content of the conclusion notes was included in this analysis. This result must be treated with extreme caution: not only is the number of degrees of freedom small, we are also dealing with groups rather than individuals, so that it is not even clear if we are dealing with within intra-group social issues or individual issues. Suffice it to say for now, that there is at least sufficient evidence for gender differences here (and in chapter 3) to warrant a careful study of them in future research on computer supported collaborative learning environments.

5.4 Case Studies

In this section, two case studies of groups on CSILE are presented. Rather than examining if students developed a research program and helped others to advance their knowledge, as in study 1, the cases probe the end-of-unit knowledge claims of the groups. Taking the work on CSILE to be a work in progress, could the students hold up their end of a critical conversation with someone knowledgeable in the subject matter they had researched? The cases are mostly developed from the interview data, see Appendix D for the protocol. Both the interviews and the cases were framed by dimensions 2-4 of scientific literacy described at the end of section 5.1.1. As in study 1, the cases presented were selected to show different types of performance. The first case, G1, is a very

---

8The distributions over the levels of interview knowledge claims (Figure 5.2) were also examined for a gender difference using an analysis of homogeneity, but no significant effect was found, $\chi^2(3, N = 96) = 3.74$. 
articulate group of girls who discussed the nature of science. This group reconstructed its knowledge during the conversation to maintain a commitment to empiricism. The second case, BI, was the second most successful group of the class (cf. Figures 5.1 and 5.2) and focused more on scientific content learned. All students in the cases were in Grade 6.

5.4.1 Group GI: Empiricism and the nature of science

Only two of three girls in this group could be interviewed, as the third girl was absent on the day of the interview. Susan and Christina wrote a slightly less than average number of notes; their test result was slightly above average, and their conclusion note was below average — especially its content (cf. Figure 5.1). Their mean incoming (grade equivalent) CTBS language composite score was 8.4 (SD = .77), well above the pooled mean for Grade 6 of 6.6. As Figure 5.2 shows, they made a substantial number of claims about knowledge building and science content. The content claims, however, were not made while the topic of conversation was the science they had researched: although Susan and Christina said they had learned much about energy, they could not articulate their knowledge. The parts of the conversation focusing on knowledge building and the nature of science were more satisfying.

5.4.1.1 Content

The group studied energy. What did they say they learned about energy?

1. Christina: I thought that energy was just simple — just energy — but then when I did it on the computer and I did more research I found out that energy isn’t just energy, it’s more than energy.
2. Jan: So — what is it? What is it that you know?
3. Christina: Well, uh .... (thinking hard, looking at Susan) I — It’s pretty much everything. I mean, when you wave your hand, that’s energy [SC], anything is energy.9
4. Christina: Power is sort of — energy comes from the sun [SC] and power is what the sun is giving. I really don’t know how to say it.
5. Susan: It’s how the sun emits the energy to people, and power is not something that’s just there (waving hands) like energy, but you need to just ....
6. Susan and Christina: Do something, to get power. [SC]

9Examples of the interview transcript coding are indicated by phrases in italics, followed by the code in brackets.
7. Susan: And it was interesting to learn that if you just do anything you create energy.

8. Christina: Yeah, I didn't know that. Energy, when I started the project I thought it was either solar energy or electric electronic energy [SC] and I didn't know anything else about — when you move you really cause energy.

In line 3, the group said it had learned that energy can be stored in the motion of a hand. This utterance suggested the scientific idea of kinetic energy; however, the group did not elaborate what scientific understanding of kinetic energy would entail (e.g., how it depends on speed or its relationship to work). Other utterances indicate that the group's concept of energy remained quite inarticulate. For example, the group did not differentiate between energy and power (4-6). Energy, in their view, is natural (5), but to get power you "have to do something" (6). In addition, some claims are likely to have been based on everyday experiences, not the class's knowledge building. Students may have learned from a TV program that 'energy comes from the sun' (4).10 Similarly, the claim that one has to produce power was used by other students (B3) in connection with house wiring and the generation of electricity at power generation stations at Niagara Falls and Pickering. Given that both of these facilities are in the geographical area in which the students live, it is plausible that students would have developed this understanding of power as made from TV programs or experiences prior to the unit.

What has been quoted is essentially all the group volunteered on the topic of energy. The group was able to keep the conversation going for some time, but the conversation was not focused on understandings of energy as a scientific concept. For example, it might have given an operational definition of energy, a classification of different kinds of energy, or demonstrated a recognition that energy is a scientific concept among other scientific concepts such as power and heat.11 While I did not use scientific terminology to ask my question (2), it left this possibility open for students. In chapter 2, I argued for a view of conceptual change which entailed a shift, in scientific contexts, to engaging in discourse

---

10 Even if they did not, the analysis of concept articulation, section 5.6, suggests that a considerable number of students made similar statements to this one early in the unit, so they may have known this fact prior to the unit.

11 An operational definition of a concept specifies what one must know experimentally to know that concept (see Arons, 1990 for a discussion). For example, to know the average speed of a car, one can measure the distance it has traveled, the time in which it did this, and divide the former by the latter.
practices that resemble scientific discourse more closely. The discourse of this interview did not meet that standard because of its lack with scientific concepts.

An examination of the science vocabulary provided with the science kit from the school board (Table 5.1), and the activities of the kit (a paragraph above Table 5.1) reveal that there were problems with the science kit from the point of view of developing conceptual understanding; this may have contributed to the noted lack of subject matter knowledge. The first four activities of the kit are all concerned with friction. Activity 1 — Measuring the force needed to move an object — focuses on Newton’s first law (also known as the law of inertia) and an interpretation of friction as a force that must be balanced by an applied force before there can be a change of motion. There is a deep result from Newtonian physics to be learned from this activity. While according to Aristotelian physics an applied force is always required to maintain motion, Newton’s first law states that no *net force* is required to maintain motion at constant speed. Frictional forces are introduced to account for the fact that an applied force is often required to maintain motion at constant speed. Unfortunately, the scientific terminology to develop this result (e.g., 'inertia', 'force of friction') is missing from the prescribed vocabulary. Activities 2-4 each consider one variable that affects friction, but here too, the vocabulary that can bring out the relevance of the observations is missing (such as a name for a variable that can describe how the force of friction depends on the *material* from which the substances that move against one another are made). In consequence, activities 1-4, which could have led students to (partially) articulate a coherent framework for understanding friction, were mismatched with their accompanying vocabulary, and were likely to leave students with a set of experiences that seemed to have little in common. A similar argument can be made for the last three activities, which are all about rocket propulsion. In sum, we have here a set of activities that is reasonably well aligned with knowledge building epistemology (it is capable of providing a rich context for progressive problem solving), but its accompanying vocabulary has limited capacity for shifting early discourse to a more scientific discourse of the topics studied.

The LKB model can guide an attempt to make improvements to the unit in several ways. First, the previous paragraph implies that it is important to start the unit with activities and/or instructional materials that are well suited to the
goal of conceptual understanding of the subject matter, or, if that is not possible, to at least be aware of their shortcomings and compensate for them later in the unit. For example, we could have done more to help students articulate what could be concluded from each of activities 2-4, and how together these activities began to examine the conceptual territory of friction, helping the students to develop specific research questions that they could pursue as the unit progressed. Alternatively, we could have modified the prescribed vocabulary to include some important words that were missing. However, it is worth noting that Roschelle (1996) has observed that convergent conceptual change is possible even when students do not engage in discourse akin to the hypothetico-deductive talk of scientists; and students may invent their own words for scientific concepts (Roth, 1995; Roth & Duit, in press). Helping students to articulate questions that can drive conceptual change is more important than introducing the concepts in scientific terms. All this requires that the teacher to have considerable knowledge of the conceptual domain of the unit, and in my opinion it is an unrealistic expectation that teachers will develop this in all subject areas they have to teach. Outsiders who have this kind of knowledge may be able to support the teacher, but in the present case this did not work because the outsider (i.e., the author) did not fully appreciate the pedagogical issues connected with learning the subject matter with the science kit used. (This is likely to be due to inexperience as a researcher.) In such cases, one should not expect students to develop conceptual understanding (although some students may succeed).

The quoted material may also reflect difficulty with summarizing what has been learned (1, 8). One often finds this in notes of the What We Have Learned thinking type, not only in this database. Frequently there are claims that something has been learned, while descriptions of what has been learned are lacking. A clear example is Christina's conclusion note:12

Doing this project has been very informative. I have learned; how humans capture energy, about atoms and matter, what the most recent form of energy made by humans, and why energy is useful. This project has changed my feeling for energy.

The LKB model suggests that there are two windows of opportunity to address the need to learn to summarize effectively: (a) when students formulate research

12This group did not write a collaborative conclusion note. Instead, two of three students in the group each wrote their own conclusion note.
questions, which requires that they are aware of, and can justify, problems of understanding; and (b) when the class engages in discourse to “validate” (Kuhn, 1970) knowledge advances. John’s What We Have Learned discussion notes at the end of the unit of study 1 were an effective strategy for the latter of these: the students improved the exactitude of their claims via commenting, and John participated in this process.

5.4.1.2 Ideas about knowledge building.

In the interview with this group, considerable time was spent on the topic of knowledge building. Often, the girls acted as informants of what had taken place in the class’s work on CSILE, but in some cases their comments indicated understanding of knowledge building ideas.

9. Susan: Well, one thing I liked was that we could all, everyone could give their own theories, we could read everyone’s different theory and kind of come to a conclusion about something [KB], or all of the different theories that everyone made so we could get a broad (showing with hands) understanding.

10. Christina: I liked how you could comment on other people’s work and if you were doing the same topic then you could say what you thought and you would give everybody an ideas of what the whole class thought about one subject [KB].

11. Jan: So, why do you like doing it that way?

12. Susan and Christina: Well —

13. Susan: Well, you don’t just have one person’s theory on something, you have a lot of people’s theories on the same thing. So you don’t just have one idea of what something is.

14. Christina: Yeah, really. And all the questions were discussion questions, they weren’t just this is the answer, there is no other answer, there were so many answers [KB], every group came up with a different answer for pretty much everything.

The group described powerful knowledge building strategies: considering a variety of theories (9), and commenting (10) as an aide for consolidating collective knowledge. The claim in line 14 is an indication that the group saw knowledge as representable from different perspectives, which is level 4 on the Bereiter and Scardamalia knowledge building scale (1996b; see Appendix E). According to the group, the class did much commenting:

15. Susan: We did a lot of commenting and read all the other theories, and talked about that, and I thought that’s a really good idea. That’s really good, what you learn.
16. Christina: A lot of people commented on our work, too. We had, I think, more than 20 comments on our, um, note.
17. Susan: Almost everybody in the class commented on everybody’s note.
18. Christina: Almost every group — almost everybody from every group commented.

Their estimates of the commenting activity by the class, however, is not confirmed by the analysis of the comment thinking type in section 5.7, although Susan did comment on notes from almost every group.

What was the nature of the comments?

19. Jan: So, can you give me an example of a comment ... that made you change your mind about something you had written? ...
20. Susan: Well, I think that there was — there were — a few comments that, that said like I added another question to my note, or um, I added another theory to my note. Just from reading theories on your note. But there wasn’t a lot of changing the theories.
21. Christina: A lot of people had definite theories, they stuck to their theory, but they still, um, kind of agreed with the other person who wrote their theory. But most people stuck to their own theory.

22. Christina: Well, I didn’t totally change my theory but somebody commented on, on, well I told what capturing energy is, what you have to do to capture energy, but I didn’t show actually how you do. ... And for my um, (thinking) graphics note I put how they capture energy, but first I was just showing what energy was and how it was cap — . Not really how it was captured, but how it was there and, um, how to begin to capture it. But then I showed how to capture it, so ...
23. Susan: Yes, and people did comment, saying that might not be a very correct theory or that really doesn’t make any sense. So you would go back and change it.
24. Christina: And sometimes people help editing and everything.

Thus, it appears that commenting had a relatively small impact on theory improvement: students’ theories were fixed (20-21), although some comments did challenge them (23). Other comments asked for information on graphic notes in order to make them clearer.

5.4.1.3 The nature of science.

The group was asked in what ways they thought their work on CSILE was like the work scientists do. The group noted that scientists also made theories, did experiments, drew diagrams, and so on; eventually, the conversation shifted to a discussion of creationism and evolution, which included some of the group’s best content claims. Sections of this long episode are quoted intact to preserve a
sense of the conversation and to evaluate whether the group was collaboratively reworking their knowledge in the face of new information or challenges to their theories.

25. Christina: I think that religion and science, there’s total difference, and I’m actually an evolutionist. I believe that the world didn’t start from God …


27. Christina: And, um, before now a lot of people didn’t know anything about evolution or anything like that, they just thought God made the world —

28. Susan: It was scientists making theories, and yeah.

29. Jan: What gives you confidence to side with evolution rather than with the creation story?

30. Christina: If you think about that one, that’s really unrealistic, the one how God made the world, and how did God get there (waves arms) anyways! I mean, no one knows how God got there.

31. Susan: Yeah, it’s really what came first, the chicken or the egg. It’s, I, I am an evolutionist, I believe that, like, *humans evolved from bacteria* [SC], and but um, there’s, there could be a slight bit of truth in the creation story. I’m not just going to say I don’t believe at all, because you’ve got to be open.

32. Christina: Yeah, but I, I, I believe it’s really unrealistic and my, what I said before, how did God get there, that’s what I’d like to know if I’m going to believe in the creation story, how did God get there?

33. Jan: OK, so, um, so you learned about force and energy. How did that, um, you were saying before that energy comes to us from the sun, right?

34. Susan and Christina: Yes!

35. Jan: So, how did the energy get to the sun? I mean, where — how did that start?

36. Susan: I think that *the sun is actually energy, and it’s made up of lots of energy and a little bit of matter and stuff, and gases, but it is mostly energy* [SC]. And it just comes out of the sun and to us here.

37. Jan: But — you’re not worried about where that came from? How did the sun come to have energy, or, or where did the sun come from?

38. Susan: Good question!! (smiles) I — I — I wouldn’t know how to answer that one.

39. Christina: It’s, um, ah — that’s a really hard question. I —

40. Susan: (laughs)

41. Christina: I mean, I have a little bit of an idea, but I really wouldn’t know how to explain it, or um —

42. Jan: Well — I think that the answer doesn’t matter so much, but what I’m trying to expose for you is that when we look at a scientific theory, there are still some gaps there.

43. Susan: That are left hanging.

44. Jan: That’s right, we don’t really understand where they come from. And I think that’s what you were saying about evolu — sorry — about creationism, that there’s just too many holes.


46. Christina: And, um, they found fossils of early humans and primates, so we can believe that more than nothing. You don’t even know if God is actually there, or not there.
47. Susan: It's just like explaining power, energy and force. Like, if you were to explain that to someone, it's kind of difficult. Because you understand it in your head, but it's hard to articulate. But, ah—so it's just the same with how did the sun get there and how did the energy get there with the sun.

The group began by asserting its alliance with evolution (25-26), although it is not clear if they saw this as a theory or a fact. In line 27, Christina recognized the creation story as a people's attempt, albeit a naive one, to account for the existence of the world and their place in it: had they known about (the theory of) evolution, then they would not have needed the creation story (27). Christina interpreted the creation story literally, and rejected it because it is "unrealistic" (30): the world is not likely to have come into existence that way. Many biblical scholars would agree with her on this point, as would scientists. Einstein, for instance, has said he believed in a supernatural power, that is, a god, but not in a god who is interested in a personal relationship with us (see March, 1996). Susan was willing to allow that there might be an element of truth to the story, adding that one must be "open." Her claim that "humans evolved from bacteria" (31) is striking because it suggests (to me) evolution from simple to more complex living organisms; but she provided no information that would allow us to see how this happened.

In this episode, a central question became, Where did God come from? Christina had introduced it in (30); Susan agreed with her ("Yeah", 31) and attempted to explain the dilemma; finally, Christina restated it (32), adding that she would have to have an answer to this question before she could accept the creation story. What struck me during the interview was that the students were not willing to accept that religion leaves something unexplained, and I wanted to know if they would require evidence of the same standard in science. Therefore, we returned to an area we had discussed before — energy from the sun (33). The claim Susan made in line 36 is interesting because it had several ideas necessary for a scientific explanation of the sun as source of energy: hydrogen gas is converted in thermonuclear reactions to heavier elements like helium and lithium, and carbon. I pointed out to the group that their explanation, like the creation story, left some things unexplained (i.e., where the sun's energy and/or material came from, line 37). Christina must have thought that this meant the group's argument against creationism was flawed; she attempted to save their
previous explanation (31) by referring to empirical evidence for the theory of evolution (46) — fossils. Finally, Susan explained that sometimes, although we do have knowledge, it is too complex to express verbally (47). What she was doing, it may be argued, was to try out an experience on her theory that some things cannot be explained. If she was correct, she would be justified in accepting the theory of evolution, although it does not explain everything.

Since it appeared we would not be able to solve the 'How did God get there?' question, I shifted the conversation to experiments:

48. Jan: Why do you think scientists do experiments? How does that help in understanding —
49. Christina: Well, if they're making a theory, they have to see if the theory is actually somewhat close to the actual — I mean, they can't just say I have this theory and people will believe them. I mean, they have to have proof that this theory actually works [NS], so they do experiments, and then they have proof [SC]
50. Susan: They have to somewhat prove what they're saying, even though it is only a theory. They can't just create a theory out of thin air. There has to be some basis.
51. Christina: I ... [inaudible, talking over Susan].
52. Jan: OK, so you're saying that they have to do the experiments first, so they can build the theories from ... ?
53. Susan: Yeah.
54. Christina: Or they can do the theory first, and then prove the theory with the experiment.

Christina made one important distinction between theory and opinion (49): theories must be based on something, must have proof. In her view, experiments provide such proof. Susan, more than Christina, seems to have had the idea that scientific theories are often built up from empirical evidence.

5.4.1.4. Summary.

Section 5.4.1.3 indicates that Susan and Christina were highly articulate: they were confident in their claims, defended their position against the creation story, and introduced new evidence in doing so (46-47); this is consistent with their high CTBS language scores. They had important insights about knowledge building. Thus the interview reveals that they were able to hold up their part of a conversation about science and religion; the discussion of the role of experiments was shorter, but suggested that they understood this role to some extent. However, they were able to contribute only a few claims about science
content; thus their performance was weak for dimension 4 of scientific literacy, as described in section 5.1.1 (detailed knowledge of a small number of science topics). This is unfortunate because both students were academically successful at the time of the unit, and should have been able to reveal knowledge of a quality similar to that of Matt, Sean, Jerry, and Emma in study 1.

I have already mentioned that the LKB model should be capable of guiding the early part of the unit — by paying adequate attention to helping students frame research questions and evaluating knowledge advances. Another way in which the LKB model may be useful is by stressing in curriculum development that the process by which knowledge advances are made receive adequate attention. Beginning from “ill-defined problems” (Roth, 1995; Roth & Bowen, 1995), defining a research program, and communally validating knowledge advances are aspects of the process of science (see Latour, 1987), and valuing them in school science presents to students a philosophically more valid image of science. The LKB model makes such processes explicit in what is to be considered in designing a curriculum unit.

5.4.2 Group B1: The science of rockets

This group’s test scores were just above the class mean; their CTBS incoming composite language score was 6.7 (SD = 1.4), close to the pooled mean. The group ranked first on the number of notes written and the structure of their conclusion note; the content of the conclusion note and the number of knowledge claims made in the interview ranked similarly, see Figure 5.1. However, two group members — Stephen and Earl — wrote very little. One of the most prolific writers in the group, Tim, could not be interviewed,¹³ and is not included in the case; another student, Nathan, was absent on the day of the interview.

5.4.2.1 Cars as rockets.

Stephen, Earl, and Doug had studied rockets, and the conversation began by discussing what they had learnt about rockets:

1. Stephen: I especially realized that there wasn’t just one kind of rocket like the space rocket, but like missiles and war stuff, and there’s other stuff, that are like rockets as well [SC].

¹³Permission for the video-taped interview could not be obtained for him.
2. Jan: Such as?
4. Earl: Rockets that some kids use for scientific ...
5. Jan: So — so, what's the principle? What's the idea that makes all of these things alike?
6. Stephen: It's just something that is ... (thinks) pushing.
7. Earl: It's force ...
10. Stephen: Yeah, there are other ...[inaudible] that are very different, for example when you drive your car [SC].
13. Earl: Well, it is. The only thing similar from a car to a rocket is the exhaust pipe, because the exhaust sort of comes out the back. But a car doesn't use the exhaust to power, whereas a rocket shoots, like, flames out the bottom, but that pushes, gives it lift off.
14. Jan: OK, so, if you try to power a car from its exhaust, what, what would have to be different about the way the exhaust comes out?
15. Earl: Well —
16. Doug: It would have to be more powerful, so it can push the car.
17. Stephen: In a car — in a car, most of the power goes into the wheels ...
18. Doug: Yeah.
20. Doug: Yeah ...
21. Earl: The exhaust is just to get the fumes out of the car, so it doesn't all come, like, build up in the car and you'll die.

In the first few speeches, the group had a better start than the group discussed in the previous section did. Initially, Stephen seems to have associated 'rocket' with space travel (1), but now he "especially realized" that missiles were also rockets. In line 3, he claimed that jet planes also use rockets, which indicated a misunderstanding of jet engines to me. Earl completed Stephen's account, referring to (model?) rockets that some kids use (4). Thus the group summarized effectively what it had learned in this part of the interview, although not yet in terms of scientific concepts.

I decided to turn the conversation to scientific concepts in line 5. Stephen and Earl said that force and energy were responsible for propelling the rockets (6-9); but rather than elaborating on this, Stephen introduced another example in line 10 — a car! My sense during the interview, from visual clues to the level of their engagement in the conversation, was that some members of the group had discussed this example before, although there were no notes in the database that could have been based on such a discussion. (It is worth noting that Doug contributed the conversation to underscore points made by others [18, 20]).
Rocket propulsion can be explained by the law of conservation of momentum. In typical rockets liquid oxygen and hydrogen is burnt (some use solid fuel), and gaseous products of this reaction (exhaust gases) are ejected from the rocket. In the absence of an unbalanced external force on the rocket+exhaust system, the exhaust gases impart an impulse — i.e., a change of momentum — on the rocket; the change of momentum of the rocket balances the momentum of the exhaust gases, so that the total momentum is constant. Thus important features of rocket propulsion are the co-occurrence of a forward momentum change of the main body and the backward ejection of gases. Jet propulsion also has these features, and so does the motion of a car. Both of these examples (3, 10) may be instances of the misapplication of a too-limited concept of rocket propulsion. Lines 14-21 represent my attempt to understand the car-as-rocket example better, without introducing momentum or Newton’s laws. In retrospect, I might have learned more about what students understood about rocket propulsion by introducing these, but this would have required some visual representation, an “inscription” (Latour, 1987) as well as, perhaps, other artifacts that could have mediated the conversation (Roth, McGinn, Woszczyna, & Boutonné, in press). It is impossible to conclude from this interview segment if the group could have worked out a more conceptual understanding of rocket propulsion if I had taken the conversation in that direction; however, in their conclusion note the group provided a concise description of how a rocket works in terms of force (see also appendix C):

To make a rocket work fuel goes into a sphere with a hole in the bottom. The fuel then ignites and tries to get out of the circle by putting force against the walls. The fuel can’t escape any other way so it is forced to go out the bottom, the force of the explosion then gives the rocket lift off. (Earl)

5.4.2.2 Knowledge building.
This group, too, discussed aspects of knowledge building.

22. Stephen: ... other people get to discuss on, um, what other people like to, like what they did. So, you don’t just get, um, one theory all the way through. You have about ten theories of other people. They give you what is a rocket.

---

14 In jet propulsion, slow-moving air enters the front of the engine and is accelerated as it passes through the air; this results in a forward momentum gain of the engine.
23. Doug: Yes, that’s always really good, because if you just got one person, you figure that’s the, that’s the answer to the question, but if you have 10 different people saying something, saying different things, then you have more of a chance of getting a real answer [KB], if it’s hard actually to find.

24. Earl: ... some theories are totally off — being right — and some are right. And then a lot of them (softer) contradict each other, like general theories. But it sort of helps you get the answer. If you see one side — it could be this, and it could be this. (pointed to different sides)

This explanation of the power of collaboration (23), building up knowledge from what a group knows collectively, is congruent with what G1 said they had liked about the unit. Of more interest is the information B1 provided about the interaction between work on CSILE and in face-to-face mode.

25. Stephen: ... we should have had more time to go down and research our topic. Like, um, we didn’t have too much time to research, we just put our theories down.

26. Jan: How much do you estimate you spent, say, every day?

27. Earl: About 20 minutes.

28. Jan: And that’s both on CSILE and ...

29. Earl: More on CSILE. Because we have CSILE groups, and they each go on for about 20 minutes, I think.

30. Stephen: And the problem with that is that, um, you can’t get a joint effort in your group, because you can only explain it either at your desk or through ... the computers, and ...

31. Jan: Did you do a lot of talking around your desk about what you were doing?

32. Earl: Not that much.

33. Jan: Not much, so you — you ... 

34. Earl: You could sort of call a person over to the computer if you wanted them to see something that you had done.

According to the group, there was insufficient time for research (25); their estimate of time on CSILE (27) is consistent with one given by G1. This episode suggests that there was little face-to-face discussion about the unit among this group, apart from showing each other something while one group member was at the computer (31-34). For logistical reasons, some groups on CSILE were not perfectly matched to groups that had been together for the experiments.15

15While the experiments had been conducted at times the whole class was together, this was not true of the subsequent work with CSILE.
5.4.2.3 **Dynamic lift.**

I asked the group about how an airplane stays up. This had become an important topic in two other databases (Lamon, Reeve, & Scardamalia, 1997; van Aalst, 1997), and provided an opportunity to discuss Newtonian concepts and to observe the group's knowledge building in process.

35. Jan: Can you tell me, how does a plane stay up, once it starts flying?

36. Doug: I think that when they use the jet engine to push off, it goes so fast that when they're in the air they're still gliding. Because, when you throw an airplane (makes a motion with his hand to illustrate) it'll just glide for a while and then it'll go down. I think they use the same... [inaudible]. Because the air, the wings are large enough, so they hold air above and below to hold it steady.

37. Jan: So —

38. Doug: And then they have the smaller engines just in case it starts to go down.

39. Jan: So it's almost like a football or something, you have to kick it hard, then it will go farther?

40. Doug: Yeah, it will go higher, then it'll start to drop.

41. Earl: Aerodynamics.

(Doug was a more active participant in this part of the conversation than before.) I was skeptical about Doug's theory (36), which seemed more appropriate for paper airplanes than jet planes, and asked the group how Doug's small engine keeps the plane air-bound:

42. Jan: ... You talked about little engines to keep it up, if it starts to fall down. How does that work?

43. Earl: Because there are wings. Fast air goes, I think it's either...

44. Doug: *Hot air goes to the bottom to push up and cold air goes to the top to push down* [SC].

45. Earl: Yeah.

46. Doug: So they're equally pushing.

47. Earl: So it just hovers.

48. Doug: Yeah, keeps a balance between the two.

49. Earl: It's the... [hesitates] Bernoulli's Principle.


51. Jan: What does that principle say?

52. Earl: It's just how — the plane sort of hovers through the air, and there's this hot air on the bottom, and the hot air is going that way (shows air below the wing flowing towards him), so it's also pushing, *forcing against the wing of the plane* [SC].

The conversation proceeded with difficulty. Earl's "they're equally pushing" (46) suggests an association of the equilibrium situation (i.e., that the plane maintains its altitude) with balanced forces, but it does not take the force of
gravity acting on the wing into consideration. In fact, the air must provide an upward force (component) to balance the downward force of gravity. Such an upward force component exists if the pressure below the wing is higher than above it; according to Bernoulli’s principle, this requires that the air flow faster over the wing than it does under the wing.\textsuperscript{16} Earl and Doug had evidently heard of the Bernoulli principle, but their best explanation, in line 52, does not suggest they understood it. Their attribution of dynamic lift to a temperature difference is puzzling. To sum up, although the group may have invoked the Newtonian idea of balanced forces as a condition for dynamic equilibrium, they did not consider the relevant concepts (gravity, air velocity, pressure) and focused on an irrelevant concept (temperature). That the group did not understand dynamic lift is not surprising or troublesome (the Bernoulli principle was not explicitly targeted by the unit, and an examination of the CSILE transcript revealed that the group had written no notes about it), but if the group had developed parts of a Newtonian understanding of force from their research, one would have expected them to consider concepts related to force — the force of gravity, pressure, and mass, to name a few, that were targeted by the unit (Table 5.1). My impression during the interview was less negative than this, however. The group had seemed very enthusiastic and confident in their claims (as group G1 was); Doug and Earl jointly constructed an explanation of what they thought they understood (e.g., elaborating on each other’s claims, lines 46-48). What is troubling to me — as a teacher and subject matter specialist — is that the group seemed to have thought they had understood something deeply when they had barely begun to explore the subject matter.

\textbf{5.4.2.4 Summary.}

This group had a mix of achievement levels, as the CTBS scores, Flesch-Kincaid reading levels, and writing output indicate. Like G1, the group had strong convictions, and they provided insights about knowledge building similar to those of G1. They had evidently thought about rocket propulsion, but their

\textsuperscript{16}Logically, the differential in air velocity must be \textit{prior} to the differential in pressure, if it is to explain the cause of dynamic lift. The Bernoulli principle does not explain the velocity differential. As K. Hoyle explained in a Knowledge Society Network database, “the incorrect theory is based on the ‘fact’ that if molecule A and molecule B encounter the leading edge of the wing at the same time they will both leave the trailing edge at the same time. While this is possible, it is not necessary. What is necessary for the air foil to create lift, is for the air foils shape to deflect air downwards (called the Coanda Effect).”
best explanation (in their conclusion note) revealed that they understood it only at a descriptive, not a conceptual, level. The application of the group's understanding of rocket propulsion to the motion of a car was based on only a few aspects—a change in forward momentum of the main body, together with the backward ejection of gases—of a scientific explanation of rocket propulsion. The group's explanation of dynamic lift did not invoke all the required Newtonian concepts. On the whole, however, the group's performance in the interview was stronger with respect to subject matter knowledge than that of G1; they also provided more evidence of thinking on their feet in the domain they had researched.

Could this group have benefited more from a unit designed with the LKB model? The group did have a problem of understanding (understanding rocket propulsion, plus several situations that seemed akin to rocket propulsion), but it was not framed in terms of variables and concepts that could have helped them to articulate parts of a Newtonian understanding of mechanics. With a prior awareness by the teacher of the conceptual terrain the students were expected to learn, the group's research might have considered variables like mass, weight, and pressure. The conceptual relationship between activities 5-7 of the science kit (rocket propulsion), and the other activities should have become clearer. The cases, being certainly among the better in this class at the time of the unit, suggest that an important role for curriculum development with the LKB model is to give the unit a scientific focus—the progressive inquiry should lead to scientific knowledge, however incomplete it may remain by the end of the unit.

5.4.3 Discussion

I intended the interviews to be conversations that could simulate knowledge building experiences students had had during the unit in as naturalistic a setting as possible. In interviews for protocol studies and for assessment studies the interview conditions are controlled carefully. In assessment, for example, one tends to standardize the amount of help that is given to a student (Brown, Campione, Webber, & McGilly, 1992); in protocol studies, the interview is framed by a protocol that is designed ahead of time, such as a problem solving task that involves an Atwood machine or software (e.g., Throwbridge & McDermott, 1980). In the current study, I was interested in the students' ability to use their
knowledge in joint problem solving — could they, with me, jointly construct explanations of issues relevant to what they had studied? For this reason, I actively participated in the conversation, sometimes contributing information; on the other hand, I did not want to be “knowledge-telling.”

Compared with interviews conducted for the CSILE project in other classrooms (CSILE Video Archive, 1989-97), the results were weak. The two groups discussed in the case studies were both enthusiastic and confident in their claims, but their explanations did not invoke scientific concepts. For other groups (e.g., G2) matters were worse. Moreover, most of the groups said that they liked the idea of learning from each other’s ideas, but from the interviews there was little evidence of collaboration beyond the group level. B1 and G3 both worked on rockets, yet these groups had no knowledge of each others’ theories. G3, for instance, gave a better explanation of rocket propulsion than B1 (they had entered this explanation into the database as well). However, as I argued in the case studies, these negative findings are not surprising, given that the way the unit was started was not conducive to progressive discourse toward conceptual understanding, and neither Paul nor I recognized this and made appropriate accommodations for this weakness as the unit progressed. The LKB model may help to avoid such situations in future units.

5.5 Test Item Analyses

The test scores reveal only part of the story of student achievement that the test can tell. Analyses of several of the questions are reported here to describe achievement with respect to subject matter knowledge and knowledge building.

The sun as source of energy: During a classroom visit, B4 was discussing energy production in the sun by the burning of gases. Question 1(A&B) was designed to probe whether this knowledge had spread to the rest of the class. This appeared not to have been the case: only six students referred to the sun as making energy by burning its gases. The most dominant view, expressed by 9 of 46 respondents (20%), was that the sun makes energy from heat inside it; two students claimed that the sun gets its energy from space.
The lever: Activity 1 of the science kits (section 5.2.2) made use of a lever as an example of a 'machine', and the word 'lever' was included in the science vocabulary from the school board (Table 5.1). Question 3, answered by 46 students, probed students' use of these terms in a science-related context. The question asked for a definition of 'machine' and an explanation of how a lever, an example of a machine, works. Most students referred to the lever as a switch. In addition, 19 students (41%) gave a definition of a lever rather than an explanation of how it works.

Force, energy, and power: Question 4-A asked students to explain to a Grade 4 student the different meanings, for scientists, of force, energy, and power; it was answered by 54 of the 57 students who wrote the test. The responses were classified using the scheme shown in Table 5.4. For instance, in the following response energy is associated with force and power, and force with a push or pull:

Energy can be turned into force and power. Force is what many things use to push or pull. I think that there is no difference between energy and power.

The expected interpretation of force as a push or pull and as causing motion was common. Students used the words force, work, and power rather indiscriminately, but with force and energy they also seemed to associate meanings useful in scientific explanations, e.g., for energy, that it enables work. Some students employed a political or interpersonal meaning of 'power'.

<table>
<thead>
<tr>
<th>Table 5.4: Question 4-A on written test, force, energy, and power.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Association</strong></td>
</tr>
<tr>
<td>Force</td>
</tr>
<tr>
<td>Energy</td>
</tr>
<tr>
<td>Power</td>
</tr>
<tr>
<td>Push or pull</td>
</tr>
<tr>
<td>Creates movement</td>
</tr>
<tr>
<td>Enables work/activity</td>
</tr>
<tr>
<td>Lets things live/grow</td>
</tr>
<tr>
<td>Natural/from sun</td>
</tr>
<tr>
<td>Heat</td>
</tr>
<tr>
<td>Used to run/make things</td>
</tr>
<tr>
<td>Human relationships</td>
</tr>
<tr>
<td>Made by human beings</td>
</tr>
</tbody>
</table>

*Note: Answered by 54 students.*
Question 4-C, answered by 51 students, asked about a flying golf ball and touches on the medieval impetus explanation, in which a quantity of motion, *impetus* (‘impulse’ in modern terms), is stored in a body and expended as a body moves, none remaining when it has come to rest. McCloskey (1983) concluded that first year students at Johns Hopkins University provided explanations that conformed with impetus theory. Six students who responded (12%) made statements indicating the *continuous force* *p-*prim (diSessa, 1993, p. 218). In addition, 17 (33%) made statements indicating that the force was imparted at the hit from the golf club, without making clear if this force was stored in the ball and depleted during its motion. Finally, 8 students (15%) made statements that indicated a correct treatment of force, such as “... that happens because the energy that is put into the ball by the force of the club runs out and gravity pulls the ball down.”

The next entry of a discussion note: In Question 2, students were presented with the first and second entries of a discussion note, and asked to make a third. The information presented to them was:

Grade 5 students in a CSILE database have started the following discussion note:

MT: Energy can neither be destroyed nor created.
INTU: That cannot be true because I have read that we must conserve energy or we'll run out.

Most of the 54 responses were declared to be My Theory (MT) entries by the students (76%), and the number of responses with no thinking type was also substantial (20%); the remaining 4% were New Information (NI) and I Need To Understand (INTU) entries. The Thinking-Type-In-Use scale was used to rate the responses (section 5.2.3.5). The results are shown in Figure 5.3. A total of 56% of responses were judged to be opinions rather than theories supported by a reason or evidence. On the other hand, 39% of responses were critical of at least one of the two initial entries of the discussion note. There was surprisingly little question-asking (4%).17

17 The lack of questioning may have resulted from the way the task was set. Originally the question had been “what would you add to the discussion” or something of that nature, but after discussion with John it was changed to the form used. In retrospect, “What do you think” seems to elicit a theory more than it would a question.
Figure 5.3: Percentage of responses to Question 2 of test that are of each of seven Thinking-Types-in-Use.
Op — opinion; C-Op — critical opinion; C-ST — critical science theory; ST — science theory; C-MFT — critical mixed framework theory; MFT — mixed framework theory; INTU — I Need to Understand.

Logical Argumentation: Question 4-B asked if students thought gases are affected by gravity, and if so, why. Of the 41 arguments, 16 (41%) concluded that gases are affected by gravity. Of these arguments, 62.5% were correct (Type IV, Figure 5.4); most gave as evidence the existence of atmospheres, such as “Yes, because if there was no gravity then there would be gases floating around everywhere.” Such statements may be anthropomorphic (Hakkarainen, 1998), but I have interpreted them as ‘What if ...?’ reasoning: if gravity did not act, this would be the consequence. Of the arguments leading to the incorrect conclusion that gas is not affected by gravity, 48% started from a true premise but used illogical argument (Type II); the converse (Type III) was true of 28%.

Summary: These test item analyses revealed a relatively bleak picture of achievement, also indicated to some extent by the case studies. Many students said that most of our energy comes from the sun, but few could articulate the origin of that energy. Thus the AAAS benchmark (#2 in section 5.2.2) on this topic was only superficially met.
Similarly, students made little reference to Activity 1, in which they had investigated the effect of the lever arm on the effort required to move an object; the majority of responses referred to levers on machines, not as machines. Perhaps most troublesome about achievement with respect to subject matter was the lack of differentiation between force, energy and power. For physicists, 'energy' describes, in part, the state of an object (or the system), and 'force' interactions with the environment; power is rate at which a force does work on the object or its environment. The lack of differentiation of terms was evident in several parts of Question 4, including the one that probed if students would use the impetus explanation.

The results appeared to be better when the conversation was focused on for knowledge building aspects of the students' inquiry. Although many students mistook unsubstantiated claims for theories, there were many responses that took issue with the first two entries of the discussion note provided in the test, indicating that they paid relatively close attention to the text of these entries and may have had intent to improve the knowledge they represented. Similarly, the

---

Figure 5.4. Types of arguments for students who answered yes and no to the question, "Is gas affected by gravity?"

*Type I* — invalid premise, invalid conclusion; *Type II* — valid premise, invalid conclusion; *Type III* — invalid premise, valid conclusion; *Type IV* — valid premise, valid conclusion.
arguments provided in connection with the hypothesis that gases are affected by gravity included 25% correct arguments, with no substantial evidence for anthropomorphism. Moreover, more than 30% of respondents who came to the negative conclusion, did so on the basis of logical argument.

5.6 Concept Articulation in the CSILE Transcript

The remaining analyses are based on the CSILE transcript, and are essentially analogous to their counterparts in study 1.

5.6.1 Scientific terms usage

The use of scientific terms in the database was examined in analogous fashion with study 1; a list of terms was compiled from the Hewitt (1987) glossary and the vocabulary provided with the science kits. The use of the resulting 70 terms was analyzed in 20 randomly selected notes (the same as the total number in study 1). The following variables were examined: term frequency, sentences, term density, and term working. Term frequency and term density were significantly lower than in study 1, but the effect sizes were small in both cases — .23 and .25, respectively. The mean number of terms per note was 1.08 (SD = 1.28) and the mean number of terms per word in the note (the term density) .056 (SD = .068). The number of sentences per note and term working (the number of terms divided by the number of unique terms) were not significantly different from study 1, p > .30, although term working was .44 standard deviation lower (M = 1.80, SD = .87). Thus, there were only small differences between the two classes as far as these aspects of how students used the terms is concerned. The largest effect was that in study 1 students appeared to work the terms more.

Which terms were used is more informative than how they were used in notes. The terms used in at least five notes are shown in Table 5.5: they cluster around understanding energy in the universe, rocket motion, and friction. To

---

18 Using Mann-Whitney tests. For term frequency, U (N = 141) = 1959, p < .03; for term density, U (N = 134) = 1880.5, p = < .02.

19 Sixteen occurrences of ‘work’ were ignored. These all occurred in sentences such as ‘How does X work?’ or ‘Thank you for giving us something to work on’. Sixty occurrences of ‘rocket’ were also
some extent, the class went in a direction that the Phase 1 activities did not anticipate: sun, earth, space, Big Bang, and universe could be concepts central to a unit on physical aspects of the universe. Striking about Table 5.5, in contrast with Table 4.4, is that there are very few terms in the subject matter domain (cf. section 5.1). For instance, a connection between force and motion was almost completely absent. Motion, air resistance, acceleration, escape speed, and kinetic energy were not used at all, and speed only on a few occasions (too few for the term to be included in Table 5.5).

<table>
<thead>
<tr>
<th>Term</th>
<th>Frequency</th>
<th>Term</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>power</td>
<td>42</td>
<td>universe</td>
<td>15</td>
</tr>
<tr>
<td>sun</td>
<td>36</td>
<td>atmosphere</td>
<td>10</td>
</tr>
<tr>
<td>friction</td>
<td>32</td>
<td>heat</td>
<td>7</td>
</tr>
<tr>
<td>earth</td>
<td>22</td>
<td>nuclear power</td>
<td>6</td>
</tr>
<tr>
<td>space</td>
<td>21</td>
<td>fuel</td>
<td>5</td>
</tr>
<tr>
<td>Big Bang</td>
<td>16</td>
<td>spaceship</td>
<td>5</td>
</tr>
</tbody>
</table>

A Newtonian understanding of force and motion seems difficult without invoking such terms as mass, newton, weight, action, and reaction; but these terms were not used. In sum, the analysis of the use of these terms reveals that there was little writing by the class that included vocabulary that could become scientific. The quantitative measures showed that this was not likely to be due to aspects of how terms were used.

5.6.2 Distinguishing energy, power, and force

As in study 1, the database was partitioned into T1 and T2, with T1 accounting for the first 1/3 notes (see section 4.2 for details). Notes that contained any two of ‘energy’, ‘force’, and ‘power’ were retrieved using a NUD*IST search, see Table 5.6. All groups except G2 and G3 contributed to the retrieved notes.
Table 5.6: Articulation of the relationship of energy, power, and force

<table>
<thead>
<tr>
<th>T1 (n = 8)</th>
<th>T2 (n = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All energy comes from the sun.</td>
<td>12. There are different kinds of energy but they all originate from the sun.</td>
</tr>
<tr>
<td>2. You cannot destroy energy.</td>
<td>13. There is a difference between energy and force.</td>
</tr>
<tr>
<td>3. Force you use to do things and energy gives you power to move.</td>
<td>14. Force is a form of energy and power.</td>
</tr>
<tr>
<td>4. All living things need energy because energy gives them power.</td>
<td>15. You can get force in different forms just a simple push or an Atom Bomb.</td>
</tr>
<tr>
<td>5. There is a big difference between force and energy because force is what energy does; you couldn't have force if you didn't have energy.</td>
<td>16. Without friction we would not have force.</td>
</tr>
<tr>
<td>6. Energy can make force.</td>
<td>17. Maybe energy created force!</td>
</tr>
<tr>
<td>7. Power could mean strength which has a connection to energy because you use strength energy is involved.</td>
<td>18. Force pushes and energy creates.</td>
</tr>
<tr>
<td>8. Power could also mean force and force also has energy involved.</td>
<td>19. Energy creates force; if there was no energy, there would be no force.</td>
</tr>
<tr>
<td>9. Force is pressure on whatever is trying to move.</td>
<td>20. Power is a strong force that energy creates.</td>
</tr>
<tr>
<td>10. The relationship between energy and power is that energy creates power even in humans because if a person eats food they have more energy so they have more power to work.</td>
<td>21. If there was no energy there would be no power.</td>
</tr>
<tr>
<td>11. A rocket is the source of energy that powers a spaceship.</td>
<td>22. To do anything you need energy.</td>
</tr>
<tr>
<td></td>
<td>23. Energy is a source of power.</td>
</tr>
<tr>
<td></td>
<td>24. There is no limit to where energy comes from; energy can come from even little pieces of grass; energy can come from people just walking around.</td>
</tr>
<tr>
<td></td>
<td>25. There is no difference in energy or power.</td>
</tr>
<tr>
<td></td>
<td>26. There is a difference between power and energy.</td>
</tr>
<tr>
<td></td>
<td>27. Power is man made and energy is natural.</td>
</tr>
<tr>
<td></td>
<td>28. When you eat food that is energy, but the body uses the food to make power so we can move around.</td>
</tr>
<tr>
<td></td>
<td>29. Nuclear energy needs the sun to work.</td>
</tr>
</tbody>
</table>

T1 Conceptions: As Table 5.6 shows, the students in T1 already appreciated the importance of the sun as a source of energy, although they tended to express it in too absolute terms (1); the claim that energy cannot be destroyed (2) is also important. Examining the list of bench marks in section 5.2.2, suggests that no additional bench marks were met at this time. Some of the notes of T1 imply that energy is logically prior to force and power (5, 6, 10). Energy is natural, force and power are more associated with strength or the ability to do work, according to the student writing.

T2 Conceptions: What changed in T2? There was an elaboration of the claim that energy comes from the sun (12). The notion of energy giving rise to force and power was, it appears, unshaken; the students had evidently been thinking
about power and energy, but they did not reach consensus regarding a distinction between the scientific meaning of these terms (25, 26). Some students saw power and force as consequences of energy, with power as a strong form of force (20). Some of the concepts were elaborated: for instance, there were different kinds of energy (12) and force (15), but the differences were not spelled out.

5.6.3 Summary
This analysis agrees with the test item analysis of the previous section that showed students differentiated only weakly between scientific terms like force, energy, and power. Compared with study 1, few terms from the science glossary (Hewitt, 1987) were used, despite the fact that the students had been given the vocabulary they were expected to learn at the outset, with instructions to add keywords to their notes. Table 5.6 is noticeably less rich than its counterpart in study 1, Table 4.5.

5.7 Discussion Notes, Questions, and Comments

5.7.1 Discussion notes
The analysis of discussion notes of study 1 had to be modified for the present study. First, the discussion notes, according to Paul's instructions, started as discussions to be shared by the small groups (although with time there was some writing in the notes of other groups). This has important implications. For instance, the number of discussion notes would be expected to be smaller, as would the diversity of ideas that could be brought into contact with each other. In addition, no learner type variable was used in this study, so that an analysis such as that of contribution mix in study 1 was inappropriate. Instead, the number of groups that contributed to a discussion, a measure of the extent to which the discussion became known throughout the database, could be examined.

The students wrote 14 discussion notes with a total of 164 entries; these accounted for 66.1% of the database entries. The average discussion length was greater than in study 1, $M = 11.7, SD = 7.2$, compared with $M = 6.19, SD = 5.48$. 
There were fewer notes with less than five entries, 33.3%, compared with 48.1% in study 1. The proportion of entries written by students not in the group who created the note was 34.8% (SD = 16.8%); and, on average, 2.6 (SD = .7) of 8 groups participated in a given discussion. Several students, however, including Susan, contributed to discussion notes created by most of the groups. The largest number of groups involved in any discussion note was four groups, half the class. Of course, students may have read notes without adding entries, but this is not likely to have been a large effect. The percentage of notes opened in Paul's class, at the time of the Hewitt and Webb (1995) study, was well below the 21% quoted in chapter 3 for John's class of that year. That statistic is not likely to have changed in recent years, as Paul did not adopt Knowledge Map or introduce other strategies to improve individual students' familiarity with the database. It is also possible that insights gained in the discussion notes were transmitted by face-to-face discourse.

The discussion notes had a somewhat different character than in study 1, due to the fact that most began with a list of, for most groups, 5-10 questions in a single note entry. One note is discussed here in some detail to illustrate this difference. The note, What is energy? was created by Susan (B1), and began with the following questions in entry 1 (question form ratings are indicated in parentheses, see section 4.2.3.4 for a description of the scale):

- How do humans capture energy? (1)
- In what ways are human beings a source of energy? (2)
- What is the most recent source of energy discovered by humans? (2)
- Does all energy originate in nature? (3)
- Can humans manufacture energy? (3)
- Why is energy useful? (1)
- Why is solar energy not used as often as electricity? (1)
- Do you think the world would be better without man-made energy? If so, then Why? and if not, Why? (1)
- What's the difference between, energy and force? (2)
Figure 5.5. Susan's What is energy? discussion note
Shaded rectangles represent entries by Susan's group (G1), and unshaded rectangles by students from two other groups; entry #15 was made by the author. Arrows indicate connections between note entries, either physical or suggested by content. Note the saliency of entry #1.

Figure 5.5 shows the entries and their semiotic connections, as in section 4.6.2. What is good about this note (and other discussion notes in this study) is that the list of questions in the first entry generated considerable writing: 12 entries referred to the questions. Unfortunately, the writing was not sustained — only 2 of the 12 entries (7, 8) were followed up by a second entry. Of the 24 entries, I judged 11 to be mixed framework. Such notes were based on commonsense knowledge, and conjectured to be locations where shifts toward scientific understanding could begin. However, the data did not reveal such shifts, presumably due to a lack of sustained inquiry. Also striking about this discussion is the absence of questions after the first entry (except for the one posed by the author, which asked the author of entry 14 to elaborate on a point). That also occurred in at least one other of the more extensive notes (What is force? with 23 entries, created by G4).
5.7.2 Questions

There were 67 note entries with questions, or 31% of the database, the same as in study 1. These questions were written by 14 of the 30 students in the class, coming from all eight groups. While there was proportionally an equal number of level 1 questions (33%), there were more level 2 (57%) and fewer level 3 questions (10%). An analysis of homogeneity, however, showed that these differences were not significant, \( \chi^2(2, N = 265) = 3.76 \).

5.7.3 Comments

There were 12 comment notes (6% of the database, compared with 10% in study 1). These were written by 7 students (compared with 19), who were from five of the eight groups. There were fewer comments with constructive feedback, 25%, compared with 58%, and more with substantive contributions, 67%, compared with 34%. Analyses of proportions showed both of these effects to be significant, \( z(74) = 4.21 \) and \( z(74) = 4.62 \), respectively, \( ps < .001 \). The relative lack of constructive feedback suggests a lack of focus on helping others advance their knowledge, the fact that the comments provided more substantive input suggests the converse. The analysis of the use of terms in section 5.6.1 makes it doubtful that when comments provided substantive input, it was input relevant to the central concepts of the unit.

5.7.4 Summary

The initial questions received much attention from students, although mainly from students in the same group, but this was not followed up by sustained knowledge building effort — most theories only saw one follow up entry. Nevertheless, this phenomenon reveals the dynamic the LKB model stresses in Phase 2 of generating diverse ideas and working these into an overall research program. In this class, no such research program was formulated, and the groups basically functioned as micro-communities (Woodruff, 1995; Woodruff & Meyer, 1997). The frequency of questioning was the same as in study 1, but the class was less able to generate the kinds of questions it was argued in chapter 4 are needed for hypothesis testing (level 3). Thus, even at the level of micro-communities, fewer good research questions were asked than in study 1.
Only a few students wrote comments, but those who did, wrote comments that had more substantive input.²⁰

5.8 Discussion

This study has continued the investigation of several aspects of knowledge building relevant to the LKB model. But while in study 1 the standard for evaluating the discourse was acceptability to a subject matter specialist, in this study it was curriculum coverage. Together, the two studies provide a picture of what was and was not working in these classrooms, and can guide further development of the model and its implementation. In addition, the study concerns a model B teacher, rather than model C teacher (Lamon, Lee, & Scardamalia, 1993). Finally, the main data sources were different; it was interesting how the test, interviews, and conclusion notes revealed diverse insights about the class’s performance.

Most groups made little conceptual progress during the unit. The best explanations — in the interviews as well as conclusion notes — were in terms of surface (i.e., observational) features of the domain. Some groups not discussed in the case studies could not demonstrate much subject matter knowledge in observational terms: G2 could tell about friction only that it is caused by two objects rubbing against each other. A longitudinal study by Hakkarainen (1998) that involved three years of earlier classes taught by John and Paul also revealed that Paul’s classes produced more explanations in terms of surface features of the domain than John’s class. Hakkarainen attributed this to the teacher, but in this study I conclude that the nature of the activities — in this case work with the science kits — may also play a role in mediating progressive inquiry. The focus on surface features is common in education (Bereiter & Scardamalia, 1993, 1996a) and has been reproduced in other CSILE classrooms; for instance, early results of a four-year longitudinal study of Finnish students indicated “fact-based epistemologies” (Hakkarainen, 1998, p. 272; Hakkarainen, Järvellä, Lipponen, Lonka, Lehtinen, 1996; Lipponen & Hakkarainen, 1997). And, there were many

²⁰However, as the analyses of section 5.6 indicate, the substantive contributions were often not focused on concepts.
similarities between the work of students with LFC achievement in study 1 and the present study.

Another feature of the discourse was the low volume of writing. While I concluded from study 1 that writing output of students with LFC achievement was an important factor in lower knowledge construction, and that it should be improved by several hundred words per student, in this study the mean number of words was only 280, or 45% of that for the LFC achievement level in study 1. Moreover, 280 words over 13 weeks of work on CSILE amounts to only 22 words per week. Using the quantity of writing measure of one word per minute on CSILE reported by Scardamalia et al. (1992) (see section 3.2.2), that reflects 22 minutes of work on CSILE per week. Of course, there are good reasons why the class might only use CSILE a few times in some weeks, but the same measure, based on 10 weeks of work on CSILE, came to 63 minutes per week, for students with LFC achievement in study 1. If this estimate of computer usage is correct, there is room for improvement.

There were a number of gender effects that should be investigated further. According to the discussion in chapter 3, girls participated considerably less in a class meeting about the project than boys did, although they did make important contributions when they spoke up. For instance, Susan's question, #7 in the first interview excerpt in section 3.3.2, "Has gravity remained the same throughout history" would be recognized by philosophers of science and scientists as a significant question. Domination of classroom discussion by boys is common in elementary schools (Gallas, 1995; see also Kahle, 1988). Gender effects in the present study were of a different sort, however. Since all groups were single-gender groups and the discussions were to a large extent confined to groups, there was no way to examine the domination of boys in discussions. However, boys gave performances substantially above average more frequently than girls with respect to three measures: the number of notes, and the structure and content of the conclusion notes. Hakkarainen found related gender differences, both in John's classes and in a Finnish classroom (1998; Hakkarainen et al., 1996); for instance, in John's classes, boys were more likely to "truncate" a discussion before the first follow-up question. Research in Paul's classroom by Andrew Cohen (1995) showed the disparity between high- and low-contributors to be less in the CSILE- than in the enhanced face-to-face condition (see section 3.2.4 for
further details of this study). Cohen's study, however, did not examine gender effects specifically. The results suggest that gender effects such as have been reported here are less dramatic with CSILE than without it. That interpretation is consistent with previous studies which have shown such effects to be smaller in electronic discourse media (Harasim, 1986, 1990).

Paul had agreed at the beginning of the unit to the idea of introducing benchmarks into the database, and attempting to get the children to work toward meeting them. The procedure for doing this was described earlier in the chapter. At the time the study began, no benchmarks were available for the relatively new Ontario Common Curriculum (Ontario Ministry of Education and Training, 1995), which is why the AAAS (1993) benchmarks were used. However, the general framework for the AAAS benchmarks was very similar to the one for the Ontario curriculum, and included attention to the nature of science and Science-Technology-Society education. Paul knew neither the AAAS framework nor the most recent version of benchmarks developed by the school board; he said he did not have time to study them (interview, Feb. 20, 1997). Thus, while Paul may have had an interest in seeing the children researching their own questions (cf. section 3.3.2), he seemed less committed to matching benchmarks. During our interview, he had also made clear that, in his experience, parents were more concerned about reading and writing than about science. In addition, it became clear during classroom visits that Paul had limited subject matter knowledge, and he did not contribute notes to the database or read many notes. In Paul's class, it seems, the work on CSILE was not very central to the education Paul wanted to provide, in comparison with the development of basic skills. There is considerable evidence and argument that inquiry-based learning is not effective when learners are left on their own to pursue their inquiry (Brown & Campione, 1996; Lampert, 1995; Hintikka, 1982).

In this unit, students worked in small groups; in agreement with Phase 1 of the LKB model, they began by focusing on getting enough ideas for research into the database. The most striking effect of this particular way of organizing the activities in CSILE was revealed in the structure of the discussion notes. While most notes in the whole-class discussions in study 1 tended to be linear chains of entries (see also van Aalst & de Jong, 1997; Hewitt, 1997), in this study, many notes were reactions to one of the questions asked in the first entry, with only a
few having a second follow-up entry. Such an effect, if intended, could be an indicator that Phase 2 of the LKB model was working effectively, that is, if the group authoring the discussion note discussed the various questions in face-to-face discourse and selected a small subset from them for further research. However, there was no evidence available to suggest that this was the case. Generally, the discussion notes suggest there was a lack of sustained lines of thought. A second characteristic feature of the discussion notes was that there were very few entries by students outside the group who created them.

The three main probes used in this study — the conversation, the test, and the conclusion note — each had strengths and limitations. Taken together, they revealed a complex classroom with evidence for students who, at the time of the unit (a) were characterized as shy (G3), (b) essentially avoided the unit (B3), (c) were academically weak (G2), and (d) avoided writing, but had considerable science knowledge (B2). These findings have important implications for evaluation and assessment. Here I focus on the conclusion notes. In section 4.9, I argued that the mixed framework notes contained an internal tension and that factual notes, which report factual information without such elaborations as examples and statements of the limits of applicability of the reported scientific fact, lack this tension. I also argued that factual notes could be produced by having students extend their notes, that is, by having them explain the fact in their own words, give an example, relate the fact to other knowledge of the group/class, and so on. In the present study, the analysis of the conclusion notes revealed that students had much difficulty writing expressive notes. As study 1 did, this study therefore points to a need for attention by the teacher to writing at the note level. That is, between projects, there should be reflection on past performance, and goals should be set for writing notes that are more capable of generating discourse than was the case on the previous project. That is a place for teacher-directed activity.

5.9 Concluding Remarks

This study has provided a rather negative account of a unit in which CSILE was used. However, that makes the study valuable from the perspective of
informing continual improvement. It is the purpose of the final chapter of the dissertation to examine, and possibly elaborate, the LKB model in light of the findings — positive and negative — of both of the studies, as well as to articulate problems of understanding that, if researched, may be expected to yield improvements in knowledge building achievements in diverse contexts.
CHAPTER 6

IMPLICATIONS FOR CONTINUAL IMPROVEMENT OF KNOWLEDGE BUILDING PRACTICE

In section 2.5, I proposed the LKB model as a model for guiding curriculum development and continual improvement of teaching and learning practices, with an aim to align these with knowledge building epistemology (see Bereiter and Scardamalia, 1993, 1996a; Scardamalia & Bereiter, 1996). In this final chapter, I examine the model in the context of the studies (section 6.1) and other research (section 6.2). This is followed by the presentation of an elaborated version of the model (section 6.3), and by a brief discussion of the role of the teacher (section 6.4); finally, I indicate some possible directions for further research (section 6.5).

6.1 Summary and Discussion of Results

In the presentation of the LKB model in chapter 2, I related the features of the model to the main points of the chapter. In this section I elaborate that frame, adding how the findings of the studies lend support to the model. Some of the findings lend support to the model because they refer to experiences or types of achievement that, from the perspective of a teacher and subject matter specialist, should be replicated in future units; others lend support because they suggest areas for research and improvement.

6.1.1 Theoretical underpinnings of the LKB model

In chapter 2, I argued that children, when they begin to study an area of science formally, use theories that work well for them, but that explain relatively little from our perspective as educators and subject matter specialists. According to diSessa (1993, 1996), such theories can be represented as a rather loose collection of p-prims. By means of research, and treating the theories as objects of scrutiny and debate in Popper's world 3 (1972, see also chap. 2), students can develop more comprehensive
explanative systems. These might be represented by PDP networks of p-prims and connections between them (see chap. 2 for details), so that the same collection of p-prims can account for a variety of explanations, given particular input data (including clues about what sort of explanation is needed in the given context). In this framework for understanding learning, commonsense explanations play an important role — they are not to be replaced or eradicated but become the building blocks of a more powerful explanatory system (see also Hills, 1989; Lonergan, 1957). Of course, building up this system also requires new learning of practices and ideas that do not have counterparts in everyday discourse. Roth describes conceptual change as a transformation of discourse practices (Roth, 1998; Roth, McGinn, Woszczyna, & Boutonné, in press). In science, words have specific meanings that may be at variance with meanings they have in everyday conversations. Sometimes, the required distinction is ontological (Chi, 1992), but more often than Chi seems to have acknowledged in her theory, the distinction is semantic (Klaassen, 1995). Other aspects of learning scientific discourse are appropriate use of "inscriptions" (Latour, 1987; Roth & McGinn, 1998b); using increasingly rigorous standards for evidence (Roschelle, 1996); and developing an increasing inclination to formulate testable questions and theories.

As explained in more detail in chapter 2, the LKB model is based on developments in epistemology in the last four decades: Lakatos's progressive research program (Lakatos, 1970) is central to the model as a whole, and the idea of continual improvement; and Kuhn's (1970) emphasis on scientific knowledge as validated by a community of scientists (i.e., a community of practice), is featured in Phase 3. From an educational perspective, the model attempts to address Hodson's (1990) call to pay more adequate attention in science education to doing science; one must begin from ill-defined rather than fully articulated problems of understanding (Roth, 1995; Roth & Bowen, 1995). Moreover, the model takes the perspective that these problems of understanding must be "constitutive" problems — problems that can advance the current limits of understanding of a community (Bereiter & Scardamalia, 1993). From the perspective of these philosophical underpinnings, the purpose of Phases 1 and 2 is to articulate these problems of understanding, and from them, a research program. The distinction between conjectures and hypotheses in Phase 1 (see Figure 6.1) stresses that the problems raised are at first ill-defined, and must be refined. A variety of modes of inquiry in Phase 1 are meant to illustrate progressive problem solving — e.g., doing an experiment can lead to questions to
study in the library, or to modeling. In addition, the model is intended to be a scaffold for a progressive research program in which the aim is to improve the educational value of the experiences created for and with students, from one unit to the next. As I argued in chapter 2, the teacher has an important role in Phases 1 and 2: to ensure that the students develop a research program for the unit that is based on worthwhile problems of understanding related to the subject matter, and to ensure that the larger continual improvement agenda is also advanced.

I also argued that the so-called learning paradox (see chap. 2; Pascual-Leone, 1976) presents a challenge to developing worthwhile research programs, as well as to the validation of knowledge advances in Phase 3. The two-part paradox is that if the views considered are not sufficiently diverse, constitutive problems may not result from the discourse, but if, on the other hand, they are too diverse, no one may be able to see coherence in the ideas that are being considered. Phase 1 attempts to address this problem by ensuring that a sufficient range of phenomena and ideas is considered by the class early in the unit. Finally, I found the idea of M capacity useful in constructing the LKB model: a group of students working with a technology to advance each student’s understanding of (communal) problems of understanding has more information processing capacity than an individual student. (The LKB model is such a technology.) Admittedly, the potential increase in information processing capacity could be undermined by the nature of the social interactions, which remains a topic for further study.

6.1.2 Studies 1 and 2 and the LKB model

6.1.2.1 Phases 1 and 2.

In Figure 6.1, Phases 1 and 2 are annotated with the most important theoretical issues that the model should address in this early part of the unit, and relevant findings from the studies.
In study 1, I classified student achievement on the basis of a focus on concepts. Students who wrote explanations that were predominantly rated as scientifically acceptable (S) late in the unit, wrote more "Mixed Framework" explanations (MF) early in the unit than other students did. This finding suggests that making commonsense knowledge explicit early does not impede the writing of scientific explanations later. If this assertion is correct, it is an important finding, because there is a concern among researchers that students, as a result of their constructive activity, "invent misconceptions" (Brown & Campione, 1996, p. 298). While "misconceptions" may remain hidden after traditional instruction, constructive activity often has the effect of explicating them. The finding agrees with previous research on knowledge building (Chan, 1993; Chan, Burtis, & Bereiter, 1997; Hakkarainen, 1998). Burtis, Chan, Hewitt, Scardamalia, and Bereiter (1993) concluded from examining the frequency of misconceptions in a student-generated
database on physical forces that students did produce misconceptions, but these were interpreted by students as facts to be explained, rather than theories; when misconceptions were stated as theories, students often made progress toward resolving them.

An implication of this study 1 finding for improving knowledge building practice is that it is probably a good strategy to attempt to increase the amount of writing early in the unit — in Phase 1. This requires that the teacher monitor the writing output of individual students. It would also be valuable to conduct an experimental study to clarify this effect. Does making commonsense knowledge explicit facilitate the writing of scientific explanations, or is the effect predominantly due to the volume of writing output? The factors that contribute to students’ willingness and ability to make their knowledge explicit early (e.g., factors attributable to subject matter knowledge, socio-economic status, and emotion) also need to be better understood. These questions could be investigated by a class by attempting to increase their writing output with an aim to informing these questions. Study 2 was less informative about the effect of writing output on knowledge at the end of the unit, although the number of notes tended to correlate positively with test scores and the number of knowledge claims in the interviews (see Table 5.1).

In study 1, even students with HFC achievement were unsuccessful with some topics (see Figure 4.2); I attributed this to a lack of subject matter knowledge with respect to the topic at hand. Of course, the role of subject matter knowledge in conceptual change is well known (Carey, 1985), and numerous writers have advocated that the teacher should be aware of the initial subject matter knowledge of his/her class (e.g., Ausubel, 1968; Goldman & Bendall, 1995; Hestenes, Wells, & Swackhamer, 1992). In study 2, students I found that students contributed relatively little subject matter knowledge in the test, the interviews, and conclusion notes); in addition, some students had expressed frustration because they could not find information that they could use (class meeting, March, 1997). Thus, the introduction of resources (information and artifactual) for knowledge building in Phase 1 seems important. An important conclusion of study 2 is that introducing such resources is of little value if they are not suitably matched to the goals of the unit, or used in ways that undermine their potential value.
Equally important, it seems, is how the teacher structures the class’s work with CSILE. No causal inferences can be made from the data of study 2, but it can be argued that Paul’s approach was less aligned with knowledge building epistemology, and may have undermined the success of the unit. I have already discussed Paul’s approach to using CSILE in chapter 3. The point that is to be made here is that just as the teacher must be able to anticipate the conceptual territory all students should learn, so should he/she be able to anticipate the implications of the way activities are structured. This requires knowledge of the epistemology and sociology of knowledge building. I did not make this issue sufficiently clear in the LKB model as presented in chapter 2, but if we amend Phase 1 (see Figure 6.1) to include a component that focuses on knowledge building aims, it becomes clearer that each unit has goals with (at least) two temporal scales (see also Lemke, 1995; Roth, 1998): (a) to develop understanding of the subject matter at hand in the unit, and (b) improvement of knowledge building practices (e.g., questioning, commenting, and summarizing) in the context of a progressive research program that spans at least several units.

The studies also revealed several issues that lent support to Phase 2. In study 1, the analysis of discussion notes showed that learners with low focus on concepts achievement (LFC) tended to contribute less to notes created by learners with high focus on concepts’ (HFC) achievement; they also wrote fewer mixed framework notes early in the unit. It is likely that the first effect can be explained, in part, by the second. There was some suggestion of stonewalling by learners whose achievement was classified High Focus on Concepts (HFC). There also was evidence that these learners were more successful at generating hypotheses than students with LFC achievement (study 1, see Figure 4.5); they also tended to be more successful at this than the class of study 2. These findings suggest that there may be social conditions in classrooms that lead to differential opportunity for knowledge building activity. Such conditions should be studied with an aim to reducing the noted inequities. Phase 2 makes explicit that questions and ideas must be elaborated (see also Hakkarainen, 1998) before they can form the basis of a research program. It is therefore a suitable place for supporting students who need it with the necessary elaboration processes, and is a natural place to study social conditions.
6.1.2.2 Phase 3.

In Figure 6.2, Phase 3 is annotated with the most important theoretical issues that the model should address in this part of the unit, and relevant findings from the studies.

![Figure 6.2: Support for Phase 3 from theory and studies](image)

For most of Phase 3 students work in micro-community, but consolidating a macro-community research agenda and validating results in the macro-community can be competitive. Study 2 showed a lack of finding others on the same research topic; in the studies there were good examples of summarizing of knowledge advances as well as the of lack thereof.

The cases of study 1 demonstrated articulation of personal research agendas and helping others advance understanding. (The cases of study 2 were based on interviews which lasted only 23-35 minutes, and were less capable of providing this...
kind of information.) Some students, like Matt and Sean, worked their ideas in diverse contexts; whenever they met new situations, they would try their idea out on them. We saw Matt become interested in the role of density in explaining the effect of gravity on gases, Sandy explain that “when a mammoth gets hot, its molecules shake and it gets tired” (cf. section 4.4.4.3), and Sean that “goosebumps are molecules that need more space to move faster” (line 19, Appendix A). Reading the database, one can imagine Sean, talking with his grandmother, trying out his molecular model on almost any subject, such as the flow of syrup over his pancakes. Clearly, many such attempts overestimate what a molecular model can explain, but each time the model is applied, insight into what it can and cannot explain becomes more fully elaborated. It is possible that isolated statements, such as Sean’s about goosebumps, are examples of trying out an idea rather than reflecting an established belief. This sort of activity is what is at the heart of the idea of sharing knowledge — others are helped with the problem they are working on, but the student who does the trying out also learns from it.

Finally, in neither of the studies did the students engage in experimental work after the science kits. The work with the science kits provided students opportunities to develop an empirical base from which to theorize; later in the units, theorizing was more directly related to previous argument, reading, and classroom discussion. What was missing in these units, however, was theory-driven empirical work. While this is not likely to be a problem if it happens occasionally, a knowledge building culture that does not include empirical work to test emerging ideas and questions would be a distortion of science (Hodson, 1986, 1990, 1996). There was a point in one of the discussion notes of study 1 (see section 4.6.1), where I thought that the teacher should have done more to encourage the students to design an experiment. At line 7 for the discussion note, Is gas affected by gravity? (see section 4.6.1), Jerry asked:

Does gas weigh anything? What would happen If you had put some sort of gas in a box and the box could contain it. What would be the effect? Would the gas weigh anything inside the box because the gas is contained?

For me, this episode was what ethnographers call a ‘rich point’: a salient feature of the data that drives theory development. First, the questions, about air pressure in a large context of understanding the force of gravity, are conceptually important. I was reminded of Minstrell’s (1989) classroom discourse. Second, the questions were
historically significant (Bernal, 1971). Kuhn (1970) argued that children should be displaced historically, so that they can see themselves as contemporaries of the contributors of important scientific advances. From my interaction with other researchers on the CSILE Project, I had become aware of some examples where students had come to see themselves in this light. One student remarked, for example, "that Mendel had worked on Anna's problem." Third, progress on the questions would have required considerable thought on experimental method (some development of experimental ideas would be necessary). Fourth, students would be able to learn about the epistemological role of empirical work in science. Fifth, if all of these elements came together — understanding of conceptual and historical significance, subtle experimental design, and epistemology — the children would, in my opinion, have had the science education experience of a lifetime. Even if I (or another subject matter specialist) had acted as a mentor in an apprenticeship relationship, we might not have made much progress. I am not arguing that Grade 5/6 students should experience all that I have described here. What I am arguing is that we should aim to develop knowledge building communities that: (a) learn to recognize situations rich in knowledge building potential better, and (b) consider empirical work as a cultural tool, to be used whenever the discourse requires it, and not only at the beginning of a unit.

6.2 Comparison with some other approaches that stress inquiry

In this section, I compare the LKB model with two other approaches that stress cognitive aspects of inquiry: (a) Fostering Communities of Learners (Brown & Campione, 1994, 1996), and (b) ThinkerTools (White, 1993a, 1993b; White & Fredericksen, 1998). Other open-ended inquiry approaches (e.g., Roth & Bowen, 1995) also appear to fit with the model but stress sociocultural aspects of knowledge building more than the two approaches I discuss here.

6.2.1 Fostering Communities of learners

The Fostering Communities of Learners (FCL) approach of Brown and Campione (1994, 1996) is grounded in ideas from several research traditions: (a) reading and reciprocal teaching (Palincsar & Brown, 1984), the Jigsaw strategy (Aronson, 1978),
Jerome Bruner’s spiral curriculum (1969), Vygotsky’s (1978) Zone of Proximal Development, “consequential tasks” (Scardamalia, Bereiter, & Fillion, 1981), and the use of subject matter specialists. Reciprocal teaching was an approach in which peers took turns at the teacher role, as small groups of students attempted to comprehend reading material; in the FCL approach that technique has become exploited more fully. A topic of “enduring appeal” is selected that is of sufficiently large scope to allow each of about five groups of students to specialize in a subproblem. The Jigsaw technique is used, but from time to time there are whole-class activities to allow the whole class to learn what progress has been made on the problem. Brown and Campione (1994) called this “cross talk.”

Many of the components of the FCL approach have analogs in the knowledge building approach. Reciprocal teaching, for instance, is mediated by writing commentary notes, and notes aimed at summarizing what has been learned. Both approaches are committed to allowing students to “major,” that is, to develop specialized knowledge of a small area of the subject matter, which is to be shared with the rest of the class. Consequential tasks have not been stressed in the LKB model, but they are consistent with it. A consequential task that would seem to be suited to the goals of the Knowledge Society Network (Scardamalia & Bereiter, 1996) is making artifacts available to other learning communities, so that, for instance, next year’s class can examine what a previous class has already found out about a topic they are about to study. In CSILE this could be accomplished by creating published notes or a portfolio notes. (A portfolio note is a note that brings together a collection of other notes, and presents a framework for understanding them.) Consequential tasks in which children have an opportunity to present their findings to third parties, such as parents and school-board representatives, can have motivational value. Berenfeld (1993) has reported how students at Pease Middle School, an inner city school in San Antonio, measured the quality of the air throughout their school. Students presented their findings at a meeting with school district officials. The school district subsequently measured the CO₂ levels with their own equipment and confirmed the children’s conclusions; this resulted in the school’s ventilation system being repaired. Pedretti reported a similar experience with Grade 6 students at a school north of Toronto, involving a “septic tank dilemma” resolved by the children (Pedretti & Hodson, 1995; Pedretti, 1997). Of course, it is important to ensure that the consequential task does not become the end of the students’ inquiry (Bereiter, 1997). Finally, the FCL idea of cross talk has analogs in the LKB model in “class
meetings," for instance, class meetings in which the whole-class research program is consolidated (at the beginning of Phase 3), or knowledge advances are validated (at the end of Phase 3).

The FCL approach is committed to a spiral curriculum. Brown and Campione (1994) argued that:

In order to be responsive to the course requirements of normal schools, we believe it is necessary to set bounds on the curriculum to be covered. In general, our approach is to select enduring themes for discussion and to revisit them often, each time at increasingly mature levels of understanding. (p. 237)

They argued that the "increasingly mature levels" should be guided by developmental trajectories of children's ideas concerning the content to be learned, adding that little is known of these trajectories for most of the science curriculum. A notable exception is Carey's (1985) study of the development of children's theories about living things.

While the FCL approach appears to be faithful to Bruner's idea of the spiral curriculum, the latter is often distorted in implementations in which one repeatedly returns to old learning, adding only minor extensions. For example, in an Ontario (general level) mathematics curriculum, 'using ratio, rate, and proportions' was an extension topic in Grades 10 and 11, and a review topic in Grade 12. The same pattern existed for 'substituting in a formula involving up to four variables and solving the resulting linear equation' (Ontario Ministry of Education, 1985). In Grade 12 physics, students were not to solve mechanics problems involving strings (Ontario Ministry of Education, 1988), which was to be left for a more advanced course. Such examples are not isolated, and it would seem prudent to explicate the sense in which a spiral curriculum is consistent with the LKB model.

According to the findings of study 1, some students could develop explanations of high scientific merit; these require only minor modifications later. What remained lacking most at the end of the unit was an integrative framework — laws within which such explanations can fit. While only a minority of students in study 1 were able to develop such high-level explanations, other studies with CSILE have presented additional evidence for such high-level explanations (e.g., Cohen, 1995; Hakkarainen, 1998; Hewitt, 1996; Oshima, 1995). It remains to be seen to what
extent such performances can become representative of the knowledge building abilities of Grade 5/6 students in learning environments similar to the one John created for his class. However, one may assume that advanced understanding can result as students continually "work" their knowledge in diverse situations, not only within a given unit, as was the case for some students in study 1, but also across units. In other words, if a student develops a reasonably accurate understanding of the concept 'kinetic energy' in a unit on force and matter, one may expect that concept to be part of the student's repertoire in a subsequent unit on heat and energy. In the latter unit, he or she may develop a deeper understanding of the concept by relating it to other kinds of energy, such as thermal energy. Thus, factors that limit levels of understanding are subject matter knowledge and the ability to formulate integrative frameworks, which is, in turn, limited by the development of M capacity (see section 2.3.2).

6.2.2 *ThinkerTools*

Barbara White has pointed out three key problem areas with learning physics: (a) acquiring abstract, generally applicable models; (b) linking formal models to real-world phenomena; and (c) acquiring knowledge of model construction (1993a, pp. 4-5). The ThinkerTools approach she developed with John Fredericksen is designed to address these three dimensions of learning physics. It uses a series of progressively more complex computer microworlds. Within each microworld, students work through a learning cycle with four phases: (a) motivation, (b) model evolution, (c) formalization, and (d) transfer. In the model evolution phase, "students collaborate to solve problems and perform experiments in the context of the computer microworld (usually with two students per computer). The purpose of the problems and experiments is to enable the students to discover the causal principles and concepts embedded in each microworld. White's hypothesis is that, through this combination of increasingly complex microworlds and focused problem solving, the students will gradually acquire the desired conceptual model. In the formalization phase, children invent laws, although initially they get considerable help (from the teacher). "For the 'incorrect' laws, they have to be prepared to prove to the rest of the class that they are incorrect" (p. 13). In the transfer phase, the objective is "to get students to appreciate how the rule that the class selected as the best law applies to real-world contexts" (p. 14). Students are initially given the experiments and problems to do; later they design their own.
A major strength of this approach is the progression from simple models to more complex models. The simpler models are usually scientifically correct, but lack explanatory power, compared with the more complex microworlds. For instance, students can learn a number of aspects of Newtonian mechanics in a one-dimensional microworld. In a one-dimensional microworld, the direction of motion is necessarily the same as the direction of an (unbalanced) applied force. That no longer is true in a two-dimensional microworld. Or one might conceive a three-dimensional microworld in which frictional forces are, for instance, only 2% of what they are in the physical world (World 3); in such a microworld one would have to bank curves in roads at a greater angle. One could then progress from such a microworld to one in which frictional effects are larger. Building progressively more complex models is central to scientific inquiry, and has much to offer for teaching about the nature of science. The LKB model has nothing analogous to a progression through microworlds, but it is consistent with it. In Phase 3, research can include a variety of activities, including model-building. Thus, a pair of students might work through a series of ThinkerTools modules in Phase 3, in an effort to answer, at progressively deeper levels of sophistication, a research question they had formulated earlier in Phase 3. A shorter series of ThinkerTools modules might be used as part of Phase 1.

The LKB model serves to underscore that the central activity of the ThinkerTools approach (i.e., working through the microworld modules) must be situated in a range of activities, such as library research, experimentation with familiar materials, and extended discourse. This appears to be the case with White and Fredericksen's work in classrooms, but one can imagine implementations in which "discovering" the causal principles embedded in the microworlds, known to the designers of the software and accompanying curriculum modules, becomes the sole curriculum. Such a curriculum would have the same philosophical problems as many "discovery learning" approaches (see Hodson, 1996). It is important to allow students, after developing extensive experiences with microworlds (i.e., after several years of working with them), to progress from using microworlds to creating their own to model the phenomena they are attempting to understand. Finally, while the intermediate level of abstraction provided by simulating the microworlds facilitates conceptual change, one may also physically constrict phenomena to progressively more complex microworlds, and use microcomputer-based laboratory tools (Mokros & Tinker, 1987; Thornton & Sokoloff, 1990) to conduct research in each physical
microworld. (This would be analogous to the study, by scientists, of optical properties of materials in *thin films*, or the calculation of the electric field set up by an *infinitely long* charged wire before treating more general cases, and can teach an important aspect of scientific research.)

### 6.3 Elaboration of the LKB model

The LKB model has several strengths. First, it underscores (a) that open-ended inquiry (discovery) is "theory-laden," and that students must therefore have adequate theoretical perspective (a "working knowledge") before they can proceed with it. Second, it is based on knowledge building epistemology. Traditional didactic approaches tend to neglect the first issue, and many constructivist approaches neglect open-ended inquiry. Third, the model explicates that attention must be given in designing curriculum units to the steps involved in developing a research program and in validating knowledge advances. And fourth, it explicates the possibility that students engage in new empirical research — which can lead to new questions — relatively late in the unit. The model is not an instructional model but, in a sense, a model for developing instructional models. The inquiry approaches discussed in the previous section are specific approaches that can be mapped onto the LKB model with reasonable success.

The model (Figures 6.1 and 6.2) also has weaknesses that need to be addressed. First, the phase boundaries are meant to facilitate discussion of component processes, not to create a rigid system of use. Their exact locations are somewhat arbitrary. Students must be cognitively active and engaged in self-regulatory process throughout the unit. At the same time, the teacher is continually engaged in helping students to become more reflective. Second, the model must say more about learning programs that are suitable for open-ended inquiry. Third, the model does not make sufficiently clear that the curriculum unit is part of a continual improvement process — each unit has not only aims related to the learning of science but also aims related to improving knowledge building practices. I discuss each of these issues briefly below.
Utility of the phases: Early in the development process of the model, I had two phases in mind, which were to characterize learning and knowledge building, respectively; the knowledge building phase was conceived as operating with a progressive research program, and Phase 2 as a transition between the two modes of activity — it allowed me to conceptualize the processes necessary for creating a successful progressive research program. With hindsight, Phase 2 no longer seems necessary. Nevertheless, the distinction between Phases 1 and 3 remains useful for considering three practical issues: (a) assessment, (b) teaching and learning activities, and (c) teacher education. Looking forward to the progressive research program, the aim of the early activities of the unit must be to develop a sufficiently rich problem space from which a research agenda can be framed. The learning program, questioning, and hypothesis formulation all contribute to that. But while developing deep understanding is key to knowledge building, I have argued throughout the dissertation that deep knowledge must be “well grounded.” Scientists must be able to relate their own research programs to their field and to science in general. Similarly, small groups of students must be able to relate their specialized knowledge to that of the rest of the class (otherwise they would not be able to use it to help others advance their understanding), and they must do justice to the topic studied. (They could develop too impoverished a problem space.) For this reason, I argued in section 5.1.1 that national curriculum frameworks (e.g., AAAS, 1993; Canadian Council of Ministers of Education, 1997; DES/WO, 1995; NAEP, 1995) can be used to develop assessment instruments to check that deep knowledge developed by students is appropriately supported by familiarity with a range of related concepts, facts and procedures. Because we are concerned with what we can reasonably expect most (if not all) students should know about the topic at hand, teachers should be comfortable with the subject matter at this level. This includes knowledge of the conceptual and phenomenological terrain of the topic, and familiarity with difficulties students may have with the subject matter. For instance, the teacher should know that students at a particular grade level may not differentiate between ‘force’ and ‘energy’, tend not to associate agency with inanimate objects (diSessa, 1993; Klaassen & Lijnse, 1996), or may think that part of the image of an object will disappear if part of a lens is covered (see Heron & McDermott, 1998). I see this as a matter of teacher education (pre-service and in-service). To sum up, early in the unit (i.e., in Phase 1), the teacher controls the curriculum, but the students are cognitively very active, and are the forming conjectures and questions that will become the basis of their research program; we
also see that we have a triangular structure of curriculum planning, assessment, and (minimum) teacher subject matter knowledge.

When the research program begin to take shape, things look quite different. The students take charge of setting goals and designing the research program; the teacher assists with clarifying goals, when necessary. When the class's research program is consolidated, students become accountable to each other, as well as to the teacher, for the problems they agree to research. They can go up blind alleys and pursue new questions (e.g., more interesting ones), but they must be able to explain to the class why they did not do the research they proposed to do, and what they learned from the research they did do. The teacher can no longer have answers to all the students' questions, in all subject areas, but this is not necessary or even desirable. The teacher's role becomes one of supporting students while they formulate research questions, conduct their research, and argue about findings. To do this, I conjecture that the teacher must become a connoisseur (Hodson & Bencze, 1998). This entails having available a variety of strategies for finding information, understanding knowledge building epistemology, and understanding the social dynamics of computer-supported knowledge building, and can only be built up by reflective practice. No amount of training can prepare a teacher to become a connoisseur of knowledge building; nevertheless, teachers can develop connoisseurship by practice and use it as a basis for judging the class's knowledge building performance holistically.

The learning program: The learning program must address both the need to prepare for developing the research program and the need to ensure this is well grounded. As we saw in study 2, knowledge building goals can be undermined by how the learning program is used to start the unit; in addition, the model stresses question asking in Phase 1. Minstrell and Stimpson (1996) argued that questions are best rooted in shared experiences, and Arons (1990) against introducing scientific terms before the ideas they represent are clear. In connection with this, it is perhaps worth reiterating that the questions in Phase 1 result from the learning experiences provided. Students are actively involved in knowledge construction throughout the unit, and benchmark lessons (Minstrell, 1992; Hunt & Minstrell, 1994; Minstrell & Stimpson, 1996) or guided inquiry curricula can be used (e.g., Laws et al. 1996; McDermott et al., 1997). It is important that the program used allows students to
develop a preliminary understanding of the meaning given to words and activities in the context of science (Klaassen & Lijnse, 1996).

Improving knowledge building practices. The LKB model divides a curriculum unit into component activities, which can provide contexts for attempting to improve knowledge building practices. Improved practice can be mediated by several factors. For instance:

- New technologies (e.g., discussion notes, see chap. 3).
- The teacher and/or students have gained experience with knowledge building and/or learning technologies.
- The teacher has reflected on knowledge building practices or engaged in knowledge building with other teachers, subject matter specialists, and students (as in the Knowledge Society Network).
- Changing social conditions (e.g., a better understanding of collaboration).

Some examples of efforts to improve knowledge building practice, based on the findings of the studies, and matched to activities made explicit in the LKB model, are:

- Ensuring that all students are contributing to the discourse by the end of Phase 2, by checking database use and being aware of face-to-face interactions that co-occur with database entries.
- Helping students to develop problems of understanding and hypotheses from their initial ideas by the end of Phase 2.
- Helping students with the writing process, so that they can learn to write more expressive notes.
- Testing subject matter knowledge targeted by Phase 1 instruction, before Phase 1, and at the end of Phase 2.
- Experimenting with new ways of structuring the students' work with CSILE.
- Undertaking steps that make subject matter expertise available to the class (e.g., video tapes, books, visits by/to scientists, virtual presence of subject matter specialists).
- Experimenting with new ways of reporting knowledge advances (e.g., developing the idea of writing a book about knowledge advances by Paul's class further).

Whatever initiatives are adopted at a given time, they become part of the curriculum for the unit.

Two-unit presentation of the LKB model: In Figure 6.3, I situate the LKB model in a context of two sequential curriculum units. The figure makes possible ways of using the model for guiding continual improvement of knowledge building practice more explicit than Figure 2.2 did. The presentation of the model for each unit has been
simplified in order to show several processes that were not shown before, while keeping the complexity of the diagram to a minimum. In each unit, the aims of the learning program are: (a) development of a problem space, and (b) ensuring that this is well grounded in the topic at hand and in science. I have represented the execution of the large-group research program by ‘progressive problem solving’ to stress that there can be several stages of inquiry that probe a question at progressively deeper levels. In the language I used in Figure 2.2, an experiment can be followed by further empirical work via modeling. A new feature is feedback from these activities to the articulation of the problem space. This allows for the possibility, for instance, that after a "break-through", the class re-examines what it has held to be the constitutive problems of the class, and returns to formulating problems of understanding.

In the diagram, black arrows refer to processes aimed at continual improvement. For example, checking curriculum coverage (e.g., mapping end-of-unit knowledge claims onto a curriculum document) can inform improvements to the learning program used, or of the way it was used (for the next time the teacher plans to do a unit on the same topic); at the same time, it can inform such issues for the next unit the current class is to do (unit 2 in the diagram). In addition, the class should reflect on the curriculum coverage it attained in unit 1, as well as its success with the knowledge building strategies it used, and this can inform the selection and operation of the learning program in unit 2, and contribute to the problem space for that unit. For instance, the class may conclude as a result of their experience in unit 1 that many of their comments were ineffective, and make this a problem of understanding in unit 2, to be studied in the context of the subject matter of that unit.
In each unit, the learning program is aimed at creating a problem space sufficiently rich for developing a research program and at ensuring that the research program is well grounded in the subject matter and science. Black arrows show linkages associated with the goal of continual improvement of practice; issues for improvement arising from the first unit are incorporated in the problem space of the second unit.
6.4 Teachers, Researchers, and Subject Matter Specialists

The purpose of this section is to discuss the role of the teacher in promoting knowledge building, as well as a number of mechanisms that are being created/studied to support the teacher in this (the LKB model being one of them).

6.4.1 Modeling knowledge building

Brown and Campione observed that the role of the teacher in constructivist classrooms is still “relatively uncharted” (1994, p. 230). The term ‘facilitator’ is frequently used by teachers when they describe their role in their constructivist classrooms, but it is an inadequate metaphor for what is required (van Aalst, 1995). Often ‘facilitating’ is taken to mean providing an environment in which students can conduct guided inquiry: all the effort goes into developing curricular materials that provide standard guidance (i.e., the same hints for all students), classroom procedures, technological supports, and so on. Students then work through the resulting program, and the teacher assumes the role of manager, making sure that students keep up with the schedule of activities, setting out materials for use by students, and giving assistance when time allows it. Lampert made an important point when she wrote

... irony is not lost on students who are invited by teachers, in courses they are required to take, in schools they are required to attend, to “inquire” into a subject and “discover” its truths. The invitation ... is even more complicated by the fact that students often feel that unnecessary work is being asked of them. “Teachers wouldn’t be teachers if,” they say to themselves, “if they did not know what I need to know, so why am I being put through the difficult task of finding things out for myself?” (1995, p. 215)

Although it requires much effort to be a good facilitator, most facilitators are Model B teachers; they tend not to put responsibility for metacognitive strategies in the hands of students.

What should a teacher do to support knowledge building? In the previous section, I already mentioned examples of specific teacher activities aimed at improving knowledge building practices. Prior to those, however, is the teacher’s commitment to knowledge building epistemology — a commitment to understand it better, to take an active role in fostering the infrastructure it requires, and to understand the subject matter better.
Understanding knowledge building epistemology is an active and dialectic process. It is not sufficient to "know" that children learn by doing. The teacher must also understand ways in which learning activities can enhance or undermine knowledge improvement; such understanding results from closely observing the discourse, and from teacher contributions. An example of what I have in mind is provided by an interview with John that was conducted by James Hewitt. John noted:

I think [My Theory] is really important for them because it provides a starting point, it gets them thinking in some depth about a problem. In the beginning, their theories are usually fairly brief, but when they come to get the theory published, if I can see a way that I can encourage that student to put more detail into that theory, I’ll ask them to go back and do that before I publish it. ... (quoted in Hewitt, 1996, p. 118)

Like Minstrell and Stimpson (1996), he seems to have understood this was an essential difference between knowledge building and traditional educational practices: at an address at a conference on computer-human interaction in 1992, he noted about discussion notes that

... we have never allowed students before to express [their theories]. In fact, the name of the game in school is to keep them hidden, not to bring them out in the open, not to ask in case it’s a stupid one, not to write something down in case it’s the wrong answer. Instead now with this [discussion] format we are encouraging students to be unafraid of saying what they actually believe and then work towards seeing whether or not they are correct (cited in Hewitt, p. 132)

It is not clear if John developed such understandings, in part, from his own contributions to databases. Although he did not contribute often to the database of study 1 by writing notes, database records show that he did contribute to other databases in which teachers discussed their work with CSILE. What does seem necessary is that the teacher is intellectually engaged with knowledge building — his/her class becomes a laboratory in which to develop the necessary knowledge. Like the scientist (see Latour, 1987), the teacher must work to maintain the viability of his lab, acting as a spokesperson for knowledge building, establishing necessary resources, and producing tangible performance indicators.

From the students' perspective, an advantage for students of the teacher's active engagement with knowledge building is that they can observe the teacher model knowledge building. Whether or not it is necessary that the teacher also contribute
directly to the students' inquiry is less clear at the present time, although there is some evidence from pilot studies that suggests a limited amount of teacher- or researcher presence in the database is desirable (van Aalst, 1997; van Aalst & de Jong, 1997). The teacher must carefully consider his or her contributions, because there is a significant danger of the "knowledge-telling" (Bereiter & Scardamalia, 1987a). If the teacher explains too much, students may think there is no need for them to contribute additional notes. Ideas like peripheral participation (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991) and the Zone of Proximal Development (Vygotsky, 1978) may be helpful in further conceptualization of the level of the teacher's contribution to the students' database discourse. A second problem with teacher control of the communal discourse is a potential upset to the authenticity of the discourse, especially if comments are evaluative rather than aimed at advancing community understanding of the subject matter. Also, results to date suggest that the teacher should not always be asking questions for which he or she already has answers.

6.4.2 Multicultural research

The proposal, in the previous section, that teachers should model knowledge building, can make substantial demands on teachers over and above what is typical in traditional settings. As a minimum, although substantial efforts have already been made in this direction by proponents of research by teachers (e.g., Stenhouse, 1975), and of action research in particular (Carr & Kemmis, 1986; Elliott, 1991; Hodson, in preparation), we must continue to create a teacher profession that is committed to "reflective practice." I have argued, following study 2, that it is important that teachers develop deep understanding of knowledge building epistemology, and that this is a dialectic (or "reflexive", see Latour, 1987; Vygotsky, 1978) process. Brown and Campione (1996) noted that "without adherence to first principles, surface features tend to be adopted, adapted, and ritualized in such a way that they cease to serve the 'thinking' function they were originally designed to foster" (p. 291); numerous other authors have also noted the failure of top-down implementation, when teachers do not understand the theoretical ideas behind an innovation, have been documented extensively in the research literature (see Pedretti, 1994, pp. 52-53 for a discussion).
A heuristic for sustaining reflective practice is what Brown (1992) has called the *design experiment*. It has been the heuristic of the CSILE Project, as well as the FCL approach and Roth's communities of learners (see Roth, 1998), and has become a central heuristic in the Telelearning National Centres of Excellence (NCE). The basic idea is the knowledge building heuristic in another context — a pool of talent from diverse backgrounds jointly articulates and solves a set of constitutive problems. Thus, successful educational change committed to knowledge building values requires changes in classroom practice (i.e., in our understanding of what it means to teach and assess), further elaboration of theoretical ideas, and further development of technical infrastructure. The problems are mutually dependent and must therefore be part of the same problem space. I call research in such an environment 'multicultural' because its success depends critically on crossing cultural boundaries. For instance, schooling, educational research, and engineering have very different funding structures, and hence different obligations to meet. Researchers must publish; school boards must demonstrate effective program delivery with such indicators as success rates in finding employment and in post-secondary education, and performance on standardized tests (e.g., College Board Exams, TIMMS). Educational institutions (school districts, ministries of education, etc.) do not recognize research and software design as teacher responsibilities; nor do they fund such activities. An implication of this is that if teachers are to engage in such activities, there must be changes to educational funding structures and/or reward for teachers — expecting teachers to engage in them in addition to what they are already mandated to do is not likely to be a sustainable model for supporting knowledge building.

In the remainder of this section, I describe several initiatives that begin to address a number of such cultural issues. First, teachers and researchers have jointly published or presented findings based on their collaborations (e.g., Bereiter, Scardamalia, Cassels, & Hewitt, 1997; Cognition and Technology Group at Vanderbilt, 1990; 1992; Lamon, Caswell, Scardamalia, & Chandra, 1997; Lamon, Reeve, & Scardamalia, 1997; Lamon et al., 1996; Tumblin & McAuley, 1997; van Aalst & Marjoribanks, in preparation). Similar alliances exist between researchers and designers (e.g., Hewitt, Webb, & Rowley, 1994; Scardamalia et al., 1989; van Aalst, Teplovs, Burtis, & Scardamalia, in press). Some of this work has been funded by corporations (or government agencies whose mandate is the promotion of technology), but in most cases the researcher held the research grant and the schools
constituted the researcher's *client communities*. In some cases, Ministries of Education have matched such funding with so-called transfer grants, with both the researchers and teachers (or school principals) as principal investigators. A push to increase the practice of matching traditional research funds in this way would be a step in the direction of funding non-traditional teacher activities (i.e., research); it would require a considerable cultural shift in the educational enterprise, because it would inevitably drain funds from traditional teacher activities. Although this cultural shift may be a positive step, continuing to view teachers as part of researchers’ client communities may undermine the design experiment heuristic. Rather, to draw from Latour’s (1987) analysis once more, students, teachers, educational jurisdictions, and funding agencies should be seen as co-constructors of knowledge. (The funding agencies, just like teachers, are delivering education, just as agencies that fund science are “doing science.”) The client community of this joint knowledge construction is *society*, which gains citizens who are equipped to lead productive and rewarding lives (see Bereiter, in preparation).

I have argued at length that teachers should become reflective practitioners who inquire into the class’s knowledge building practices, but to do this they need tools. Such tools, which originated with the tracking tools of CSILE, are currently being developed. The *Vocabulary Analyzer* (Hewitt, 1998) is a web-based program that allows teachers to study the use of vocabulary as the class’s discourse progresses. Students and teachers can use it to study the evolution of discourse practices around particular words, like force, energy, or friction. It can also be used to study *who introduces* such words. The *Analytic Toolkit* (Burtis, 1998) is another web-based suite of programs to study database use. It is being used by some teachers and researchers to study who is reading whose notes, how many notes are being read, who is using keywords, the variety of problems that students work on, and similar questions. Data produced by these tools can be introduced, as a note, into a Knowledge Forum™ database, where it becomes an object of inquiry (i.e., a "knowledge object," see Bereiter & Scardamalia, 1993; Popper, 1972). Finally, van Aalst, Teplovts, Burtis, & Scardamalia (in press) are exploring the use of latent semantic analysis (LSA) for studying not only what words are being used, but how they are being used to convey meaning.\(^{50}\) The eventual goal of this research is to

\(^{50}\)Latent semantic analysis was developed to address the problems of *synonymy* (several words have the same meaning) and *polysymy* (one word has several meanings) in the context of document retrieval systems (Berry & Fierro, 1996; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990).
develop software components that teachers, researchers, and students can use to study how the discourse is evolving — i.e., does it gain the characteristics of a scientific discourse?

In chapter 1, I mentioned that the idea that synergy can develop when multiple groups share a common discourse medium was an underpinning of the Knowledge Society Network (Scardamalia & Bereiter, 1996). Another aspect of that approach is the possibility of having subject matter specialists deliberately contribute to the knowledge of a group, whose goals he or she does not share. Thus a university student could develop a “view” of a database constructed by children on the basis of her knowledge, for instance, by grouping some notes in a new way, explaining in a super-ordinate note why she grouped them in that particular way, and perhaps explaining what she learned from doing this. In this way, an outsider could model knowledge building at the same time as introducing a new perspective into the discourse. It would seem to be important that the outsider become enculturated into the discourse and have a personal relationship with the students. Research by Kevin O’Neill (1997) indicates that personal (although not necessarily face-to-face) contact is central to successful telementoring relationships. The needed personal relationship could be fostered by a classroom visit, email contact, and regular contributions to the database that are of smaller scope than views, such as commentary. The educational advantages of this proposal are that several people, including the teacher, can model knowledge building, and that the success of the discourse is less vulnerable to (possible) weaknesses in the teacher’s subject matter knowledge.

Finally, initiatives in teacher education should be mentioned. Knowledge advances are often propagated by people when they move from one institution to another (Latour, 1987). Thus, an effective way to promote educational change may

The LSA procedure computes a singular value decomposition (SVD) of a terms-by-documents matrix A (where ‘document’ stands for a text unit) consisting of the frequency of each term in the various documents. If the rank of A is low, the SVD leads to a dimension reduction of the problem (Berry & Fierro, 1996). Information about both the terms and the documents is retained in the SVD, so that one can make queries involving both. Terms and documents play a symmetrical role (Landauer & Dumais, 1997), unlike in factor analysis, which would involve only one of them. In LSA, the SVD is approximated by a reduced SVD, by keeping only a certain number of the largest eigenvalues and the corresponding eigenvectors. This has the effect of further reducing the dimensions of the computational problem, but also that of neglecting minor differences in word usage. LSA uses the cosine between vectors (term-term, document-term, or document-document) as a measure of similarity.
be via new teachers. For instance, pre-service teachers can collect data on their own teaching during their practica, while they benefit from the experience of their associate teachers. By bringing a high concentration of pre-service teachers to a given school, and supporting multiple schools with a common database accessible via the Internet, one can expect to enhance the discourse around such data. This approach, building on previous work on Professional Development Schools (Holmes Group, 1990), is being taken in the Telelearning NCE by Therese Laferrière and her co-workers (e.g., Laferrière, Abdous, Perreira, & Benoît, 1998).

6.5 Conclusion and Implications for Teaching and Research

The CSILE Project has produced a suite of technologies aimed at developing new educational practices — educational practices for the knowledge age, one might say. None of these technologies are fully developed or static, however. Viewing the evolution of these educational practices as a dynamic “super-discourse,” they can be seen as mediational means in this discourse, each with its particular strengths and weaknesses (Lemke, 1998). The full set might then be seen as encoding a set of self-consistent constraints that govern the overall dynamics. A roughly specified set might be:

- Knowledge building epistemology.
- CSILE or Knowledge Forum™.
- Teacher education as a mechanism for building a teaching culture aligned with knowledge building and the reflective teaching practice it requires.
- The Knowledge Society Network (KSN) and other mechanisms for (e.g., telementoring) that can extend the discourse beyond traditional classroom boundaries, with an aim of offsetting shortcomings in subject matter knowledge (content and process). The architecture of the KSN is a network of networks (Scardamalia & Bereiter, 1996).
- The LKB model. This is aimed to guide the design of curriculum units, but it can also be used as a scaffold for developing research questions in the context of reflective practice. (Developing and testing a benchmark lesson to introduce a unit on force and energy might be a focus in the teacher’s agenda for his own knowledge building activities during the unit.)
- Analytical tools (i.e., Analytic Toolkit and Vocabulary Analyzer) in the service of researching these questions (but also in the service of quality control).

The analysis tools produce database notes, graphs, and so on (inscriptions), that become foci of database discourses (and the super-discourse); they also mediate
cultural change toward reflexivity. Similarly, the LKB model guides the evolution of cultural practices (knowledge building, reflexivity) and produces inscriptions (databases, unit plans, etc.). Research into latent semantic analysis may lead to discourse software that is more effective at bringing notes with similar conceptual content together than currently available software is. It could be quite illuminating to study if such a model of educational change can be represented adequately by a connectionist network (see chapter 2). If so, this may begin to put the design experiment (Brown, 1992) on a more scientific basis.

Within this framework, I will in the few pages remaining in the dissertation, briefly outline two research problems: (a) the Matthew effect, (b) and assessment and data analysis.

6.5.1 The Matthew effect

The Matthew effect refers to the idea that the rate of gain is proportional to the initial endowment (Walberg & Tsai, 1983, quoted in Shaywitz et al. 1995). Stanovich (1986) talks about the rich getting richer:

The effect of reading volume on vocabulary growth, combined with the large skill differences in reading volume, could mean that a "rich-get-richer" or cumulative advantage phenomenon is almost inextricably embedded within the developmental course of reading progress. The very children who are reading well and who have good vocabularies will read more, learn more word meanings, and hence read even better. Children with inadequate vocabularies — who read slowly and without enjoyment — read less, and as a result have slower development of vocabulary knowledge which inhibits further growth in reading ability. (p. 381)

Previous research with CSILE has shown that students who had used CSILE longer made greater gains on depth-of-explanation scales (Scardamalia et al., 1992). However, this question requires further study because there were several (descriptive) variables in study 1 that differentiated between students who wrote mainly notes rated scientific late in the unit (students with HFC achievement) and other students (LFC). Studies 1 and 2 suggest the following questions:

- What is the role of initial subject matter in concept articulation? That is, do students such as those with LFC achievement in study 1 write fewer mixed framework notes because they have less subject matter knowledge (content and/or science process), or because they are less willing to publicize ideas they suspect to be incomplete or incorrect?
• If the latter of these turns out to be confirmed, can something be done to encourage students to contribute more mixed framework notes? If so, what strategies work best?
• What social conditions lead to differential participation in the database discourse? What is known about such conditions?
• What characteristics of human-computer interactions contribute to differential knowledge building? Are short-term improvements feasible?
• If analysis tools are available, how are they used, and to what effect?
• Can the value of mixed framework notes to conceptual change be established more firmly? (Including: do mixed framework notes contribute to deep explanations by others, even if there is no evidence that they facilitate learning for a student with LFC achievement.)
• What are the characteristics of instruction in Phase 1 that optimize the potential for providing a context for subsequent discourse? (E.g., are gender effects such as were observed mediated by the technical nature of the subject matter? Does starting the discussion in CSILE by making knowledge problematic help?) (See Bereiter, 1992; Chan, Burtis, & Bereiter, 1997.)
• What are the characteristics of discussion notes that are sustained beyond a few (e.g. 1-5) entries? For instance, do the first few notes clearly identify a problematic?

Such questions require detailed study of the face-to-face discourse, the database discourse, and the mediational characteristics of the activities used to start the unit — in short, ethnographic study of the early stages of the unit. For the last question, it would also be desirable to examine the value of mixed framework explanations when computers are not being used to record them. And despite the focus on ethnography, a certain amount of control is desirable. For example, an effectively functioning Knowledge Society Network may be able to facilitate a project in which teachers jointly teach a unit on the same topic, starting it with jointly developed benchmark lessons, using the same probes of depth-of-explanation; it might also buffer differences in teacher subject matter knowledge and experience. Control of this kind could lead to substantial insight into knowledge building, but requires substantial advances in the functioning of the Knowledge Society Network and in current ability to analyze large volumes of discourse data.

Lemke (1995) and Roth (1998) have conceptualized the dynamics of community development with several time scales. The above questions all operate within individual units, and mostly within their early parts (phases 1 and 2). Additional questions have to do with time scales that span several units. For example,

• Can students, over several units, advance from a predicament of weak initial subject matter knowledge? (For instance, if pre-tests in units 1 and 4 show that the student starts with similarly weak initial knowledge, and similar interest in the unit, that the student is able to gain more on depth-of-explanation measures in unit 4 than in unit 1?)
Can similar developments be established with respect to views about the nature of knowledge, such as measured by the Bereiter and Scardamalia knowledge building scale? (1996b; see also Appendix E). (E.g., do comments, over several units, begin to have more substantial content, do a greater proportion of questions become hypotheses?

Can such effects be correlated to similar effects for the teacher or to teacher knowledge building efforts? (e.g., in the context of reflective practice or by contributing to the class’s database.)

What improvements seem to be needed to the LKB model and analysis tools?

Do students, over several units, develop a greater inclination to consider a variety of possible approaches to research — modeling, experiments, etc. — late in the unit (i.e., in Phase 3?)

Studies of these questions relate to the parts of the LKB model that were not articulated until this chapter (e.g., inter-unit planning to address shortcomings in current knowledge building practices), and depend even more critically on the ability to analyze large volumes of discourse data than the within-unit questions did. They also require performance indicators that are robust across subject matter domains. Latent semantic analysis appears to be the best currently available strategy to develop in this direction.

6.5.2 Data analysis and assessment

A variety of performance indicators are required for addressing the questions I described in the previous section. Some paper-and-pencil tasks can be used (e.g., the depth-of explanation scale used by Scardamalia et al. [1992], and the inquiry test developed by White and Fredericksen [1998]), but the process of continual improvement via reflective practice will probably require automated ways of producing the necessary data and analyses. The situation is similar to the use of microcomputer-based laboratory tools (MBL; see Laws, 1991; Mokros & Tinker, 1987; Thornton & Sokoloff, 1990), which have automated data collection and analysis in science classrooms, allowing students to probe the conceptual content of the subject matter more deeply than was possible without them. In addition, in the 1990s, there has been a shift in educational policies in North America and Britain toward increased emphasis on standardized curricula and testing (AAAS, 1993; Canadian Ministers of Education, 1997; DES, 1988). The administration of standardized tests is expensive and causes a significant disruption of normal classroom life; in addition, such tests may have validity problems because they do not probe students' best work in a supportive environment (taking a test can be
stressful), are based on only a small sample of writing, and distort the curriculum. It is conceivable that data produced automatically by the students' database discourse may provide an alternative to paper-and-pencil standardized tests. Some ethnographic case studies could serve to triangulate assessments based on electronic data with that obtained from paper-and-pencil tasks or interviews. But even without them, there is reason for optimism about the potential of latent semantic analysis as a basis for automatic assessments. In several case studies, assessments of knowledge quality of short essays, as determined from analyses based on LSA, compared favorably with those of professional graders (Foltz, 1996; Landauer, Laham & Foltz, 1998). Analyses based on LSA also performed better than chance on a simulation of classification (Laham, 1997) and on a simulation of multiple-choice test taking (Landauer et al.). Additional studies, in diverse content areas, are required to confirm such findings. It is worth pointing out, however, that even if the reliability of LSA-based techniques for grading declarative knowledge claims is inferior to that of human graders, it may be good enough for our purposes. From the point of view of continual improvement, what matters is the extent to which the performance indicators can reveal significant problems to study. From the point of view of assessment, with a commitment to knowledge building values, declarative knowledge relatively uninteresting, and the increased reliability that might be obtained by using human graders might not justify the associated expense. From a knowledge building perspective, what is more important to assess is that the students' discourse becomes more scientific — the discourse must reveal understanding of scientific concepts, and evidence for scientific reasoning (Latour, 1987; Roth, 1998), such as increasing standards of evidence (Roschelle, 1996) and oscillations between alternate explanations in a period of conceptual change (Kuhn, 1989). LSA and tools for lexical text retrieval such as the Vocabulary Analyzer may provide effective tools for extracting the relevant features of the discourse, if not to evaluate them. There are many issues connected with the kinds of text analyses I have been discussing that need considerable conceptualization — there is a significant danger that such analyses might be used inappropriately, as critics of LSA are quick to point out. The point of automating assessment is to measure what can be measured adequately this way, thereby making more (teacher and student) time available for experiments, discussion, and study, not to take the human grader out of the educational enterprise. (For a related discussion, see Pedretti, Mayer-Smith, & Woodrow, 1998).
Without going into full details, I sum up by giving some examples of the kinds of performance indicators that are required:

- Demonstrating adequate curriculum coverage by mapping student statements onto curriculum guidelines or similar documents. Indicators based on the instructional program of Phase 1 are required for individual students, groups of students, and the whole class.
- Demonstrating convergence on scientific explanations as a unit progresses, for a multiplicity of units. (This includes the use of scientific vocabulary.)
- Demonstrating improvement to the extent to which students comment on other students’ notes, use keywords, summarize what has been learned, and engage in other such knowledge building strategies.
- Demonstrating evidence for increased competence at such inquiry skills as forming hypotheses, designing experiments or models, analyzing and discussing the data, and drawing conclusions, across units.
- Demonstrating increased ability, over several units, to retrieve information and report it, also providing sufficient context for the information.
- Identifying periods of transition within a unit, by evidence such as a loss of coherence in the discourse caused by oscillation between alternate explanations (Kuhn, 1989) and the use of metaphors (Roschelle, 1996), and the “trying out” new knowledge identified in study 1.

As explained in chapter 4, in a given unit, one student may not make a claim if another has made a similar one recently (in a knowledge building community there might be no reason for doing so), and much of the discourse is in face-to-face mode. This means that in a given unit only a few students may provide evidence suggesting conceptual change, “trying out ideas,” and so on. There are two approaches to addressing this problem, and I suggest that both should be used to provide complementary performance indicators. One approach is to use a large unit of analysis (groups of students and/or the whole class) and study the phenomenon for these, but repeating the analysis for each unit. (Focusing on group performance, as I did in study 2, circumvents the face-to-face problem to some extent if groups on CSILE match groups in face-to-face discourse.) The second approach is to pool data across multiple units. One would expect most students, if one examines a sufficiently long corpus of data, to write notes that suggest conceptual change, trying things out, and so on. Examining the data in this way, one can look for individual differences. For the more prolific writers, one could also examine the development of views about knowledge.
REFERENCES


Hodson, D. (Ed.) (in preparation), *Changing science education through action research: some experiences from the field.*


APPENDIX A

Concept Articulation for Students with HFC achievement (Study 1)

<table>
<thead>
<tr>
<th></th>
<th>T1 (n = 9)</th>
<th>T2 (n = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Milk is <em>denser</em> than water; so is honey.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Greater <em>density</em> means more particles.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Everything that has matter has weight.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>In gas the amount of molecules is less which means less weight and so <em>gravity</em> is not strong enough to keep gases from rising.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>The <em>molecules</em> in gases are few and far between. This is why gas is barely affected by <em>gravity</em>.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>The existing <em>molecules</em> get larger.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td><em>Molecules</em> in color are weak and cannot spread out far.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>In wood the <em>molecules</em> are few and far between.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9. Everything that is hard is a solid.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10. <em>Molecules</em> in gases are few and far between, compared with liquids.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11. When atoms move around an object fast, they take up more room. The hotter the object the more room the atoms take up.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12. <em>Molecules</em> want to expand very much and get away from the heat.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13. When the <em>molecules</em> take in too much air the <em>molecules</em> can't expand and burst.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14. Solids have more <em>molecules</em> than gasses and liquids.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15. Gas is affected by <em>gravity</em>, everything is. But since gas is so light <em>gravity</em> cannot keep it on the ground.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16. If gas wasn't affected by <em>gravity</em> then all the <em>oxygen</em> would float out into space.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17. Some gases are affected by <em>gravity</em> and some aren't.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18. When heat comes close to a solid like a [tree] the <em>molecules</em> in the tree will start to shake, that would cause some of the leaves on the tree to dry and fall off.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19. Goose bumps are <em>molecules</em> that need more space to move faster.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20. Ice and water are both forms of the same thing, two main forms that its <em>molecules</em> can be in.</td>
</tr>
</tbody>
</table>
HEAT AND ENERGY

21. *Friction* which is the force caused when two objects rub against each other.

22. When the atoms in an object stop moving the temperature is called *absolute zero*.

23. A substance cannot be cooled to *absolute zero*, substances will always have some heat energy.

24. If a gas is cooled 1°C, it would decrease \( \frac{1}{273} \) of its volume.

25. When the two sticks are rubbed together their *molecules* will mix due to the speed of the contact.

26. Heat can be caused by *friction*.

27. The amount of heat depends on the amount of matter.

28. Temperature depends on how fast the atoms are moving around.

29. The water *molecules* which make up an ice cube have less *kinetic energy* than those in hot water.

MAKING GLASS

30. When *glass* comes in to contact with a heat source it will *melt* or explode.

31. The *form* of the sand when it is melted is that of *glass*.

32. Sand consists of so many grains it will not be able to reach a true liquid state.

33. When the *molecules* of sand come in contact with EXTREME heat the *molecules* spread away from each other slowly forming a liquid.

34. All you have to do to turn sand into *glass* is, heat it at an extremely high temperature.

FIRE

35. In order to *melt* or burn or give off smoke, the *molecules* in the object need to be moving very fast.

36. The reason for water putting out fire is that the water smothers the fire with a very large amount of *molecules* so that the fire can’t get enough *oxygen* to survive.

37. *Molecules* are moving around and need *oxygen* to energize them.

38. "Natural" solid like *wood* will burn because it is dry and its *molecules* aren’t very compact.

39. Water can make *wood* expand.

40. The *molecules* in paper can link up quickly with the *molecules* of *oxygen* in the air. This linking produces heat.

41. Besides *wood* and other kinds of solid fuels, fire needs *oxygen* to burn.

42. With looser *molecules* like *wood* can get hot faster and that other solids like metals take longer to get hot because the *molecules* in the metal are tighter together so the metal takes a longer time to catch fire.
ENERGY TRANSPORT

43. Conduction of heat is when heat is transferred from one object to another.

44. Solids such as plastics and wood are insulators, poor conductors of heat.

45. Convection is more frequent in liquids and gasses, because it involves heat only reaching one point in matter then moving around causing the whole thing to be hot.

46. If air were a good conductor of heat, as metal is, there would be neither hot nor cold areas, and the air would not move.

47. Sun rays are a form of radiation.

48. Metals are better conductors of heat than wood because there are more molecules in metals and that results in the metal getting heated a lot faster.

49. Convection is when hot air rises.

50. The trade winds, cool winds which flow towards the equator, are huge convection currents. They are caused by warm air rising from the equator and near the equator and cold air flowing in to replace the warm air.

51. Heat travels quickly through metals.

THERMAL EXPANSION

52. Objects expand when heated, contract when cooled.

53. Metal will expand much faster than another object such as a brick.

54. Objects expand because the molecules move quicker.

55. The molecules are so hot that they want to get out of the little space they are in, so they expand.

56. Heat can make solids shrink, melt, expand and even bend.

57. Heat expands glass, but not as much as metal.

58. Solids and liquids expand only a little bit. Gases expand much, much, more.

59. Mercury is uses in thermometers because its molecules expand a lot faster than any other liquid when it is heated.
PHASE CHANGES

60. The boiling point of water is 100°C.
61. Denser liquids have higher boiling points.
62. Evaporation follows boiling.
63. In boiling, bubbles which have formed in the water begin to pop out onto the surface, and steam is produced off the top of the water.
64. In most cases heat will evaporate the liquid.
65. Denser solids (steel) have higher melting points.
66. The higher the melting point is for the metal, the more it will expand before it melts.
67. Different types of solids have different molecular structures and, therefore, melting points.
68. A heated liquid evaporates because when the molecules expand, there is no more room in the pot.
69. Plastic will melt because its molecules are more compact and dense.
70. When molecules expand too much the solid expands into a liquid, then into a gas.
71. Metal has a lot of molecules, but they are not as crammed as the ones of wax. This causes the molecules to expand, but not to melt like wax.
72. When cold comes in contact with most solids the molecules become closer together because the object is trying not to freeze.

73. The temperature needed to turn a liquid into a gas is the boiling point.
74. The boiling point of water is 100°C.
75. At high altitudes the boiling point of water will drop considerably.
76. When you boil water and the water gets to 100°C, the temperature will not go any higher.
77. The boiling point of a liquid depends on the density and the pressure of the surrounding air.
78. All substances boil if they are heated at a high temperature.
79. Water at room temperature can evaporate; at higher temperatures it evaporates faster.
80. When the sweat evaporates into air people will lose body heat and then feel cool.
81. If the water is heated some more it evaporates into a gas, but when water (vapor) cools, it condenses into water droplets.
82. When water freezes it expands.
83. Clean water freezes at 0°C. Clean ice melts at 0°C.
84. The temperature of the ice stays the same during melting.
85. Brass will melt.
86. If you add salt to ice, the melting point is lowered.
87. We can get metal from some rocks by heating them (smelting). Copper, tin, iron, zinc, and other metals can be made this way.

88. When ice starts melting the molecules want to expand because they can’t stand the warm temperature.

89. If a liquid is cooled enough, its particles lose the ability to move around freely. They become fixed in place. The liquid freezes at this point and turns into a solid.

90. When ice is exposed to heat warm enough, the crystals will separate, and the ice will be in the form of water.

91. Solids melt because the molecules can’t take getting so hot. The molecules would want to take more space than there really is.

92. If you heat an ice cube it will melt into water because the heat makes the molecules start to move around in the ice cube and the heat that is made melts the ice cube into water.

93. When a solid melts and becomes liquid, the network of bonds that held its molecules together breaks apart, and the molecules wander all over the place. That is why liquids flow into any shape.

94. Bubbles are molecules that are expanding and taking in air. When the molecules take in too much air the molecules burst.

95. When water is heated for a long time its molecules spread so far apart that it evaporates.
APPENDIX B
WRITTEN TEST USED IN STUDY 2

1. According to scientists,

"The sun is the main source of energy for people and they use it in various ways. The energy in fossil fuels such as oil and coal comes from the sun indirectly, because the fuels come from plants that grew long ago."

(A) What is your theory about how the sun makes its energy?
(B) Apparently, plants store energy (the energy we get from fossil fuels like coal and oil). What is your theory about how plants store some of the energy they get from the sun?
(C) Scientists have long been interested in the question, do living things exist "out there." Do your theories in (A/B) help you to understand that question? For example, what things would you accept as evidence that living things existed at one time on the moon? Why?

2. Grade 5 students in a CSILE database in Nijmegen (in Holland) have started the following (translated) discussion note:

MT: Energy can neither be destroyed nor created.
INTU: That cannot be true because I have read that we must conserve energy or we'll run out.

What do you think?

3. (A) Give a definition of the word "machine" that you think a scientist would accept.
(B) A lever is an example of a simple machine. Explain in one or two sentences how it works. Use a diagram if you think it is helpful.
(C) Give another example of a simple machine.

4. Vocabulary list: energy, power, force, weight, gravity, speed, machine, friction.

(A) The words energy, force, and power have exact meanings for scientists. Explain these words to a Grade 4 student.
(B) Are gases affected by gravity? Why or why not? Use as many words from the vocabulary list as you think necessary.
(C) A golf ball that is "driven" by a golf club flies upward in a curved path for a while and then begins to fall down again. Why does the golf ball not start falling immediately after it is hit by the golf club? Why does it not keep going up forever? Use as many words from the vocabulary list as necessary.

The scale shown in Table B.1 was used for scoring the test.
Table B.1: Scale used to grade written test

<table>
<thead>
<tr>
<th>Item</th>
<th>Type of response</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-A</td>
<td>Commonsense</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sun gathers heat from space</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sun makes heat and releases it if the temperature gets too high</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Sun is a burning ball of gas</td>
<td>4</td>
</tr>
<tr>
<td>1-B</td>
<td>Commonsense</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Anthropomorphic</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Leaf or root</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mention of fossil or organic material of plant</td>
<td>5</td>
</tr>
<tr>
<td>1-C</td>
<td>Evidence of currently living forms</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Insufficient evidence or illogical argument</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Correct and logical argument</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Commonsense</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Anthropomorphic</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Makes contact with earlier entries</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Elaborates an idea</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Advances the discussion or introduces a promising new idea</td>
<td>5</td>
</tr>
<tr>
<td>3-A</td>
<td>Unclear/incorrect</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Technical appliance</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Makes work easier</td>
<td>5</td>
</tr>
<tr>
<td>3-B</td>
<td>Incorrect</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Definition rather than example</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Correct</td>
<td>5</td>
</tr>
<tr>
<td>4-A</td>
<td>Incorrect or commonsense</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>One concept correct</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Two concepts differentiated</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Three concepts differentiated</td>
<td>5</td>
</tr>
<tr>
<td>4-B</td>
<td>Not affected</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No, gas is lighter than air</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Correct explanation</td>
<td>5</td>
</tr>
<tr>
<td>4-C</td>
<td>Commonsense</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>'Initial speed' is a factor or 'gravity pulls down the ball'</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Both of these conceptions</td>
<td>5</td>
</tr>
</tbody>
</table>

Examples of codings follow.

1-A: “The sun makes energy by burning gases making heat and heat is a form of energy. The sun gives of heat and there is energy in the rays of the sun.” This response was rated a 5 for the reference to burning gases.

1-B: “Plants store energy from the sun in there leaves, that is why some plant leaves are so large.” This response was rated a 3. Responses like this were unexpected, but there was a reference in the question to fossils. The response, “Plants don’t store energy; it just takes a long time for the energy to run through the [...] illegible” of the plant” was rated a 1.

1-C: “Bacteria on the moon would be a sign that living things could have lived there” was rated a 1 because if bacteria were found there would currently be life on the moon. The following response addresses life in the past rather than present life,
and was rated a 5: “If it was a little while after the end of life I would accept plant- or life-form remains; but if it was a long time, fossils.”

2: “INTU: I think energy can be created and destroyed. I mean the sun makes energy every second of every minute, but at the same time energy has to run out.” This response was rated a 5, because it makes contact with the previous entries in the note and asserts that the sun creates energy.

3-A: “A machine is something is not alive and it helps people in some way” was rated a 1 because helps people in some way was not sufficiently explicit. On the other hand, “A machine is some thing that makes it easier to do things” was rated a 5.

3-B: “I think you pull on it and it activates a switch” was rated a 1 because this response suggests the student was thinking of, for instance, a lever on a drill press, which is a switch. “A lever works by using a down force to make an object go up” was rated a 5.

4-A: “Energy is natural and comes from the sun. It helps things grow and exist. Force is related to friction and happens when something moves.” This response was rated a 1, as it neither provided a description of what force and energy are, nor differentiated between the concepts.

4-B: Table B.2 showed the types of arguments provided.

4-C: The response “When it starts to lose its force, then gravity compensates the force” was rated a 3 for the recognition that the golf ball falls under the influence of gravity; the first part of the response, however, invokes the impetus explanation (see McCloskey, 1983). “The golf ball will not fall back immediately because of the force of the golf club hitting it send it into the air, but gravity created by the earth pulls the ball back down to the ground. Therefore, the harder the hit the farther the ball goes.” This response was rated a 5.
Table B.2: Responses — Question 4-B, arguments about gravity and gases.

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Yes</th>
<th>No</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No evidence given</td>
<td>2</td>
<td>2</td>
<td>I</td>
</tr>
<tr>
<td>2. There are atmospheres around several planets</td>
<td>7</td>
<td>-</td>
<td>IV</td>
</tr>
<tr>
<td>3. Gravity is in air and so is gas</td>
<td>1</td>
<td>-</td>
<td>I</td>
</tr>
<tr>
<td>4. The stronger gravity is, the closer to its source</td>
<td>1</td>
<td>-</td>
<td>II</td>
</tr>
<tr>
<td>5. All things [students know] are affected by gravity</td>
<td>1</td>
<td>-</td>
<td>IV</td>
</tr>
<tr>
<td>6. Weightless, but they still don’t float for eternity</td>
<td>1</td>
<td>-</td>
<td>I</td>
</tr>
<tr>
<td>7. Gravity can force to keep down the gas</td>
<td>1</td>
<td>-</td>
<td>IV</td>
</tr>
<tr>
<td>8. Gravity is so strong it can take over anything</td>
<td>1</td>
<td>-</td>
<td>I</td>
</tr>
<tr>
<td>9. Gas is like air, and air is not affected</td>
<td>-</td>
<td>2</td>
<td>I</td>
</tr>
<tr>
<td>10. If you look up it will be painful</td>
<td>1</td>
<td>-</td>
<td>I</td>
</tr>
<tr>
<td>11. Helium repels all forces of gravity</td>
<td>-</td>
<td>1</td>
<td>III</td>
</tr>
<tr>
<td>12. Gravity is another form of energy</td>
<td>-</td>
<td>1</td>
<td>I</td>
</tr>
<tr>
<td>13. That would make gases heavier</td>
<td>-</td>
<td>1</td>
<td>I</td>
</tr>
<tr>
<td>14. Gases are energy and energy is all around us</td>
<td>-</td>
<td>1</td>
<td>I</td>
</tr>
<tr>
<td>15. Gases are not solids</td>
<td>-</td>
<td>1</td>
<td>II</td>
</tr>
<tr>
<td>16. Pollution would be hanging around ground</td>
<td>-</td>
<td>1</td>
<td>III</td>
</tr>
<tr>
<td>17. There are gases on earth and in space</td>
<td>-</td>
<td>1</td>
<td>III</td>
</tr>
<tr>
<td>18. Gases can float into the air</td>
<td>-</td>
<td>1</td>
<td>II</td>
</tr>
<tr>
<td>19. Gases don’t weigh anything</td>
<td>-</td>
<td>1</td>
<td>III</td>
</tr>
<tr>
<td>20. Gases are lighter than air</td>
<td>-</td>
<td>9</td>
<td>I</td>
</tr>
<tr>
<td>21. Gases are not at the same level as we are</td>
<td>-</td>
<td>3</td>
<td>II</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>16</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

Note: Type I, invalid premise, invalid conclusion; Type II, valid premise, invalid conclusion; Type III, invalid premise, valid conclusion; and Type IV, valid premise, valid conclusion.
APPENDIX C

CONCLUSIONS ABOUT FORCE AND ENERGY

C.1 Introduction

At the end of the unit on force and energy (chapter 5), students wrote conclusion notes. These conclusions were intended to become chapters of a book, but that idea was never carried out. Not all of the groups did write a conclusion. Except for correcting spelling mistakes, the notes have not been altered.

C.2 Group B1: Rockets

C.2.1 Questions
   1. What is a rocket?
   2. How does a rocket work?
   3. How does a rocket dislodge from the space ship?
   4. What is the difference between a rocket and a jet?

C.2.2 Answers
   1. A rocket is a booster you might say because it launches space ships out of the atmosphere and also boosts air planes from place to place. There are different kinds of rockets like there are some rockets that are massive and launch space ships off the ground and there are also rockets like I said before that help planes get from place to place. But there is one thing that all rockets have in common they have a lot of power. (Tim)
   2. To make a rocket work fuel goes into a sphere with a hole in the bottom. The fuel then ignites and tries to get out of the circle by putting force against the walls. The fuel can’t escape any other way so it is forced to go out the bottom, the force of the explosion then gives the rocket lift off. (Earl)
   3. The rocket on the space ship is only used to get the ship out of the atmosphere. As soon as the rocket gets out of the atmosphere the rocket twists and comes free to float in space. Modern shuttles have built in rockets that take off like an airplane. The rockets stay on the shuttle the entire trip and are refueled every mission, unlike the rocket ship that can only be used once but can go farther in space. (Nathan)
   4. Jets are sort of like planes and can fly but only in the Earth’s atmosphere. On the other hand rockets are everything from weapons to a spaceship. So we have learned that rockets can fly both in and out of the Earth’s atmosphere. (Tim)

---

1 The original question is not stated; nor is the reason why the group chose this question. This group, however, does provide sources, and the note is well organized.
C.2.3 *Quotes*

1. "A fireworks, shell, or cylinder shot into the air by the backward thrust of burning gases." Quoted from *Nelson Canadian Elementary School Dictionary*.

2. "A self propelling, aerial fireworks, consisting of a paper tube filled with saltpeter, sulfur and charcoal, or the like, and attached to a guiding [sic] stick, which at a height bursts into a shower of sparks: used in display for signaling: also called a skyrocket-mil., a highly explosive missile, motivated by jet propulsion, and usually launched from a airplane or ship." Quoted from the *Winston Dictionary*. (Earl)

C.2.4 *Conclusion*

We hope that you’ve enjoyed our work on rockets as much as we have had working on it. As we have been working we have picked up plenty of knowledge on the way to learning about the rocket. We have learned many things from asking questions such as: What is a rocket, how does a rocket work, how does a rocket disconnect from a spaceship, and what is the difference between a rocket and a jet and asking these questions we learned a lot more then we expected and we also learned that from one simple question you can get a whole lot more knowledge than you bargained for.² By reading this note we hope that you’ve been inspired to learn some more about the rocket because after a while you get really interested in amazing invention called the rocket. Well I guess that’s our note and we also hope you have learned a lot from our discussions and our diagrams. So until we get to work on this subject again, so long!

C.3 Group G2: Friction

1. I have learned many things from this project. I discovered many things about the different types of friction and how they are used. I have also learned where different types of friction are used. I hope everyone who read this project enjoyed reading it and also learned something from it. (MA)

2. I am glad that I chose this topic, because it helped me understand friction better. I enjoyed learning about friction, and I hope that all the readers learned something from this (in my opinion) very interesting topic. (C.B.)

I would like to convey, a very special thanks to: MW, MA, and especially to [our teacher], for choosing this subject for the class to work on.

C.4 Group G3: Rockets II

We have all learned something. All of us have learned similar things but they are not exactly the same so we are going to do this separately.³

1. **AM**: What I found out is that there are two tanks in a rocket. One holds oxygen and the other holds gas. Valves control the flow of oxygen and gas. Tubes bring the

² This was judged to be an important insight.

³ The note was written by AM, rather than co-authored by the entire group. This note provides no context and does no more than state results.
oxygen and gas to a place called the combustion chamber. When oxygen and gas mix they explode. When they explode there is equal force on either side of the combustion chamber but unequal force on the top and bottom, causing the rocket to go up. This is the basic way that I found a rocket works. There are different variations that I have not mentioned. But, of course, that's why we are doing this separately.

2. IM: What I have learned is there are two tanks in a rocket. One is called the oxidizer tank and the other one is the fuel tank. There are pumps for both of them, and they pump the fuel and the oxygen into the combustion tank, where they are burned. This pushes the rocket up because the force on both sides are equal, but the force on the top and the bottom are unequal, and that pushes the rocket up. Of course, there are quite a few other types of rockets, and we have all learned different things so that is why we are all doing separate things.

3. RFM: I learnt how one particular kind of rocket works. It's called the Solid Propellant Rocket. It has a pump on the top which pumps the grain down and on each side there is a solid propellant. In front of the solid propellant there is a burning surface. The grain comes down and gets heated and comes out of the nozzle in fire. The fire keeps coming out all the time. And the book I read doesn't explain why.

C.5 Group B4: How is Energy Made?

C.5.1 Questions

C.5.1.1 questions.
1. What is energy used for?
2. How did energy start?
3. How does nuclear energy work?
4. How does energy get into food?

C.5.1.2 Subdiscussion questions.
1. If there was nothing before the Big Bang then how was the universe created without using energy?
2. If there was no energy how was energy created?

C.5.2 Answers:

C.5.2.1 Answers to questions.
1. MT: "My theory is that energy is generated from almost everything. Because energy is used in almost every thing you use, like running your car to being in the food you eat also to run your everyday appliances." (EL)
2. MT: "... I think that energy has no origin (that we know of). ... If the "Big Bang" theory is true then there was energy before the beginning of the universe. That is

---

* This is a good explanation of the conceptual issues of propulsion.
* The note had a “main question,” which has been made the title of this subsection. The organization of the note is strong, although it does not provide sources.
why it is (difficult) for me to believe in this theory. It is hard to say how energy came to be because there is no recorded history of the earth (before history).” (MN)

3. MT: “I think that nuclear energy needs the sun to work. You could still have nuclear power for a while (without the sun) but not for long.” (NH)

4. MT: “My theory on the question ‘How Does Energy Get Into Food?’ is that Energy gets into food from sun and water because food needs sun and water to grow. However if we didn’t have the sun we would not be able to have water because it would freeze and if we lost water. Plants and humans would not be able to survive. I believe that everything revolves around the sun and if we did not have the sun everything would die. It is entirely possible that in 6 billion years, when the sun is supposed to blow up, we will be able to live on another planet.” (TW)

C.5.2.2 Subdiscussion.

1. MT: “... there must have been some source of energy building up otherwise the supposed ‘Big Bang’ would never have happened. So my theory is that either there was some source of energy before the Big Bang or a different way that the Universe was created.” (TW)

2. MT: “Actually, the ‘Big Bang’ could not have made everything because... there would have to be great sources of energy to make that explosion ....” (TC)

C.5.3 Conclusion

NI: “I think that we have come to the conclusion that we really aren't sure (of the answer to “How was Energy Made?”) at all! We know that the galaxy wasn't always as it is now, but we also know that there must have been something before the “Big Bang” (for something) to happen. Maybe there was nothing to begin with but the fact that there was nothing for so long meant that there had to be something. Like evolution. Gradually, after a very long period of time, things will change and other things will change as a result of those changes. I, personally, think that this is our best guess. For now anyway.” (JW)
APPENDIX D

PROTOCOL FOR SEMI-STRUCTURED INTERVIEWS USED IN STUDY 2

1. What topic did you specialize in?
2. Tell me one thing you liked and one thing you didn’t like about this unit.
3. What have you learned about __________ (answer to #1) from this unit? What did you already know/think about this topic?
4. Did you read notes from other groups? If yes, what was something neat that you read?
5. Do you think that what you did with CSILE is like what scientists do? Why/why not?
6. Is a scientific theory any different from other theories? For example, different from religion?
7. Do you like science?
8. Do you think you’d like to be — or could be — a scientist?

(Field notes, April 16, 1997.)
APPENDIX E

DEVELOPING HIGHER-LEVEL APPROACHES TO KNOWLEDGE

(Reproduced with permission from Bereiter & Scardamalia, 1996b)

On this poster we present a provisional scheme of levels of working with knowledge. The levels may be thought of as levels of objectification, which start with viewing knowledge as a mental state and extend to viewing it as consisting of abstract objects. Objectification means the prying loose of knowledge from individual mental states and collective practices, making it an object of constructive activity in its own right. Historically, objectification emerged over the course of many centuries. For individuals, we have sketched a series of seven levels or stages that represent increasing ability to deal with knowledge as such—to construct it, view it from different perspectives, criticize it, improve it. Thus, progression through these levels represents an educational objective of particular significance to a knowledge society.

Levels of Approach to Knowledge

LEVEL 0. Knowledge as equivalent to “the way things are.” Thoughts are distinguished from things, but thoughts about things are not distinguished from the way things are; hence, the possibility of false belief is not recognized.

LEVEL 1. Knowledge as individuated mental states. Children realize that one person may know something that another does not. Thus, implicitly, there is some entity—a fact—which a person may or may not know.

LEVEL 2. Knowledge as itemizable mental content. Children can relate things they know about a topic, and often delight in doing so. Thus, implicitly, knowledge consists of sortable items.

LEVEL 3. Knowledge as representable. In trying to communicate what they know to a reader, students take into account what the reader already knows and is in a position to understand. Thus knowledge is no longer just something in the head to be expressed but is something to be represented, shared, interpreted by others.

LEVEL 4. Knowledge as viewable from different perspectives. Students see that the same knowledge can appear in different contexts and can be viewed from different perspectives. This is an important step toward objectification.

LEVEL 5. Knowledge as personal artifacts. Although constructivism is widely endorsed by teachers, it is not common for young students to view themselves as constructors of knowledge. Viewing oneself as constructing knowledge is a large step beyond viewing oneself as constructing knowledge representations (Level 3).

LEVEL 6. Knowledge as improvable personal artifacts. A theory or other knowledge object is viewed in terms of what it can and cannot do, what its virtues are and where it is in need of improvement, although still viewed as a personal possession.
LEVEL 7. Knowledge as semi-autonomous artifacts. Students recognize that knowledge objects, like other constructed objects, can take on a life of their own and may be considered independently of their personal relevance. Thus, at this level, knowledge objects become things that one can relate to, use, manipulate, judge in various ways, and have feelings about—just like other things in the real world.