

**STUDENT SCIENTIFIC UNDERSTANDINGS IN A NINTH GRADE
PROJECT-BASED SCIENCE CLASSROOM:
A RIVER RUNS THROUGH IT**

by

Valerie Lynn Talsma

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Education)
in The University of Michigan
2004

Doctoral Committee:

Professor Joseph S. Krajcik, Chair
Professor Michael J. Wiley
Professor Annmarie S. Palincsar
Assistant Professor Elizabeth Davis

UMI Number: 3150103

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3150103

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Valerie Lynn Talsma 2004
All Rights Reserved

DEDICATION

**To the Teachers and Students
in the Class of 2000 in Foundations of Science
and to the future generations of science students.**

**May all your school-science experiences be
inquiry based,
richly contextualized,
and supported by sophisticated tools designed to
facilitate your investigations.**

**SPECIAL DEDICATION TO
PROFESSOR EMERITUS WILLIAM B. STAPP (1930-2001)**

Beloved by his Stappings, of which I am one.

Bill Stapp was both inspiration and catalyst of my work from the day I arrived at the University of Michigan in January 1989 and enrolled in two of his courses. In the Advanced Environmental Education Seminar that winter term, Bill and his graduate students founded the Global Rivers Environmental Education Network (GREEN) to bring diverse groups of people together to investigate and protect river water quality. That summer we students split into teams and conducted GREEN workshops in 22 different countries on four continents. GREEN later spread to 135 countries and merged with Earth Force in 1999. With GREEN, Bill published a number of books on multicultural cooperation, environmental education, and water quality monitoring. Bill was the chair of my Master's thesis on the Rouge River Education Project in 1992.

Bill Stapp's accomplishments in the field of environmental education are numerous. Bill was the first Director of the International Environmental Education Program at UNESCO and helped organize UNESCO's environmental education conference in Tbilisi Georgia where the field of environmental education was first defined. He has published many books and journal articles. Bill has been recognized

with numerous national and international awards and was nominated for the Nobel Peace Prize in 1993.

At the time I began this dissertation, Bill was Professor Emeritus at the School of Natural Resources & Environment at the University of Michigan, where he had founded and chaired the Environmental Education program from 1970 to 1993. Although he had retired for health reasons from the University, he agreed to serve on my dissertation committee as I researched student learning in classrooms that were using his *Field Manual of Water Quality Monitoring* as their textbook. His guidance was invaluable through the proposal and early analysis of student outcomes and I regret that he passed on before he could see the final product.

Therefore, I dedicate this dissertation to



William B. Stapp

An inspirational leader and a visionary, who motivated hundreds of students
and colleagues around the world

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support and encouragement of many people. Foremost in that list is my Advisor, Dr. Joseph S. Krajcik. It was Joe's encouragement that set me on the path to pursue my PhD and his research program that provided the context for this study. Joe was also the person, after each of the medical setbacks that plagued the completion of this document, who insisted we set a new schedule and complete the program. I owe the completion of this document to his enduring optimism.

My other committee members also deserve recognition. Dr. Annemarie Palincsar for her thoughtful critiques always presented in kind manner and her leadership in the Spencer fellowship program. Dr. Mike Wiley, who contributed significantly to my own understandings of steam ecosystems. Dr. Bill Stapp, whose spirit continues to inspire me. Dr. Ron Marx and Dr. Betsy Davis who willingly stepped onto the committee in the ninth hour after Bill's passing.

I would also like to acknowledge the contributions of the teachers at Community High School in Ann Arbor, Michigan, first for their bravery in implementing project-based science in their classrooms to improve student learning. And then for the even greater courage of letting a graduate student study the teaching and learning that occurs in their classrooms. The students in the Class of 2000 were also phenomenal, allowing me probe into their learning and video tape their activities in science class.

My peers in the graduate program also made valuable contributions to this study. I would like to especially recognize Nathan Bos for his collaboration and for his willingness to score student artifacts for the analysis on inter-rater reliability. Also, Michelle Wisnudel-Spitulnik whose parallel studies in Foundations of Science and willingness to share her research greatly informed the early stages of this study.

Drs. Shirley Campbell and Barbara Seels deserve recognition for their support and collaboration at the University of Pittsburgh during the hiatus in the completion of the dissertation when I had to earn some money. My sister, Pam Krueger, whose nit-picking exacting standards normally drive me to distraction but were an immense help with the grammatical editing. Finally I would like to acknowledge my parents, John and Pat Talsma, both Spartan Alumni, for their unfailing support during my graduate studies at “the other university.”

PREFACE

Science education reform as outlined in the *National Science Education Standards* (NRC, 1996) emphasizes the importance of active learning. Teaching for understanding requires performances or the production of artifacts that demonstrate the connections learners make between concepts. Yet, the range and nature of these tasks in specific science contexts are not well understood. The issue becomes more complex when students are offered choice in “what” or “how” they will represent content understandings. Therefore, analysis of student work in artifacts is important in order to determine their overall educational value, identify misconceptions or gaps in conceptual understandings, and assess students’ achievement levels.

This study examined student generated artifacts produced over the course of an 18-week project. This study is significant because: 1) It provides a rich description of inter-disciplinary scientific understandings as demonstrated in student work, 2) it provides data on what understandings may be demonstrated in different forms of artifacts; and 3) it tests the feasibility of using the *Standards* as an assessment tool for multiple representations of student understanding in flexibly designed assessments in extended projects.

TABLE OF CONTENTS

DEDICATION.....	ii
SPECIAL DEDICATION TO PROFESSOR EMERITUS WILLIAM B. STAPP (1930-2001).....	iii
ACKNOWLEDGMENTS	v
PREFACE.....	vii
LIST OF FIGURES	xi
LIST OF TABLES	xiv
LIST OF APPENDICES	xv
CHAPTER 1 OVERVIEW OF THIS STUDY.....	1
Significance of this research	2
Classroom environments.....	4
Research on Scientific Understandings	6
Research Design.....	8
The setting.....	10
Data Sources and Analysis.....	11
Chapter Summary	12
CHAPTER 2 THEORETICAL FRAMEWORK	14
Defining Scientific Understanding	14
Conceptual Understandings: Understandings of concepts and subject matter content	22
The National Science Education Standards: Defining what students should understand	30
Assessment: Identifying and Characterizing Scientific Understandings ..	35
Artifacts and Understandings: Theoretical Rational.....	38
Assessing Artifacts.....	46
Understandings in the Classroom: Applying Constructivist Theory to Practice.....	48
Project-Based Science.....	53
Chapter Summary	58

CHAPTER 3 RESEARCH DESIGN	61
Overview	61
Context	63
The School	63
The Classroom	64
The Students.....	65
The Teachers	66
The Researcher.....	67
The Classroom Culture and Curriculum	67
Research approach	74
Delineation of the curriculum and mapping course content to standards	75
Identification of content expectations in the selected measures	80
Analyzing Student Artifacts for Conceptual Understandings.....	90
Statistical Analysis of Student Understandings in Artifacts	99
Comparison of course content and artifact specifications to actual demonstrations of understanding across the artifacts	100
Analysis of Student Understandings across Time	101
Summary of Research Design.....	103
CHAPTER 4 FINDINGS: CONTENT, STANDARDS, AND STUDENT UNDERSTANDINGS	105
The Match between the Standards, Course Content and Artifact specifications.....	105
Standards assessed in the Artifacts.	107
Standards assessed in the Pre/Posttests.....	111
Summary of the Assessed Standards	111
Demonstrations of Student Understandings within Artifacts	113
Essays.....	113
Stream Water Quality Reports	120
Creek Models - Cycle 1: Learning how to model ecological phenomena	127
Creek Models - Cycle 2: The Final Artifact	136
Understandings demonstrated on the Pre and Post Tests.....	153
Demonstrations of Student Understandings across Artifacts and Time	161
Chapter Summary	167
CHAPTER 5 DISCUSSION.....	171
What do students learn? And what are the implications for instruction and assessment?	171
Ecological Understandings	173
Implications for Instruction: Modeling and Feedback.....	174
Implications for Instruction: Engagement in Inquiry.....	178
Implications for Assessment.....	181
Implications for Additional Research	185
Summary of Learning	186

Using the <i>National Science Education Standards</i> as a assessment tool ..	187
Validity	191
Reliability.....	196
Sensitivity	200
Impact of the assessment on instruction and classroom practices	203
Final Thoughts	205
APPENDICES	208
BIBLIOGRAPHY	268

LIST OF FIGURES

Figure 2.1: Understanding Dimensions: Understanding is partially a function of knowledge accretion and knowledge integration. Theoretically, it is possible to accumulate a large number of discrete bits of knowledge without creating relationships between those pieces of knowledge or only a few weak relationships (cells a and c). It is also possible to have a limited amount of knowledge with each piece of knowledge well <i>integrated</i> or linked to other knowledge, perhaps in multiple ways (cell b).....	17
Figure 2.2: A hypothetical conversation in a ninth grade science classroom synthesized from comments across several students and class meetings.....	18
Figure 3.1: Example of mapping standards onto a response to Pre/Post Test Question 15.....	89
Figure 3.2: The Planning Stage for Chase’s Model. The worldview on the left and the notepad, on the right, where the modeler can define the purpose of the model and plan the objects and factors that will be needed in the model.....	95
Figure 3.3: The “Build” window for Chase’s Model. The factor map is on the left. The notepad on the right lists a summary of the objects, factors and relationships that have been built. Selecting one of these features (e.g. fire) causes Chase’s description or explanation to be called up in the lower box of the notepad.....	95
Figure 3.4: The relationship editor for the relationship Chase created between stream:total solids and stream:water quality. In this relationship, Chase has built a quantitative relationship using the table view option and then clicking on the graph. His explanation is both scientifically accurate and elaborate in that he includes more than one causal reason for the decrease in water quality (D.3.3 Interactions of Earth systems Level 3).	96
Figure 3.5: Chases' relationships, which demonstrate his understanding of the sun as a source of energy (Standard D.1.1).	97

Figure 4.1: Conceptual Understandings Demonstrated in Student Essays. Level 0 (LO = red) results are included only for expected standards (See Table 4.2).....	115
Figure 4.2 Sample of student work - Introduction from scientific report written by Ezra, Leah, Gene and Kiley.....	121
Figure 4.3: Expected conceptual understandings demonstrated in water quality reports ($n_a=23$, $n_s=84$). C.3.5 –classification, C.4.2 – food chains, C.5.5 – niches, D.3.3 – interactions of earth systems, F.5.2 - human caused hazards.	122
Figure 4.4: Additional conceptual understandings demonstrated in water quality reports ($n_a=23$, $n_s=84$).....	125
Figure 4.5: Factor and Relationship Map from Casie, Lilith, and Taylor (Block 3, model 1) showing factors and constructed relationships. In this model, the students have created two objects - Traver creek and clams, six factors, and five relationships. The longest factor path, from discharge to benthics, is 3.	129
Figure 4.6: Conceptual understandings demonstrated in Model 1. ($n_a=48$, $n_s=84$). Chart a shows those standards expected in Model 1. Chart b shows additional content that was observed in the models.	131
Figure 4.7: Keshia and Glynis’s model of Traver Creek.	134
Figure 4.8: Conceptual understandings demonstrated in Model 2($n_a=46$; $n_s=85$). Chart a shows Std B: Physical Science and Std C: Life Sciences Chart b shows Std D: Earth and Space Science and Std F: Environmental Sciences.....	139
Figure 4.9: Heather and Elnora’s model 2 factor and relationship map showing impacts on dissolved oxygen and the impacts of oxygen on the taxa groups.	144
Figure 4.10: Annie and Mackenzie’s model 2 factor and relationship map showing how development of new apartment buildings affects the stream.....	147
Figure 4.11: Kiley and Leah’s model on how buildings will affect the creek. Cycle 2, Block 3.....	150
Figure 4.12: Model created by Cliff, Sean and Jack, Cycle 2, Block 1 showing the effects of a golf course on the creek.....	151
Figure 4.13: Comparison of Students’ levels of conceptual understandings on the Pre- and post-test across each of the major content areas	

($n_s=96$). Wilcoxon non-parametric sign-rank tests were used to compare students' pre and posttest achievement (* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$).....154

Figure 4.14: Comparison of Students' levels of understandings on the Pre/post-test across the Physical Sciences - Standard B ($n_s=96$). (** $p \leq .01$, *** $p \leq .001$).....155

Figure 4.15: Comparison of Students' levels of understandings on the Pre- and post-test across the Life Sciences Standard C ($n_s=96$). (* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$).....156

Figure 4.16: Comparison of Students' levels of understandings on the Pre-/post-test across the Earth and Space Sciences (Standard D) and the Environmental Sciences (Standard F) ($n_s=96$). (** $p \leq .01$, *** $p \leq .001$).....157

Figure 4.17: Standard D.3.3: Interaction of Earth Systems across measures. (ns $p > .05$, * $p \leq .05$ ** $p \leq .01$, *** $p \leq .001$)163

Figure 4.18 : Standard C.5.5: The Niche Concept and Environmental Tolerances across measures. (ns $p > .05$, * $p \leq .05$ ** $p \leq .01$, *** $p \leq .001$).....165

LIST OF TABLES

Table 3.1: FOS I (ninth grade) student subjects by teacher and class.	65
Table 3.2: Excerpt from Appendix B: Matrix of Understanding Standards.....	79
Table 3.3: Data sources used to determine student scientific understandings	81
Table 3.4: Selected Student responses to Question 15 on the posttest.	99
Table 4.1: The Traver Creek Project curriculum addresses National Science Education Standards (NRC, 1996) across the four science content areas of conceptual understanding.	106
Table 4.2: <i>National Science Education Standards</i> (NRC, 1996) assessed in the artifacts and pre/post test throughout the creek curriculum. Codes inside parenthesis e.g. (C.3.5) represent standards that are afforded by the measure but not explicitly required.	108
Table 4.3: General features of student models after model cycle 1 ($n_a=48$; $n_s=84$).	128
Table 4.4: General features of student models in model cycle 2 ($n_a=46$; $n_s=85$).	136
Table 4.5: Content areas of factors in student's computer models cycle 2 ($n_a=46$; $n_s=85$).....	137
Table 4.6: Summary of expected and observed conceptual understandings along the National Science Education Standards. Values represent number of standards in that content area.	169
Table 5.1 Examples of a General Scoring Rubrics and the Coding Scheme used in this study.....	183

LIST OF APPENDICES

APPENDIX A PARENTAL/STUDENT CONSENT LETTER	209
APPENDIX B MATRIX OF UNDERSTANDING STANDARDS.	212
APPENDIX C ESSAY ASSIGNMENT	218
APPENDIX D REPORT ASSIGNMENT	220
APPENDIX E MODEL 1 ASSIGNMENT.....	222
APPENDIX F MODEL 2 ASSIGNMENT	228
APPENDIX G PRE/POST TEST FORM A.....	230
APPENDIX H PRE/POST TEST FORM B.....	250
APPENDIX I DATA SUMMARY.....	262

CHAPTER 1

OVERVIEW OF THIS STUDY

The *National Science Education Standards* (National Research Council “NRC”, 1996) present an exciting vision for science education that emphasizes scientific literacy. They outline what students need to know, understand, and be able to do to be scientifically literate at different grade levels. In this vision, students are active learners who use inquiry to explore authentic problems within a community of scientific practice and who create products (e.g. examinations, journal notes, written reports, diagrams, data sets, physical and mathematical models, and collections of natural objects) to represent their emergent understandings (NRC, 1996).

The *Standards*¹ claim to be criteria by which to judge quality: the quality of what students know and are able to do and criteria to judge progress toward a national vision of learning and teaching science (NRC, 1996, Ch 5). Assessment in terms of the *Standards* is viewed as a primary feedback mechanism for implementing this vision of science education (NRC, 1996, Ch 5). Yet, because of the different roles students are expected to take and the different types of work they are expected to produce, educators may find it challenging both to assess student achievement and to communicate student achievement

¹ “Standards” will often be used throughout this document in reference to the *National Science Education Standards* (NRC, 1996).

effectively to educational stakeholders (i.e. colleagues, parents, administrators, researchers, policy makers, etc).

In this study, I employ the content standards in the *National Science Education Standards* (Chapter 6, NRC 1996) as an assessment tool to characterize students' scientific understandings demonstrated in selected artifacts. I also propose the *Standards* as a tool to communicate student achievement, providing a language to describe student achievement across different types of student work, time, and contexts.

I begin in chapter two by defining the construct "understanding" and give examples of different understanding frameworks that are useful in science education. I also describe constructivist pedagogy and present one program, Project-based Science, which is based upon principles embodied in the *Standards* thought to support the development of understanding (Blumenfeld, et al., 1997; Blumenfeld, et al., 1991, Krajcik, et al., 1999). I then describe a semester long inquiry of a stream ecosystem and associated student artifacts enacted in the ninth grade. Because the *Standards* claim that inferences about students' understandings of the natural world and their ability to inquire "can be based on the analysis of performance in the science classroom and work products" (Ch 5), these students' artifacts were used to identify and characterize students' scientific understandings.

Significance of this research

Constructivist based instruction in which students carry out extended inquiry and

produce concrete representations of their understandings (artifacts²) is thought to have the potential to enhance deeper understanding because students need to acquire and apply information, concepts, and principles to a question or problem that is grounded in their every-day-world. However, existing research shows that the changes in science education called for by *Standards* are difficult to put into practice, create dilemmas and challenges for teachers, require changes in teachers' values and beliefs, and often demand new student roles and different kinds of student work (Anderson & Helms, 2001; Blumenfeld, et al., 1997).

Although the types of inquiry based learning environments called for by the *Standards* are thought to promote robust understandings, the difficulty in creating these environments and in characterizing the types of understandings these environments are thought to develop (Blumenfeld, et al., 1997) means we do not have sufficient evidence to support a claim of robust understanding. Exactly what students learn, in terms of depth and breadth of scientific understandings and proficiency in the *Standards*, has yet to be established. This lack of evidence has been identified as a major gap in the research base of the science reform movement (Anderson & Helms, 2001).

A reason we lack evidence about the kinds of understandings that learners can develop in these environments has to do with the complexity of creating the classroom environments when students can carry out these types of investigations. Furthermore,

² In this paper, I use the term "artifact" as a synonym with "product" to denote shareable and critiquable externalizations of learners' cognitive work. However, the everyday usage of the word "product" connotes a final, often unchanging manifestation. I preferentially use the term artifact because it recognizes even "final products" as intermediate expressions of emergent states of understanding subject to revision and improvement as student understandings develop.

research of outcomes in these types of environments require attention to the entire complex context: curriculum content, particular instructional practice, the educational goals pursued, the assessment practices employed, the prior conceptions of the students, etc. (Anderson & Helms, 2001). Prior research has tended to focus on small pieces of scientific understanding and has been constrained by traditional notions of objectivity, reliability and validity. Limitations in our current knowledge of student outcomes are elaborated in the next section in the context of classroom environments, children's ideas in science, and issues of assessment criteria.

Classroom environments

Our current understandings of student outcomes in constructivist classrooms is limited because a constructivist approach is still relatively new to classroom practice; and empirical research has addressed only limited aspects of this type of classroom learning and motivation (Roth, 1994, Blumenfeld, et al, 1995; Anderson & Helms, 2001; Keys & Bryan 2001). Several research collaborations have attempted to build these types of learning environments in science classrooms (e.g. Blumenfeld, et al., 1991; CTGV, 1992b; Edelson, 1997; Krajcik, et al., 2000; Linn & Songer, 1991b; Roth, 1995; Schauble, et al., 1995; Songer, 1996). Initially, many of these interventions to enact constructivist theories in classrooms and the evaluation of student outcomes were conducted on a scale of a few classroom hours to a few weeks. Few studies have researched the effects of long term, 9-week quarters, semesters, year long or multi-year interventions. Furthermore, outcomes of these interventions have often been assessed in the first or second cycle of enactment rather than later when they have become more integrated into the classroom environment. In addition, research points to the difficulties

teachers face as they attempt to change their practice. Anderson and his colleagues' (1994) review of the research in this area shows that changes of the type called for in the *Standards* are very difficult to put into full practice and where attempted, generally fall far short of the mark.

In addition, the constructivist interventions have mainly occurred in elementary (e.g. Brown, Ash, Rutherford, Nakagawa, Gordon, & Campione, 1993; Brown & Campione, 1991; Brown & Campione, 1994) and middle school classrooms (e.g. Krajcik, et al., 1998; Linn & Songer, 1991a; Linn & Songer, 1991b; Schauble, et al., 1995; Songer, 1996) with fewer studies in the high schools (e.g. Roth, 1994; Roth, 1995; Crawford, 2000; Schneider, et al.; 2002). There is a need for more studies of classes implementing science reforms in the public high schools (Anderson & Helms, 2001; Keys & Bryan, 2001).

One instructional approach congruent with the call for reforms and constructivist theory is Project-Based Science (PBS) (Blumenfeld, et al., 1991; Krajcik, Blumenfeld, Marx, Bass, Fredricks, & Soloway, 1998; Marx, Blumenfeld, Krajcik, & Soloway, 1997) in which students carry out investigations and create representations of what they learn. In the PBS approach, students design investigations and produce artifacts around sub-questions related to an overall driving question. These investigations may involve text related research, scientific inquiry, computer modeling, community surveys, etc. The representations, or artifacts, that students create as products of their investigations include research reports, physical representations of ideas, computer models, and hypermedia-documents. As students carry out investigations and build artifacts, it is thought they engage in the kinds of thinking and learning (e.g., explaining phenomena, interpreting

data, and solving problems; Perkins, 1986) advocated by science education reformers. However, the nature of their understandings needs additional research.

Research on Scientific Understandings

Educational research has often limited the exploration of children's ideas in science to single or discrete concepts rather than exploring students' understandings of a complex domain area (concept vs. conceptual levels of understanding). Furthermore, much of this research in science education has been conducted in the physical sciences (e.g. Osborne & Freyberg, 1985, Summary in Table 1.1; Linn & Songer, 1991b; Roth, 1995; Songer, 1996) and much less in the biological sciences. In ecology, research has been done on students' knowledge of nutrient cycles (Hogan & Fisherkeller, 1996; Leach, Driver, Scott, & Wood-Robinson, 1996a); food webs (Leach, Driver, Scott, & Wood-Robinson, 1996b); acid rain (Brody, Chipman, & Marion, 1988/89) and pollution (Brody, 1990/91). Again, these are studies about discrete concepts rather than a conceptual understanding of an ecosystem and all the relationships thus entailed.

Many studies use an "objective" instrument such as normative or standardized assessments (e.g. Schneider's, et al. (2002) use of the NAEP instrument) or other *a priori* instrumentation. Often these instruments are used as a pre/post test type of measurement and are constrained by psychometric standards of validity and reliability. In constructivist learning environments, this method of using pre-constructed pre- and post-tests may not adequately capture the range of developing understandings that occur when students pursue different lines of inquiry and/or create a variety of artifacts, especially if students demonstrate deeper understandings in certain parts of the curriculum or draw unanticipated links to other content. Consequently, additional or alternate forms of

assessment may capture a wider range of students' developing understandings and offer alternatives to some of the psychometric criterion of objectivity, reliability and validity.

The modes of learning called for in the *Standards* imply new roles for the students in terms of designing investigations, processing information and engaging in such mental processes as interpreting, explaining, and hypothesizing. The student roles described in the *Standards* imply: (1) that students will direct much of their own learning, including designing and directing various tasks, (2) that these tasks will vary among students, and (3) that these tasks will emphasize reasoning, reading and writing for meaning, solving problems, building from existing cognitive structures, and explaining complex problems. The range and nature of these tasks in various specific science contexts are not well understood; "in fact, there is a dearth of studies on desired student work" (Anderson & Helms, 2001).

Instead of or in addition to *a priori* instrumentation, alternative methods of assessing student understandings, methods that are consistent with constructivist theories of learning and the types of understandings promoted by the *Standards*, should be employed. In science education, such assessment tasks would require students to apply scientific knowledge and reasoning in real world contexts as well as in situations that approximate how scientists do their work (NRC, 1996 Chap 5). A few studies (Spitulnik, 1995; Spitulnik, et. al., 1996; Stratford, 1996) have examined student understandings exhibited in discrete artifacts (dynamic models, hypermedia documents, etc.) but not across a series of artifacts designed as assessments to represent understandings across a semester or yearlong curriculum.

In addition, it is difficult to compare outcomes reported in one study with those in another as much of the prior research has not described student outcomes in terms of a common frame of reference. By using the *Standards* as a frame of reference, information generated from alternative modes of assessment applied locally can have common meaning and value in the larger community, despite the use of different assessment procedures and instruments in different locales (NRC, 1996 Chap 5). This contrasts with the traditional view of educational measurement that allows for comparisons only when they are based on parallel forms of the same instrument.

In summary, this study addresses some of the gaps in the research literature identified above and by recent calls for research (e.g. Anderson & Helms, 2001; Keys & Bryan, 2001). The study is situated in a public high school where the teachers were initiators and co-designers of a science curriculum that had already implemented many of the classroom innovations that are also recommendations of the *Standards*. This study employs both *a priori* instrumentation and alternative modes of assessment grounded in student work to examine the breadth and depth of student understandings of a complex system, the ecosystem of a creek. In the next section, I provide an overview of the research design.

Research Design

This study was an investigation of the breadth and depth of scientific understandings acquired by students engaged in extended inquiry around a creek in well-established project-based science classrooms at the secondary level) (Huebel-Drake, Finkel, Stern, & Mouradian, 1995). The 18-week curriculum had been piloted five years previous to this study and had undergone several cycles of refinements by the teachers in

collaboration with university-based researchers. The guiding question for this study was:

What scientific understandings do students develop in a learning environment in which they pursue long-term investigation(s) of a substantial question about stream ecology and produce artifacts that represent their understandings?

The creek project was chosen for this study because (1) the content was interdisciplinary, combining content from earth science, biology and chemistry; (2) the project had the potential of meeting a number of science standards and state objectives; (3) studying the creek was a meaningful, relevant activity, as the creek was within walking distance of the school and flowed into a nearby river which was the primary source of the community's drinking water; and (4) the project provided an opportunity to build classroom-community connections.

The following sub-questions further guided this study:

- What opportunities were provided by the different artifacts for students to express their scientific understandings?
- What conceptual understandings (breadth and depth as mapped on to the *Standards*) did students demonstrate in the artifacts?
- What patterns of understandings (if any) were demonstrated by the students?

The above questions helped to frame a five-step analysis process. These five steps included: (1) The delineation of the creek curriculum and mapping the content onto the *National Science Education Standards*, (2) The identification of *Standards* assessed in the artifacts and pre/post instrument, (3) Analysis of student understandings in each of the measures. (4) Comparison of course content and artifacts' potentials for scientific

understandings to actual demonstrations of understanding across the artifacts. And (5) the examination of student understandings across time.

To show the process by which these questions were address, I first describe the nature of the learning environment. I then represent student understandings while simultaneously comparing them to standards of proficiency described by the *Standards*.

The setting

This study was conducted in a small alternative public high school in an urban university town in the Midwest. As part of a large research and development effort, the teachers in the science classrooms have been working with educational researchers from a local university to develop and implement a three year, integrated, project-based science curriculum for all students called Foundations of Science (FOS) (Huebel-Drake, Finkel, Stern, & Mouradian, 1995, and NSF TPE 9153759). Foundations of Science had been phased in as the mandatory science curriculum, replacing separate earth science, biology, and chemistry courses at the ninth, tenth, and eleventh grades. Throughout the year, students studied scientific subject matter by investigating broad questions and creating artifacts. This study focused on the ninth grade, fall semester project on Traver Creek.

At the core of the Traver Creek project was the driving question: “Is Traver Creek ecologically healthy?” Students conducted physical, biological and chemical assessments of the creek. They collaborated to make sense of their data and to hypothesize different relationships between stream factors. They used their findings from the water quality assessments and data analysis to build computer models of this ecosystem and to craft written reports that were shared with the local watershed council.

Data Sources and Analysis

Data were collected in four ninth grade classrooms ($n_{\text{students}} = 99$) in one school over the course of one semester (18 weeks). Multiple sources of qualitative and quantitative data were collected, including:

- Student constructed artifacts: essays, scientific reports, and computer models
- Student responses to essay, short answer and multiple choice questions on pre- and post-instruments,
- Classroom observations including field notes and audio/video recordings of the classroom and of student groups involved in artifact construction.
- Teacher generated classroom handouts.

These sources of data were subjected to the five-step analysis process. The classroom observations and teacher handouts were used to identify the content of the project and to map it onto the *National Science Education Standards* (NRC 1996). This mapping was used to create a matrix tool (see Appendix B) in which demonstrated understandings could be placed as they were observed and scored. Secondly, the classroom observations and papers were used to identify the content in each of the artifacts as they defined the opportunities for students to express their understandings. Each student artifact was then examined for instances of understanding. When understandings were identified, they were characterized on a four-point scale and entered into the matrix tool.

Analysis of the matrix revealed which standards were demonstrated with understanding by the students and at what frequency through each of the different artifacts and measures. This tool supported an analysis of the range of understandings

demonstrated in each of the different artifacts by looking at the whole set of student responses.

The matrix was also examined to determine if the array of measures was sufficient to assess student understandings of the content addressed by the curriculum and to determine if there were patterns of demonstrated understandings (e.g. if some constructs were over represented or if others were missing). By examining both the range of demonstrated understandings and individual understandings, I was able to describe the scientific understandings of students involved in this constructivist, project-based science curriculum using the *National Science Education Standards* (NRC, 1996) as a reference.

Chapter Summary

This study explored the nature of student scientific understandings that were constructed through their engagement in sustained inquiry. The exploration occurred in the context of a high school that was implementing a project-based science curriculum based upon constructivist pedagogy that meets the recommendations of the *Standards*. Because the sustained inquiry involved investigations of a local creek, river ecology provided a framework for students' conceptual understandings and helped to identify the specific content standards addressed by the curriculum. The *Standards* provided a frame of reference by which student understandings revealed in this study can be compared to outcomes in other contexts or at later points in time. As such, this study has theoretical, methodological, and practical implications.

Educational researchers need to specify successful learning environments and their outcomes in sufficient detail to communicate essential features to: (1) ourselves, so

we can refine the design of the environments; (2) our colleagues, so they can elaborate, help clarify, and criticize our views; and (3) teachers and administrators, so they can put the changes into place in their classrooms, schools, and districts.

The investigation described here informs educators and researchers about the scientific understandings that students develop in one such environment, and how these understandings can be assessed. This research also demonstrates how curriculum, assessments and outcomes can be mapped onto a standards document, suggesting a practical tool for educators as they evaluate progress toward meeting the Standards. In addition, this study provides evidence that student artifacts can demonstrate the types of understandings that are valued in the new reform efforts.

CHAPTER 2

THEORETICAL FRAMEWORK

In this chapter, I expand upon the need for studies that explore the scientific understandings that learners develop when involved in sustained inquiry supported by constructivist pedagogy. I begin by defining the construct “understanding” in general and “scientific understanding” in particular and give examples of different understanding frameworks, e.g. conceptual understandings that are useful in science education. I then review how learning environments can support the development of robust understandings advocated by the *National Science Education Standards*, and use one approach, Project-Based Science in the Foundations of Science program, as an example. Finally, I describe the use of student artifacts as one way of measuring and characterizing students’ scientific understandings.

Defining Scientific Understanding

Often educators seem to have taken the concept “understanding” as an intuitively meaningful one and have not attempted explicit definitions (Nickerson, 1995). Our vocabulary of understanding implies “action,” to understand something is to grasp it, penetrate it, comprehend or apprehend it, to see the point of it (Webster's, 1985). The word itself, meaning to stand under something, implies an inside view or another perspective to the knowledge in question. One way to better understand “understanding,”

is to contrast the construct with two other constructs, “knowing” and “remembering.” To know or to remember something suggests that one has information in storage, such as a phone number or an author and book title, and can retrieve it on call (Perkins, 1991). Scientific knowledge refers to facts, concepts, principles, laws, theories, and models (NRC, 1996). A learner, who *knows* and can *remember* scientific knowledge, can recite it (e.g. Avogadro’s number is 6.02×10^{23} or pH is measured on a scale of 1-14). A learner who *understands* the scientific knowledge can use that knowledge to do something effective, transformative, or novel with a problem or complex situation (e.g. use pH measurements of a creek system to predict which macro-invertebrates may be found there) (Wiggins, 1989a).

Understanding goes beyond knowing or retrieving information along a continuum, which includes readiness for a wider range of performances (Perkins, 1991). For example, suppose that a learner can *explain* a concept (e.g. dissolved oxygen) in their own words (not just reciting a canned definition), can *exemplify* its use in fresh contexts (aquariums instead of streams), can *make analogies* to novel situations (carbonated beverages, stuffy rooms), can *generalize* the law (solubility of gasses), and can recognize other laws or principles with the same form (solutions); most educators would agree that learner has an understanding of the construct in question. Understandings can be demonstrated because understanding involves action more than the possession or accumulation of cued knowledge (Perkins, 1991; Wiggins, 1993).

Understanding science requires that an individual integrate a complex structure of many types of knowledge, including: the ideas of science, relationships between ideas, reasons for these relationships, ways to use the ideas to explain and predict other natural

phenomena, and ways to apply them to many events (NRC, 1996). “Scientific understanding” in this study derives from this definition provided in the *Standards*. It is further informed by the work of Anderson and Roth (1989) where they distinguish between the structural and functional components of scientific understandings. The structural part of scientific understanding refers to the knowledge learners develop that is scientifically accurate, conceptually coherent and integrated with their personal knowledge of the world (Anderson & Roth, 1989). The functional component arises in externalization when learners develop and demonstrate knowledge that is useful for the essential functions of describing, explaining, predicting, and controlling the world around us (Anderson & Roth, 1989, Linn & His, 2000).

According to the above definition, one characteristic of understanding is the integration and structuring of accumulated knowledge (Figure 2.1). Because this structural component includes both the acquisition of knowledge and the active construction of a schema that integrates the bits of knowledge, we can conceptualize different degrees or levels of understanding. Understanding is not a simple dichotomous state (either one gets it or one doesn't), but a multi-dimensional continuum (White & Gunstone, 1992). Understanding involves not just the acquisition of pieces of knowledge, but the integration of the pieces of knowledge into some sort of whole. Understanding includes elements of both structure and function from no understandings (Figure 2.1a) to thin understandings or a crude grasp of the whole (Figure 2.1b & c) to a more sophisticated grasp of the whole or robust understandings (Figure 2.1d.). A learner can understand a little about something (Figure 2.1b) or a lot about something (Figure 2.1d).

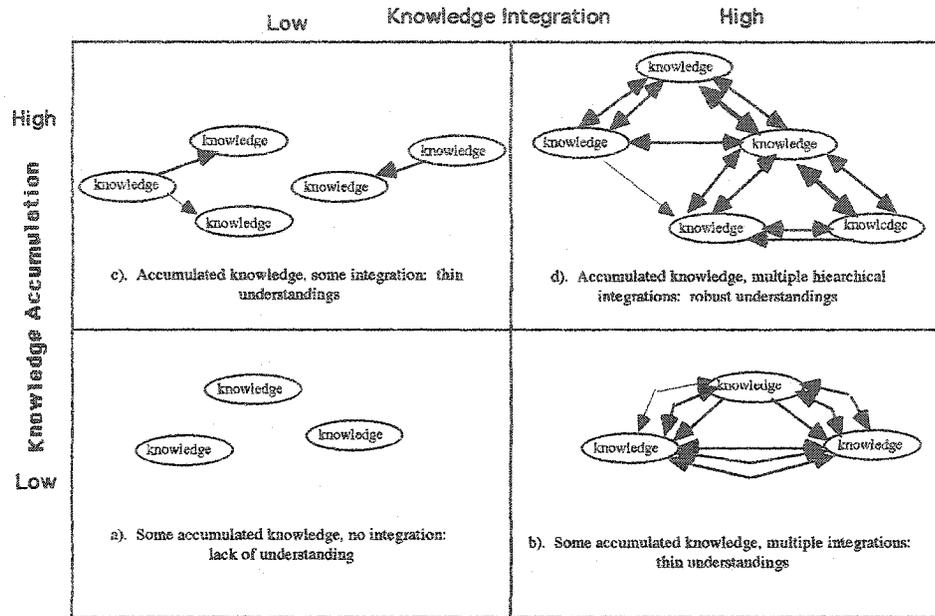


Figure 2.1: Understanding Dimensions: Understanding is partially a function of knowledge accretion and knowledge integration. Theoretically, it is possible to accumulate a large number of discrete bits of knowledge without creating relationships between those pieces of knowledge or only a few weak relationships (cells a and c). It is also possible to have a limited amount of knowledge with each piece of knowledge well *integrated* or linked to other knowledge, perhaps in multiple ways (cell b).

For example, as students begin to build a computer model of their creek study site, a conversation may develop along the line depicted in Figure 2.2.

Kreg: Juliann, did you see this article in last night's paper... the one about Allen Creek? This apartment complex wants to bury almost 300 meters of it in a huge culvert where it runs through their property and have asked the city for permission. Maybe we can use this stuff as the basis for our science project.

Juliann What do you mean, I thought we were supposed to create a computer model of Traver Creek?

Kreg: But there are several apartment complexes on Traver Creek above our study site. What do you think would happen if one of them decided to do the same thing to our creek?

Juliann: Oh, I see... we could model some of the changes that might occur in that case... like... like... I know! The sun wouldn't reach the creek inside the culvert and that would affect the photosynthesis rate... it would make it go down... because the algae need the sun to carry out photosynthesis. And that means that maybe the DO [dissolved oxygen level] would go down because algae produce oxygen when they do photosynthesis. And that might affect the kinds of benthics [macroinvertebrates] that can live in the creek. We might find fewer Taxa 1's [pollution intolerant organisms] and more Taxa 3's [pollution tolerant organisms] because Taxa 3's can survive in areas with low DO.

Kreg: Well I guess so... but I was thinking more about when they were building the culvert... because then they bring in all this big equipment... backhoes and bulldozers and stuff. And they would have to do a lot of digging and stuff to make room for the culvert. And that might cause a lot of erosion...

Juliann: Yeah... and the erosion might cause sedimentation further down stream...

Kreg: Uh-huh... and that will affect the benthics habitats, you know, like the rocks and stuff will get covered up by all this erosion stuff. So we might get more Taxa 3's because they like silty habitats.

Juliann: I think this is a good idea... let's do it!

Figure 2.2: A hypothetical conversation in a ninth grade science classroom synthesized from comments across several students and class meetings.

In this example, "Kreg" and "Juliann" take a novel situation encountered outside the classroom (local newspaper) and apply it to their study site, making hypotheses on the

possible effects a large culvert might have on their creek. In this short excerpt, we can observe the application of several concepts and relationships, the two main ones being 1). sunlight-algae-photosynthesis-oxygen production and indicator species (Taxa 1's and 3's), and 2). erosion-sedimentation-changing habitat-and community structure (more Taxa 3's).

Both of these chains indicate Kreg and Juliann's understandings as moving more toward the right side of Figure 2.1. At the same time, we can identify less robust understandings, for example, Kreg's statement that Taxa 3's "like silty habitats." In this instance, a more complete understanding would be represented by an application of the concept of "adaptation" instead of "liking." Juliann's statement that Taxa 3's "can survive" in low dissolved-oxygen environments shows a more developed understanding of adaptation than Kreg's statement about "liking," a distinction that reflects a deeper organizing principle indicated by cell d of Figure 2.1.

Demonstrations of understanding, as in the Figure 2.2 example, employ different perspectives based on different frames of reference and may be displayed at different sophistication. For instance, Juliana and Kreg touch on concepts of photosynthesis and respiration from the perspective of their stream model. A stream ecologist may view these concepts through the lens of the River Continuum Concept (Vannote, et al. 1980.) where major bioenergetic influences along the stream are local inputs (allochthonous litter and light) and transport from upstream reaches and tributaries contributing to a mix of heterotrophic and autotrophic sources of energy. A physicist's lens might be on the energetics, with little emphasis on the organisms in which photosynthesis and respiration takes place. A chemist may focus on the reactants and products of the corresponding

oxidation and reduction reactions. A cell biologist's understanding might focus on the structure of cellular membranes and the mechanisms by which photosynthesis and respiration occurs in cells. Each of these legitimate differences in perspectives represents sophisticated understandings of photosynthesis and respiration.

Legitimate differences in perspectives and sophistication of understanding will also be evident in individual student's scientific understandings of the natural world, reflecting differences in experience and exposure to science. A challenge to teachers and others responsible for assessing understanding is to decide how such variability is translated into judgments about the degree to which individual students or groups of students understand the natural world (NRC, 1996, Ch 5)

Defining scientific understanding in terms of structural and functional components may- not capture the full complexity of scientific understanding. Several educators have proposed different frames or dimensions of understanding (e.g. Schwab, 1964, Schoenfeld, 1985, Posner, et al., 1982; Brown, Collins, and Duguid, 1989; White and Gunstone, 1992; Perkins & Simmons, 1988, Perkins, et al., 1995; Novak & Gowin, 1984). These are briefly summarized below.

In his essay on the structure of disciplines, Schwab (1964) outlined three issues in his exploration of the nature, variety and extent of human knowledge. The first had to do with the organization of knowledge and the relationships between different branches of knowledge; the second with what he termed the substantive structure of knowledge, the role of conceptual structures on inquiry; and the third as the problem of syntax, the canons of evidence and proof and how well they can be applied (Schwab, 1964).

Posner, et al. (1982), in their model of conceptual change, emphasize certain features of the learner's "conceptual ecology" or understandings that influence learners' movements from naive concepts to more scientific understandings of phenomena. For instance, epistemological commitments about the nature of evidence and the importance of parsimony in a theory and metaphysical beliefs such as faith in the orderliness of nature influence the willingness of a learner to give up their naive conceptions in favor of more scientific views.

Schoenfeld (1985) discusses four factors that figure in mathematical understanding and problem solving: resources (knowledge base of facts and concepts), heuristics, control (metacognitive monitoring and control of problem solving) and belief systems (broad beliefs about the nature of mathematics and mathematical inquiry). These are similar to the four kinds of knowledge important in cognitive apprenticeships as emphasized by Brown, Collins, and Duguid (1989): domain knowledge (facts and concepts), heuristic strategies, monitoring strategies, and learning strategies. Brown, et al., (1989) claim that people who use strategies or tools within an apprenticeship actively build an increasingly richer understanding of the world in which they use the tools and of the tools themselves. The understandings, both of the world and of the tool, continually change as a result of their interaction.

White and Gunstone (1992) apply their definition of "understanding" to a range of targets: understanding of concepts, whole disciplines, single elements of knowledge, extensive communications, situations, and people. Perkins and Simmons (1988) characterize deep understanding as involving four interlocked levels of knowledge: the content frame, the problem-solving frame, the epistemic frame, and the inquiry frame. In

a revised version of this framework, Perkins, et al., (1995) explicates just three kinds of knowledge necessary for understanding: content knowledge, problem-solving knowledge and epistemic knowledge.

These frameworks of understanding help to focus the definition of understanding in this study to that of conceptual understanding. In the following section, I elaborate on conceptual understanding. Although I focus on conceptual understanding, I also recognize that conceptual understandings do not encompass the full richness of scientific understanding. One cannot have robust scientific understanding about a scientific concept unless those understandings also include how that knowledge came to be accepted (epistemological understandings) and how new scientific knowledge is constructed (procedural and strategic understandings). Likewise, one cannot carry out inquiry in ignorance. A learner needs some conceptual understanding of the phenomenon to be investigated in order to frame a question and some sort of guide to relevance and irrelevance, importance and unimportance (epistemological understandings) in order to discriminate among the findings (Schwab, 1964).

Conceptual Understandings: Understandings of concepts and subject matter content

The content standards (Chap 6, NRC, 1996) call for scientific understanding of the natural world. Such understanding requires knowing concepts, principles, laws, and theories of the physical, life, and earth sciences, as well as ideas that are common across the natural sciences. In the modern world, the volume of science-related knowledge is growing exponentially. Students today, particularly in the sciences, are inundated with information. Huge, encyclopedial textbooks and computers combined with CD-ROM

and networking technology make satellite photographs, acid rain measurements, DNA sequences and other databases available on every desk top.

Too often, in science classrooms, students are expected to master, or more specifically, memorize large amounts of material with little opportunity to apply that knowledge or to discover its relevance in their every-day-world (Brown, Collins, & Duguid, 1989). The consequences of this form of schooling are graduates with notable amounts of discrete knowledge and skills but lacking the necessary understandings to apply their school knowledge to every day life (figure 2.1 cell c). In other words, they possess inert knowledge that is often unstructured, non-functional and cannot be transferred to novel situations (Salomon & Perkins, 1989; Whitehead, 1929). The accumulation of informational knowledge is not the same as understanding.

Definition: Conceptual understanding is the set of elements of knowledge a learner possesses about a concept, plus the richness of interconnections and relationships made between concepts, plus the structure that organizes those concepts (Novak & Gowin, 1984). Implicit in this definition is the idea that conceptual understanding is a dynamic rather than a static state, for new knowledge can be added to the set, new links can be formed between things already known, and the knowledge set can be restructured based on more abstract principles.

According to White and Gunstone (1992), a person's understanding of a concept is the set of propositions (facts, opinions, or beliefs), strings (fixed or unvarying forms - like speeches), images (mental representations of sensory perceptions), episodes (memories of events), intellectual skills (capacities to carry out classes of tasks, memories of procedures) and motor skills (capacity to perform classes of physical tasks) that the

person associates with the label. In Perkins' and Simmons' scheme (1988), the content frame of understanding contains the facts, definitions, and algorithms associated with the substance of the subject matter; as well as, and even more importantly, the mapping schemes that associate concepts with referents; and the content-oriented metacognitive knowledge, such as strategies for monitoring the execution of an algorithm, or for memorization and recall.

Spoehr (1994) refers to this set or cluster of knowledge representation as a conceptual neighborhood and expands on the structural nature with the elements in a neighborhood most often organized in a roughly hierarchical way (figure 2.1 - cell d), and neighborhoods combined hierarchically to form even larger neighborhoods. Expert conceptual neighborhoods are likely to be organized using deep, abstract concepts. More sophisticated knowledge representations tend to abandon neighborhoods built out of exemplars and information that share surface characteristics, and move in favor of structures based on deeper, abstract principles (Spoehr, 1994, p 80). The earlier exchange between Kreg and Juliann (figure 2.2), is an example of understandings demonstrated within the content frame. Understandings that are more expert might be based on concepts such as energetics and 'form and function' adaptations.

Learners who have a conceptual understanding in science will not understand all the things there are to know, but they will understand quite a few concepts and will be aware that there are many more that could be looked up, learned from experience or from other people. They may know about things like solutions, dissolved oxygen, nutrients, electrons, atoms, molecules, bacteria, populations, community structures, food webs, habitats, benthic macro-invertebrates and have some mental image of what they are like.

They will also know that there are explanations of how things happen: like why algae grows in some areas of the stream and not others; or why the number of mayflies is decreasing and the number of isopods is increasing; or what happens to substances dumped in the storm sewer.

Demonstrations of Conceptual Understanding

Associated with the conceptual frame are characteristic performances that demonstrate conceptual understandings, including the familiar recall of facts and correct description of instances using specific vocabulary common to classroom environments. Yet since understanding a concept is not a simple dichotomous state, simple recall and vocabulary exercises may tell us something about a student's knowledge, but not very much about their understandings and their capacity to reason with knowledge. A more valid measure of a learner's understanding of a concept involves performances that require a learner to reveal more of the integration of structure and function of their content knowledge, e.g. generalization, contextualization, comparison and contrast, justification, application, exemplification, and explanation (Perkins, 1992). Discerning what a student knows or how the student reasons is not possible without communication, either verbal or representational, an essential component of understanding.

In the excerpted conversation in figure 2.2 (page Error! Bookmark not defined.), we can find several examples of characteristic conceptual understanding performances. Contextualization is probably the most obvious example as both Kreg and Juliann think about what might happen in a different *context* if part of their creek was placed in a culvert. They *compare* and *contrast* the state of the creek with and without a culvert. Juliann *explains* the relationship between sunlight and dissolved oxygen

mediated by photosynthesis. Kreg *exemplifies* by using the new example of construction in the creek bed as a cause for erosion. As they go about creating their model, they will have the opportunity to justify their assertions and test the model to see if it works.

Research on conceptual understandings

In the last twenty years a substantial body of research has been built up that explores and documents students' specific ideas and levels of understandings on topics within the physical, earth and life sciences. This research has often limited the exploration of knowledge to that of single or discrete concepts rather than exploring students' understandings of a complex domain area (concept vs. conceptual levels of understanding). Furthermore, much of this research in science education has been conducted in the physical sciences (e.g. Linn & Songer, 1991b; Roth, 1995; Songer, 1996; also see reviews such as Confrey, 1990; Driver & Easley, 1978; Driver, Guesne, & Tiberghien, 1985; Gilbert & Watts, 1983; Osborne & Freyberg, 1985) and much less in the biological sciences.

In the biological sciences, little has been published about students' understanding of the dependence of organisms on one another and the environment, or the flow of energy through the living environment (AAAS 1993). In ecology, research has been done on students' understandings in nutrient cycles (Hogan & Fisherkeller, 1996; Leach, et al., 1996a); food webs (Leach, et al., 1996b); acid rain (Brody, Chipman, & Marion, 1988/89) and pollution (Brody, 1990/91). Yet these, too, are studies about discrete concepts. An exception to this trend can be found in the study by Hulland and Mumby (1994) which reports on two cases of young children's sense making of information learned during a two week summer camp program on wetlands. This latter study is one

of the few that used learners' artifacts to assess understandings. The relevance of Hlland and Mumby's work to the present study is expanded below in the section on conceptual and epistemological interactions.

Understanding reflects a kind of knowing that is primarily useful in the service of some larger goal. For example, the ability to read a map is not useful in itself but does help when one needs to locate a watershed or identify land uses along a stream. Likewise understanding the concept "dissolved oxygen" is not in itself useful³. It is, however, useful in the service of some larger goal, as for example, evaluating the quality of a creek ecosystem and predicting what benthic macroinvertebrates might be found there or for monitoring municipal discharges. Therefore, it should be recognized that conceptual understandings of science interact with other forms of scientific understanding.

Conceptual and Epistemological Interactions: The nature of the interplay between conceptual and epistemological understandings can be illustrated though Posner's, et al., (1982) descriptions of features in a learner's "conceptual ecology," the understandings that influence a learner's movement from naive concepts to more scientific understandings of phenomena. In their paper, they discussed the difficulties inherent in achieving student conceptual change when the epistemological commitments of the student differ from those of the scientific community. For instance, epistemological commitments about the nature of evidence, the importance of parsimony in a theory, and metaphysical beliefs such as faith in the orderliness of nature influence the willingness of a learner to give up their naive conceptions in favor of more scientific views. Other

³ Excluding knowledge for knowledge sake or for so-called "academic" exercises, such as asking whether the concept has been accurately defined.

research has indicated that it is possible to hold multiple understandings of a concept, a conceptual profile, all of which may not be revealed in a given situation (e.g. Driver, et al, 1994, Magnusson, et al. 1997). In this case, an epistemological understanding of the domain in which one is functioning influences the conceptual understandings that may be activated during an understanding performance. For example, a student would probably describe Traver Creek differently in a creative writing class or a history class than in environmental science because the expectations, or epistemologies, of each class are different. The findings of several studies (e.g., Carey, et al., 1989; Linn, et al., 1990; Songer & Linn, 1991; Schauble, Klopfer, & Raghaven, 1991; Strike & Posner, 1992; Demastes, Good, & Peebles, 1995) suggest that a learner's epistemological framework is a factor in effecting changes in knowledge representation.

Learners' conceptual ecologies also influence the heuristics they employ in sense-making. Hlland and Mumby's (1994) study focused upon two children (ages 10-11) who demonstrated remarkably different sense-making heuristics: a recognizably scientific one verses one that involved the generation of metaphors and personalized stories. At the classroom level, such heuristics affect what learners consider evidence for or against an emerging scientific explanation, and thus affect the representation of scientific content. In addition, since science is a way of knowing or looking at the world, conceptual and epistemological understandings will also influence and in turn be influenced by understandings of scientific inquiry. Although this study focuses on students' conceptual understanding, awareness of possible epistemological may enhance the interpretation of students' scientific understandings.

Conceptual and Strategic Interactions: Students who seek to understand the world around them engage in inquiry/investigative activities, whether or not they are scientifically literate. Learners of all ages are more inquisitive about the things they know about than the things they do not. So, people who understand some science are also likely to ask questions about how things happen, what causes them to happen, what are the processes involved, how people have come to give such explanations, and what evidence supports the explanations. Scientific understanding, particularly in the conceptual and strategic dimensions, provides learners with the conceptual and technical (or procedural) tools which produce the power and precision that would not otherwise be possible (Lee & Brophy, 1996).

Early on, Schwab (1964) described the interactions of the conceptual and strategic dimensions. In his essay on the Structure of the Disciplines, he contended that our conceptual structure, especially our ignorance, leads us to formulate questions, which, in turn, lead us to investigate and collect data. Those same conceptual structures influence the way data are interpreted. He goes on to explain the reverse interplay, that of strategic understandings on conceptual understandings and the recursive nature of the feedback between the two dimensions.

There is some research evidence that supports these arguments about the inter-relationship between the conceptual and strategic dimensions of scientific understanding. For example, Stratford (1996), Penner, et al., (1997), and Finkel and Stewart (1994), found that as students engage in model building (an inquiry activity), they built conceptual understandings and that students tended to construct more in-depth conceptual understanding when building more complex models. Stratford (1996) related the model

building processes to the conceptual richness of their final models. He suggested that students who employed more strategies (analysis, relational reasoning, synthesis, testing, debugging, planning, explaining, questioning, and searching) produced models that were conceptually richer than students who used fewer strategies.

Acquiring scientific knowledge about how the world works does not necessarily lead to an understanding of how science itself works, nor does knowledge of the philosophy and sociology of science alone lead to a scientific understanding of the world (AAAS, 1993 p. 4). The challenge for educators is to weave these different aspects of science together so that they reinforce one another. Learners need to use strategic understandings of inquiry and problem solving to develop conceptual understandings. They need to use their conceptual understanding when engaging in inquiry and problem solving and their epistemological understandings to make sense of their findings. Their epistemological understandings inform their choice of strategies when involved in investigations and as they structure their conceptual understandings.

Scientific understanding includes both structural and functional aspects as well as the dimensions of conceptual, strategic and epistemological understandings. Since understanding is a multidimensional continuum, what should students understand and be able to do? And, how can we know that students understand? These issues are explored in the following sections.

The National Science Education Standards: Defining what students should understand

Three major efforts during the 1990's contributed significantly to the development of standards or targets for what students should understand (know and be

able to do) in the sciences: *Benchmarks for Science Literacy* (AAAS, 1993), NSTA's *Scope, Sequence, and Coordination* series (NSTA, 1993, 1995) and *National Science Education Standards* (NRC, 1996). States and school districts have also worked to define standards for student achievement in the subject areas.

The American Association for the Advancement of Science (AAAS), working from the foundation they helped build in *Project 2061: Science for All Americans* (Rutherford, et al., 1990), have developed over 60 "literacy goals" in science and its intersections with mathematics, technology and the social sciences. This effort, published as *Project 2061: Benchmarks for Science Literacy* (AAAS, 1993) articulates science literacy goals across levels K-2, 3-5, 6-8, and 9-12; and includes a useful discussion and presentation of the research base of the document. The parent document, *Project 2061: Science for All Americans* (Rutherford, et al., 1990) asks that students come to understand science as a dynamic, cross-connected enterprise involving mathematics and technology as well as the natural and social sciences. It also suggests that scientifically literate people can view the world through the eyes of science and develop scientific habits of mind.

Based upon exemplary practice and research, the *National Science Education Standards*, issued by the National Research Council (NRC, 1996), present a vision of a scientifically literate populace. They outline what students need to know, understand, and be able to do to be scientifically literate at different grade levels. The content standards, organized by K-4, 5-8, and 9-12 grade levels, are found in Chapter 6. These standards provide expectations for the development of student understanding and ability over the course of K-12 education. Content is defined to include inquiry (Standard A);

the traditional subject areas of physical (Standard B), life (Standard C), and earth and space sciences (Standard D); connections between science and technology (Standard E); science in personal and social perspectives (Standard F); and the history and nature of science (Standard G). The content standards are supplemented with information on developing student understanding; and they include fundamental concepts that underlie each standard. The *Standards* also describe an educational system in which all students demonstrate high levels of performance, in which teachers are empowered to make the decisions essential for effective learning, in which interlocking communities of teachers and students are focused on learning science, and in which supportive educational programs and systems nurture achievement (NRC 1996).

To compliment these efforts, the National Science Teachers Association (NSTA) has published the *Scope, Sequence, and Coordination of National Science Education Content Standard* (NSTA, Aldridge, & Strassenburg, 1995) as an addendum to *The Content Core: A guide for curriculum designers* (NSTA, et al., 1993). This supplement is designed to make the “Core” more consistent with the new *National Science Education Standards*.

Based on *Project 2061: Science for All Americans* (Rutherford, et al., 1990), is *Michigan Essential Goals and Objectives for Science Education K-12* (MEGOSE) (MDOE, 1991), used by school districts in the State of Michigan to align school curricula with the statewide achievement tests. The 212 objectives in this document are divided between elementary, middle, and high schools; and emphasize understanding over content coverage. The objectives are organized around central questions and are divided into three categories of scientific activities: constructing scientific knowledge (ask

questions, develop solutions, interpretation, and reconstruct knowledge i.e. strategic understandings), reflecting on scientific knowledge (justification, show an appreciation, seeing connections, historical and cultural perspective, relationships among STS, and describe limitations of knowledge, i.e. epistemological understandings) and using scientific knowledge (description and explanation, prediction, design, i.e. conceptual understandings) (MDOE, 1991). Each objective is defined in terms of the activities and contexts in which students should be able to use their knowledge as well as the knowledge that they will need.

Some of these documents (e.g. AAAS, 1993; NRC, 1996) describe “content standards.” Content standards specify “what” students should know and be able to do. They indicate the knowledge and skills -- the ways of thinking, working, communicating, reasoning, and investigating; and the most important and enduring ideas, concepts, issues, dilemmas, and knowledge essential to the discipline -- that should be taught and learned in school (NESIC, 1993, p. ii).

A second class of standards is “performance standards” (Shavelson, Baxter, & Pine, 1992; Wiggins, 1991). The word “perform” in common parlance means to execute a task or process and to bring it to completion. Our ability to perform with knowledge can therefore be assessed only as we produce some work of our own, using a repertoire of knowledge and skills and being responsive to the particular tasks and contexts at hand (Wiggins, 1993). Performance standards also specify “how good is good enough.” They provide not only models but also a set of implicit criteria against which to measure achievement (Wiggins, 1991). They relate to issues of assessment that gauge the degree to which content standards have been attained, and the indices of quality that specify how

adept or competent a student demonstration must be. A performance standard indicates both the nature of the evidence (such as an essay, mathematical proof, scientific experiment, constructed product, exam or combination of these) required to demonstrate that the content standard has been met and the *quality of student performance* that will be deemed acceptable (NESIC, 1993, p. iii). Progress involves successive approximations in the direction of an exemplary performance (Wiggins, 1991).

Understanding performances are the manifestation of the functional component of understanding (Anderson & Roth, 1989) and include: generalizing, comparing and contrasting, justifying, applying, exemplifying, and explaining (Perkins, 1992); offering explanations, articulating richly relational knowledge, and revising and extending explanations (Perkins, et al., 1995); and analyzing, making fine judgments, undertaking syntheses, and creating products that embody principles or concepts central to a discipline (Gardner, 1991). In science education, performances might require students to apply scientific knowledge and reasoning in real world contexts as well as in situations that approximate how scientists do their work (e.g. investigations and subsequent reports, building and testing models).

Performance standards “contextualize” content standards by identifying the manner in which the content must be demonstrated and the expected level of performance or understanding. In effect, performance standards are an interpretation of content standards (Kendall & Marzano, 1996).

The *Standards* emphasize the importance of explicit criteria for judging the quality of students' work (Chap 5). However, the *Standards* do not provide such criteria or describe intermediate levels of understanding. They describe the desired endpoint, what

the scientifically literate person should know, understand, and be able to do after 13 years of school science. But the *Standards* do not define the contexts in which understandings should be demonstrated nor how to assess different levels of understanding.

Work by researchers such as Carey (Carey, et al., 1989; Carey & Smith, 1993) Grosslight et al (1991), Stratford (1996; Stratford, Krajcik, & Soloway, 1997b) and Spitulnik (Spitulnik, 1995; Spitulnik, Stratford, Krajcik, & Soloway, 1996) have often employed three or four level scales in their assessments of student understandings or performances of understandings. These scales of performance quality indicate a proficient or mastery level of understanding, developing understandings, and novice/naïve/non –scientific understandings. The proficient or mastery level in these studies was defined by the researcher and may, or may not, be based on an outside frame of reference such as the *Standards*.

Therefore, while the *Standards* can be used to identify the content standards, other strategies must be employ to capture the *Standards* based performance standards. Once “content standards” and “performance standards” are defined, the next step is to assess students’ achievement. In Chapter 5 (NRC 1996), the *Standards* lay out expectations and criteria for assessment in science education. These criteria and associated challenges are discussed in the next section.

Assessment: Identifying and Characterizing Scientific Understandings

An essential responsibility of teachers is to report on student progress and achievement to the students themselves, to their colleagues, to parents and to policy makers (See Teaching Standard C, Chap 3 NRC, 1996). These reports provide information about:

- (1) The teacher's performance standards and criteria for evaluation.
- (2) A student's progress from marking period to marking period and from year to year.
- (3) A student's progress in mastering the science curriculum. And ,
- (4) A student's achievement measured against standards-based criteria (NRC, 1996 Chap 5).

In classrooms trying to enact the vision of science education embodied in the *Standards*, however, assessment becomes problematic. Due to the different or novel roles students are expected to take and the different or novel types of work they are expected to produce, educators may find it challenging both to assess student achievement and to communicate student achievement effectively.

In most school settings, the accepted way for a student to express “understanding” of a history lesson, scientific theory, or novel is to answer questions on a test or perhaps to write an essay (Goldberg, 1992). For example, school assessments usually ask the learner to identify the products (discourse, things, or performances) of others by recognizing the difference between two concepts, by matching scientists with their theories, or by correctly labeling flower parts or vector forces (Archibald & Newmann, 1988) often in an “end-of-chapter test.” In classrooms where the activity of answering recall questions plays a dominant role, this activity often becomes the basis for students’ operational definitions of scientific understanding (Anderson & Roth, 1989). Students, who say that they “understand” a concept or topic, often mean that they are prepared to answer recall questions about it; because, in their experience, this is the sole or primary function of scientific knowledge (Anderson & Roth, 1989).

Instead of end-of-chapter tests and other forms of *a priori* instrumentation, alternative methods of assessing student understandings, methods consistent with

constructivist theories of learning and the types of understandings promoted by the *Standards*, must be employed. In the view of assessment expressed in the *Standards* (Chap 5), assessment and learning become two sides of the same coin. The methods used to collect educational data define in measurable terms what teachers should teach and what students should learn. When students engage in an assessment exercise, they should learn from it. In conforming to the vision of science education expressed in the *Standards*, such assessment tasks would require students to apply scientific knowledge and reasoning in real world contexts as well as in situations that approximate how scientists do their work (NRC, 1996 Chap 5).

The *Standards* emphasize “embedded” and “authentic” assessments (NRC, 1996 Chap 5). Assessments embedded in the curriculum serve at least three purposes: to determine the students' initial understandings and abilities, to monitor student progress, and to collect information to evaluate student achievement. Assessment tasks used for those purposes reflect what students are expected to learn; elicit a fuller extent of students' understanding; are set in a variety of contexts; have practical, aesthetic, and heuristic value; and have meaning outside the classroom (NRC, 1996 Chap 5). Assessment tasks also provide important clues to students about what is important to learn. According to the *Standards*, assessment tasks are authentic when students are engaged in tasks that are similar in form to tasks in which they will engage in their lives outside the classroom or are similar to the activities of scientists (NRC, 1996 Chap 5). In the next section, I develop the argument that embedded and authentic science assessments result in artifacts that can be used to make inferences about student's understandings.

Artifacts and Understandings: Theoretical Rational

Authentic demonstrations of mastery in the real world often share three features: the production of discourse, things, or performances; flexible use of time; and collaboration with others (Archibald & Newmann, 1988).

One way to help students develop deeper understandings of the subject matter they are supposed to be learning may be to provide more opportunities for creating “artifacts” which demonstrate the connections they are making between concepts, and reveal what they understand about that subject matter (Gardner, 1991; Lehrer, 1993; Wisnudel-Spitulnik, Stratford, Krajcik, & Soloway, 1996). Artifacts are concrete representations of students’ emerging understandings and provide a product or result that can be assessed. In this study, I use the term “artifact” as a synonym with “product.” However, since the everyday usage of the word product connotes a final, often unchanging manifestation, I preferentially use the term artifact to indicate that the understandings expressed in these pieces of student work, even supposed “final products,” are intermediate expressions of emergent states of understanding, subject to revision and improvement as students’ understandings develop. Artifacts may include writing samples, letters, reports, journals, debates, plays and skits, scientific devices, 3-D models, drawings, computer programs and models, concept maps, videotapes, and hyper-media documents. Artifacts may be embedded in instruction or performance assessments, or serve as culminating assessment tools. They may be assessed individually or collected in student portfolios.

Some artifacts that students construct reflect mostly “recall knowledge,” e.g. worksheets and quizzes. Other artifacts, such as written reports and computer models,

require a greater time investment, are constructed over a number of days, and often (or at least ideally) undergo cycles of revision. These are the assessments where questions and tasks may be discussed, clarified, and even appropriately modified through discussion with the teacher and/or one's peers. As students engage in such thought-filled activities, their conceptual understanding may be enhanced by the development of a richer network of interconnections between mental concepts (Novak & Gowin, 1984). These are the artifacts that require the functional demonstration of scientific understanding. Therefore, in terms of understanding, this study will preferentially select as artifacts those representations that require students to integrate multiple forms of knowledge, to build explanations and other representations of science content, that provide opportunity for revision, and that provide a meaningful context or purpose so that understanding is facilitated.

Constructing artifacts provides learners with an opportunity to develop understanding as well as a context in which they may demonstrate their understanding (Papert, 1991; Perkins, 1986; 1991). The theoretical rationale comes from a synthesis of design theory and constructivist learning theory where the "construction of meaning" is viewed as a core process (Kafai & Resnick, 1996; Papert, 1991). The processes of creating artifacts requires learners to engage in many elements of design, for example: formulating questions, gathering data from multiple sources, organizing diverse and contradictory information, and presenting their findings (Lehrer, 1993; Perkins, 1986).

Blumenfeld, et al. (1991) claim that it is through the process of creating the artifact that students construct their knowledge, the doing and the learning are inextricable. As learners engage in artifact development, they enhance their conceptual

understanding as they integrate new information and build connections between concepts (Novak & Gowin, 1984). They construct and reconstruct their understandings as they synthesize information and work with ideas, forming them into a coherent structure (Papert, 1993; Perkins, 1986). Constructing new relationships with knowledge is seen as important as forming new representations of knowledge (Kafai & Resnick, 1996).

The physical representation of understanding allows for rescan and reformulation of knowledge (McGinley & Madigan, 1990). In addition, because artifacts are tangible, concrete and explicit (e.g., a model, report, videotape, hypermedia document, or computer model), they exist within a social space where they can be shared and critiqued. Artifacts are drawn from student understandings, but these understandings are reshaped, reexamined, and selected for communicating with a certain audience (Bos, Krajcik, & Soloway, 1997). This allows others to provide feedback and permits learners to reflect on and extend their emergent knowledge and revise their artifacts (Blumenfeld, et al., 1991).

The degree to which students make connections and draw relationships between concepts within their artifacts provides insight into students' understanding of concepts (Spitulnik, 1995; Spitulnik, et al., 1996; Wisnudel, 1994). The next section of this chapter will consider the types of understandings that might be revealed in student-constructed artifacts such as reports, models and concept maps.

Reports

A very familiar artifact in science classrooms is the written report⁴. Writing provides practice in using language that represents specific scientific ideas (Connolly, 1989; Keys, 1994; Keys & Bryan, 2001). Creating a written product, such as an investigation report, has the potential to facilitate the negotiation of meaning and the construction of conceptual understanding. The production of a report involves both expressive modes of verbalization in which the writer formulates personal understandings, and communicative modes in which the understandings are processed for presentation to others (Freisinger, 1980). Report writing requires the learner to retrieve, synthesize, and organize information, and to reason through a chain of events and related evidence in order to present a formulation of knowledge (Fellows, 1994; Keys, 1994). Because writing for an audience requires that language be explicit enough to be understood outside its immediate context, conceptual links are formed or clarified as writers work through the expressive and communicative state of composition (Daiute, 1992; Langer & Applebee, 1987).

Maimon and colleagues (1981) describe the act of writing as problem solving. Good writing begins with a careful definition and interpretation of the problem, defining the problem first in terms of aim, or purpose, and of audience (Maimon et al., 1981).

⁴ Writing in science is an authentic scientific activity. Scientists write to report their research, to communicate with colleagues at other institutions, to request financial support for their work, to colleagues, managers, and subordinates in their own institutional setting. Scientists write instructions and memos, and keep lab notebooks. They even write to explore their own ideas, theories and speculations (Porush, 1995). The dissemination of scientific ideas is crucial and inseparable from science itself (Porush, 1995).

Recent research on writing to learn, especially collaborative writing, has shown that the generation of language around problem-solving tasks is a key factor in enhancing student understanding (Connolly, 1989; Fellows, 1994; Keys, 1994; Roth & Roychoudhury, 1992). Further, because science writing is based on patterns of argumentation that stress the links among claims, warrants, and evidence (strategic and epistemological understandings), students engaged in report writing may enhance reasoning skills used to support report composition (Keys, 1994).

Collaborative writing tasks can structure student discussion around conceptual understandings to be included in the composition. When young children collaborate on writing tasks, they use speech to elaborate their writing and to become more familiar with words, ideas, and composing processes (Daiute, 1989b). The conversation, in turn, creates an environment in which students more readily generate text by posing ideas, evaluating, and editing each other's work (Daiute, 1989b; 1992). A collaborative writing environment should support reasoning skill development by distributing the cognitive burden, by providing a context for discussion and debate (Brown & Palincsar, 1989), by providing opportunities for negotiating meaning (Roth & Roychoudhury, 1992), and by situating writing in the more familiar format of conversation (Daiute, 1989a, 1989b, 1992; Keys, 1994). Brown and Campione (1990) found that elementary students who created collaborative products increased their use of comprehension-extending processes, such as making analogies, providing causal explanations, elaborating predictions and supplying evidence in their discussions over time.

Drawings and Graphical Representations

Text is not the only way of expressing understandings. When visuals like pictures and maps accompany a text, the text is easier to comprehend, especially if the contents of the pictures and text are related (Daiute, 1992). In addition, some people who have difficulties expressing ideas or details in writing draw very detailed pictures that express coherent themes and points of view (Daiute, 1992; Gallas, 1991). Hulland and Mumby (1994) used drawings as tools for middle school students to express their understandings about wetlands. In the two cases described in their paper, these artifacts were used as evidence of deeper, interconnected understandings of ecosystems (Stan), or weak understandings that focused on just one organism or idea at a time (Dana) (Hulland & Mumby, 1994).

In science, drawings are often used for their descriptive function, for example, in illustrating key identifying features of an organism or for mapping out key components of an environment. Drawings are also used for organizing conceptual explanations. For example, a drawing may illustrate the parts of the hydrologic or carbon cycles. As such, they serve as graphical models or representations of scientific phenomena.

Concept maps are a special type of graphical representation that help learners to organize and represent concepts in meaningful ways (Novak & Gowin, 1984). Most often, concept maps are used with the terms that make up the content of a series of lessons. After identifying concepts relevant for a particular topic, learners organize these concepts in hierarchical relationships. By mapping the concepts, learners can connect concepts in a variety of ways and can represent the personal meanings they hold for concepts - a representation of their conceptual understandings. Similarly, the modeling

tool employed in this study included a map view using iconic representations of model components and arrows representing relationships between the components.

Concept maps focus specifically on the structure and linking of concepts that the student perceives. They can be used to explore the understanding of a limited aspect of a topic and the relationships between key terms by choosing terms to direct the focus of the probe or by limiting terms to two to three key concepts and asking learners to make multiple links between these terms (White & Gunstone, 1992). Concept maps may also be used to see whether or not students relate distinct topics, and whether they can draw links between two concept maps on different topics (White & Gunstone, 1992). Wallace and Mintzes (1990) claim that they found concept maps complement other measurement techniques and are useful for revealing not only what students know, but also how they organize their knowledge.

In addition to learner generated concept maps, some researchers use a concept mapping technique to graphically represent student understandings expressed in interviews (Chi, Feltovich, & Glaser, 1981; Rye & Rubba, 1996), student writing (Fellows, 1994), to organize field observations (Barden, 1997), or to analyze a curriculum or document (Wandersee & Fisher, 1995). Mapping out understandings may also be a useful technique when analyzing other learner created artifacts such as non-linear hypermedia documents and computer models (e.g. Jackson, et al., 1996c).

Models

Models are instruments of understanding and are especially important as tools for understanding science. Perkins (1986) defines a model as any example or other representation that makes a concept more accessible by rendering it concrete, perceptual

and vivid. Examples of models include surrogate representations (photographs and drawings), captured perspectives (maps and three-dimensional representations), alternative expressions (graphs of functions), analogs or analogies, and demonstrations (Perkins, 1986). Gilbert (1991) defines science as a process of constructing predictive conceptual models. This definition unites both the processes and product of science, and identifies model building as a super-ordinate process skill. An understanding of the nature of models and model building is an integral component of scientific literacy (Gilbert, 1991; NRC, 1996; AAAS, 1993).

Hestenes (1992) proposes that students will learn the principles of a theory by using, manipulating, and building models. The construction of dynamic models encourages analyzing, synthesizing, reasoning and explaining (Spitulnik, et al., 1996). These processes help students develop both strategic and epistemological understandings as they create and revise their models; and conceptual understandings as they attempt to explicate relationships between the different parts of their models. Building a dynamic model is a concrete way to help students construct mental connections between real-world concepts by providing an environment in which they can formulate and test their mental model of a phenomenon (Spitulnik, et al., 1996).

The process of constructing models provides the context for learners to build conceptual, epistemological and strategic understandings. As students construct their models they wrestle with the content of their models (conceptual understandings), they reflect about the purpose of building models, and they must decide on the problem they wish to tackle and how they will go about finding answers or evidence to address their

problem (Spitulnik, et al., 1996; Stratford, 1996). Examining models constructed by students can provide evidence of scientific understandings.

Assessing Artifacts

Artifacts, as concrete representations of student's emerging understandings, provide a product or result that can be assessed. Some artifacts, such as concept maps, may be limited to assessing conceptual understandings. Others, such as scientific reports and the construction and testing of dynamic models, may reveal insights into broader aspects of scientific understandings. Assessment of these artifacts is important in order to determine their overall educational value, identify general misconceptions or gaps in conceptual understandings at the classroom level, and assess students' individual achievement levels (Haertel, 1991).

A one-time collection of a single artifact may not build an adequate picture of students' understandings because different artifacts reveal different aspects of understandings. A collection of artifacts (sometimes known as a portfolio) may furnish a broad, longitudinal portrait of individual performance along several standards, or they can provide summaries or inventories of many facets of individual accomplishment (Archibald & Newmann, 1988). By using a range of artifacts, it is possible to probe students' scientific understandings across multiple standards.

However, teachers may find it challenging to assess diverse artifacts and to communicate student achievement demonstrated in artifacts to parents and policy makers. Clearly relating assessment tasks and products of student work to the valued goals of science education (i.e. the *Standards*) is integral to assessment plans. Equally important is that the assessment plans have explicit criteria for judging the quality of students' work

that policy makers and parents can understand. Teachers, parents and policy makers need to be reassured that alternative assessment methods are not only as good as, but may be better than, those used when they were in school. Thus, in developing assessment strategies to compile evidence of student achievement, teachers and researchers need to demonstrate that alternative forms of data collection and methods of interpreting achievement are at least as valid and reliable as the familiar short-answer test (NRC, 1996).

A few studies (Spitulnik, 1995; Spitulnik, et. al., 1996; Stratford, 1996) have examined student understandings exhibited in discrete artifacts (dynamic models, hypermedia documents, etc.) but not across a series of artifacts that represent understandings across a semester or yearlong curriculum. In addition, it is difficult to compare outcomes reported in one study with those in another because much of the prior research has not described student outcomes in terms of a common frame of reference

In this dissertation, I employ the content standards in the *National Science Education Standards* (Chapter 6, NRC 1996) as an assessment tool to characterize students' scientific understandings demonstrated in students' artifacts. I also propose the *Standards* as a communication tool, providing a language and a common frame of reference to describe student achievement across different types of student work, time, and contexts. By using the *Standards* as a frame of reference, information generated from alternative modes of assessment applied locally can have common meaning and value in the larger community, despite the use of different assessment procedures and instruments in different locales (NRC, 1996 Chap 5). This contrasts with the traditional view of

educational measurement that allows for comparisons only when they are based on parallel forms of the same instrument.

Assessment is one measure of progress toward the vision of science education described in the *Standards*. Yet progress toward the vision also depends upon an understanding of how teachers can promote learning with understanding in the classroom. Bridging the notions of instruction and assessment is the challenge to educators to provide “performances spaces” for students to develop and demonstrate their scientific understandings (Perkins, 1991; 1992). Efforts to provide such spaces by applying constructivist theory in the classroom are highlighted in the next section.

Understandings in the Classroom: Applying Constructivist Theory to Practice

Too often, even when “understanding” is purportedly the target of instruction, there continues to be a focus on teachers teaching and students practicing decomposed and decontextualized skills (Campione, 1991); a focus that researchers need to attend to when trying to measure student understandings. In many classrooms, students do not get to practice their understandings but instead practice “remembering” (Perkins, 1992). School science tends to present science as a series of known concepts and ideas, a body of knowledge to be mastered (Aikenhead, 1982; Perkins & Simmons, 1988). For example, in an observational study of 11 junior high school science classes, only a very small proportion of tasks required higher-level creative or expressive skills; the predominate activity involved copying information from the board or textbook onto worksheets (Mitman, et al., 1987). Teachers in these classrooms stress correct answers, grades, competition, and public comparison with others. Students are often not provided opportunities to learn the critical thinking skills that permeate the cognitive repertoires of

accomplished learners (Campione, 1991) and that help develop understandings. This situation is compounded by the nature of instruction in the higher grades, where the emphasis is too often on breadth of coverage. Students are not required to explore a subject in depth, and consequently, it is not easy for them to learn to evaluate new information critically and build the multiple links between concepts that are the hallmark of robust understandings (figure 2.1-d). In the face of such instructional activities, students are likely to conclude that science is static rather than active, and that science proceeds in a linear trial-and-add-new-information approach rather than as a series of conjectures that may or may not be supported (Linn, et al., 1990).

Since understanding science requires that students be able to use their knowledge (functional component) in building explanations, models, and arguments; it makes sense that students will need experiences in doing or building, in order to learn how to do or to build. If scientific understanding is the goal, then students need opportunities to develop that understanding from a combination of direct experiences and knowledge from experts (including the textbook and the teacher) and to consider whether their own beliefs gained from prior experience are consistent with their new experiences afforded through classroom activities, with their own lines of inquiry, and with the canonical explanations of scientists (Raizen, et al., 1989 p 46).

The findings of several decades of cognitive research have pointed to the constructive nature of human learning, the complex nature of expertise related to specific subject areas, the power of intuitive conceptions, and the limitations of school knowledge for application in non-school settings (Winn, 1993).

In the constructivist framework, people are seen not as mere recorders of factual information but as creators of their own unique knowledge structures. Bolstered by this research knowledge, proponents of science education reform have begun to consider strategies for promoting greater instructional and curricular emphasis on thinking and reasoning; and have begun to develop new descriptions of competence and proficiency that emphasize such themes as thinking, reasoning, complex performance, and problem solving in addition to knowledge and skills.

In searching for organizing principles of instruction and curriculum that attend to critical relations between thinking and motivation for developing understanding, educators have repeatedly turned to Dewey's ideas of project-based education (Kilpatrick, 1918, Dewey, 1956).

Project-based education is a comprehensive approach to classroom teaching and learning that is designed to engage students in relatively long-term, problem-focused, and meaningful units of instruction that integrate concepts from a number of disciplines or fields of study. In these environments, students engage in a continual, partially self-directed process of inquiry and knowledge building, sharing their understandings with other members of the "learning community" to which they each belong. Several research groups are currently attempting to build these types of learning environments in science classrooms.

The Cognition and Technology Group at Vanderbilt (CTGV) have developed an approach where students are presented with complex, realistic problems that call for the types of mathematical and scientific understandings required in real life (CTGV, 1990, 1992a, 1992b, 1996). Students who complete the multimedia adventures develop

complex problem-solving skills, such as the ability to plan solution procedures and generate the sub-problems and sub-goals required to solve complex, multi-step problems.

Brown and Campione (Brown, et al. 1993; Brown & Campione, 1994) have developed a program known as *Fostering a Community of Learners* (FCOL). In their learning community, members (i.e. elementary students) have expertise in different content areas related to the curriculum. Each member is responsible for sharing their expertise with others and for seeking out others whose expertise can further their own understandings and knowledge.

Computer Supported Intentional Learning Environments (CSILE) is a computer environment that stores the combined understandings of a group (e.g. all the students in a single classroom) in a communal database (Scardamalia & Bereiter, 1996). CSILE users can access the entire database, contribute text and graphical notes to the database, and add comments and queries about other people's notes. A learning community evolves as a natural by-product of the CSILE system. CSILE students report that their classroom activities are directed more toward learning than toward completing tasks or assignments. There is evidence from this approach that CSILE students independently pursue collaborative knowledge-building activities (Scardamalia & Bereiter, 1996).

Schools for Thought (Lamon, et al., 1996) combines features of FCOL, CTGV, and CSILE and emphasizes the importance of sustained thinking about authentic problems that form the basis of extended in-depth inquiry in domains such as science, social studies and mathematics. Written artifacts collected from students showed a marked increase in quality over the course of the school year for every student. Students wrote more and organized it better (Lamon, et al., 1996). Students also showed increased

sophistication in their views about research as a process of knowledge building rather than knowledge telling, and they developed a good intuitive grasp of the nature and purpose of experimentation in science (Lamon, et al., 1996).

The Kids as Global Scientists (KGS) Program (Songer, 1997; Mistler-Jackson & Songer, 2000) engages students in 'real-time,' inquiry-based weather curriculum. The project focuses on the educational potential of an Internet-based middle school weather curriculum that begins with students collecting their own weather data, such as wind speed and direction, cloud type, or tests for acid rain or snow. Distinctive features of the project include: a) classroom learning that includes peer coaching and information exchange world-wide, b) learning which can capitalize on today's science developments, and c) a coordinated, yet flexible curricular shell which allows for local customization by teachers and students at each site (Songer, 1997).

Computer as Learning Partner (CLP), developed by Linn and her colleagues (Linn et al., 1990; Linn & Songer, 1991; Songer & Linn, 1990), provides a semester-long integrated curriculum for middle school students to learn about heat, light and sound. This approach begins with students' naive exploratory heuristics and concepts and guides the students through a series of predefined experiments and writing exercises, which are designed to challenge students' previous beliefs. Accompanying specially developed software, The Electronic Laboratory Notebook (E-LabBook), is designed to help students collect, display, and explain findings.

The Learning Through Collaborative Visualization Project (CoVis) project involves students collecting, analyzing and visualizing weather data and collaborating via telecommunications. The program emphasizes projects that give students the opportunity

to learn scientific concepts and procedures in practice (CoVIS, 1996). An important project design goal is to create activities centered on more open-ended project-based scientific inquiries that are, at least in part, born of the students own interests and situated in the greater scientific and political world (Gomez & Gordin, 1995). Projects are supported by the Collaboratory Notebook, a networked, multimedia knowledge-building environment that has been designed to help students, teachers and scientists share inquiry over the boundaries of time and space (O'Neill & Gomez, 1994).

Each of these programs is founded upon elements of constructivist theory thought to enhance the development of more robust scientific understanding. In this study, another approach, project-based science, provided the constructivist environment in which learners are thought to develop robust scientific understandings.

Project-Based Science

The teachers, in whose classrooms the research for this dissertation occurred, employed Project-based Science as an instructional framework to restructures their high school science curriculum. Project-based science (PBS) is a comprehensive approach that combines research from cognitive science and motivation to focus on developing scientific understanding by engaging students in investigations (Blumenfeld, et al., 1991). Within the PBS framework, students pursue solutions to non-trivial problems by asking and refining questions, debating ideas, making predictions, designing plans and/or experiments, collecting and analyzing data, drawing conclusions, communicating their ideas and findings to others, asking new questions, and creating artifacts (Blumenfeld, et al., 1991; Krajcik et al, 2000, Marx, et. al., 1997; Krajcik, et al. 1999).

The PBS pedagogy is built around five features. These features were described by Krajcik, et al. (1994) and include activities within projects that: a) engage students in investigating an authentic question or problem that drives activities and organizes concepts and principles; b) require students to develop a series of artifacts, or products, that address the question/problem; c) encourage students to engage in investigations; d) involve students, teachers, and members of society in a community of inquiry as they collaborate about the problem; and e) promote students' use of cognitive tools. These five features are elaborated below.

Driving questions: Driving questions are the framework that situates or anchors the content, investigations and artifacts over the course of a project. Krajcik, et al.'s, (1994) description of PBS states that good questions or problems are a) feasible - they can be broken down into smaller questions and students can design and perform investigations to answer the question/problem; b) worthwhile - containing rich science content and relating to what scientists really do; c) contextualized - based in the real-world; and d) meaningful, interesting, exciting and/or important to the learners. A crucial design issue for PBS has been how to ensure that the project questions are educationally rich enough that in seeking answers students must gain understanding of significant subject matter concepts and an epistemology of inquiry. The driving question addresses the constructivist point that students need to learn by addressing authentic problems (Ernest, 1995). They serve as conceptual organizers for the various activities that occur during the course of a project.

The driving questions also tend to focus classroom activities in order to explore the concepts in depth (Marx, et. al., 1997). This is consistent with the reform efforts'

emphasis on teaching less content in order to teach it better (Rutherford, et al., 1990 p. xi.). There is research that relates the depth of coverage of subject matter to student understanding of the content (Glaser, 1984; Raizen, et al., 1989). Deeper, more complex coverage of a concept or set of concepts increases the opportunity for students to be engaged in effective complex problem solving (Chi, Feltovich, & Glaser, 1981; Resnick, 1987).

Investigations: In project-based science, students pursue solutions to different questions or problems by carrying out investigations. Investigations may resemble familiar science laboratory exercises or they may include such activities as research on the Internet, surveying classmates or community members, or interviewing knowledgeable collaborators. They may be designed by the teacher or an *ad hoc* design by the students or collaborative teams. Key features of investigations in PBS are that they involve students in asking and refining questions, making predictions, designing plans, identifying resources, collecting and analyzing data, drawing conclusions, sharing ideas, and asking new questions.

Students need to have sufficient understandings to explore information pertinent to the problem to be successful in their project investigations. For example, students need to use tool skills that are necessary to undertake the project (e.g., reading maps, using a compass, or operating computer software) as well as the tools necessary to process the information. Students need to use an array of learning, metacognitive, and problem-solving strategies during projects (Blumenfeld, et al., 1991). Students further develop their understandings as they pursue solutions to different questions or problems within the project by carrying out investigations. Teachers in PBS help students develop

the understandings needed to carry out the project by designing benchmark activities, structured investigations, and other classroom experiences that will provide students with sufficient knowledge and strategic understandings to begin their investigations. (Krajcik, Blumenfeld, Marx, Bass, Fredricks, & Soloway, 1998)

Collaboration: Any learning community is limited by the combined knowledge of its members. Within traditional schools, members draw on a limited knowledge capital, the teacher. A cornerstone of the newer approaches to learning is collaboration with students in the same classroom and in classrooms located at other sites as well as community members. In PBS classrooms, students discuss and try out their ideas, challenge the ideas of others, form research groups, and specialize in one or more aspects of the project. Telecommunications allow students to interact with a wider community of other students, and outside science experts to share information, data, resources, and ideas. This feature is consistent with the social constructivist foci on the role of a learning community in facilitating understanding (O'Loughlin, 1992). Research on collaborative learning suggests that structured small-group work results in increased learning for the individual (e.g., Brown & Campione, 1990; Fellows, 1994; Keys, 1994; Palincsar, et al., 1993; Roth & Roychoudhury, 1992).

Artifacts: Artifacts have been defined in a previous section. In PBS, producing an artifact requires students to engage in many thoughtful activities: gathering data from multiple sources, analyzing topics into subtopics, organizing diverse and contradictory information, formulating questions, and presenting information (Lehrer, 1993; Perkins 1986; Stratford, 1996). Artifacts should be rich enough to promote both depth and breadth of knowledge in their creation as well as demonstrate student mastery of the

content. Blumenfeld, et al. (1991) claim that students' freedom to generate artifacts is critical, because it is through this process of generation that students construct their knowledge, the doing and the learning are inextricable. Because artifacts are concrete and explicit (e.g., a model, report, videotape, or computer program) they can be shared and critiqued. This allows other to provide feedback and permits learners to reflect on and extend their emergent knowledge and revise their artifacts. The creation and sharing of artifacts makes doing project-based science like doing real science and mirrors the performance of individuals in the work world (Krajcik, et al., 1994).

Students also construct understandings as they design their artifacts if they can do so collaboratively (Brown, et al., 1989). Peer discussion can facilitate learners' abilities to acquire concepts and solve problems, as members of a collaborative group attempt to come to consensus through a process of meaning negotiation.

Cognitive tools and technology: Technologies, especially computers, are new tools available to students in many classrooms. Constructivism emphasizes learners' use of cognitive tools that can help extend and amplify learners' mental processes (Salomon, et al. 1991). In doing projects, students need access to information and examples or representations that will help them to understand and use central ideas. Using technology in PBS classrooms provides students access to data and information, promotes laboratory investigation, and emulates tools experts use to produce artifacts. For example, probeware enables students to collect water quality data such as pH, dissolved oxygen (DO), temperature and conductivity with high reliability and at different time scales that afforded by traditional chemistry kits.

Technology allows students to easily manipulate, construct, and revise their own representations in several media, including text, graphic, video, and audio. Through this process, students can develop their conceptual understandings. Technology can enhance challenge, variety, and choice by providing multiple levels of tasks to match student knowledge and proficiency, access to numerous sources of information that allow breadth in project questions, and offers many possibilities for artifact production. In addition to text, there has been an explosive growth in the use of computer multimodal and multimedia capabilities (sound, graphs, color, pictures, and video). This variety provides for representation of single concepts in multiple, simultaneous modalities which can enhance student understanding. This technology also allows students to manipulate and construct their own representations easily and multimodality.

These five features of PBS form the background of the instructional design and enactment of the classrooms that participated in this study. The development of scientific understandings is thought to occur during the dynamic interplay of students and these features. By examining understandings that students demonstrate during a semester long project, this study will fill in some of the gaps about the kinds of scientific understandings that are facilitated by this type of learning environment.

Chapter Summary

In this chapter, I have taken the definition of scientific understanding presented in the *Standards* and expanded it with Anderson and Roth's (1989) idea that understanding has both structural and functional components. The literature on "understanding" has helped define the focus of this study on a sub-set of scientific understandings: conceptual understandings.

The *National Science Education Standards* (NRC 1996), in explicating a vision of science education, have defined which conceptual understandings students should have at the end of their K-12 education. The *Standards* also describe how assessment should be used in schools seeking to implement this vision although “performance standards” are not provided in the document. I argue that artifacts, the products of student work, allow us to make inferences about students’ scientific understandings. Artifacts alone will not promote scientific understanding, but rather, learning environments that encourage inquiry and value the process of artifact construction, including collaboration, review and revision will be more likely to encourage the construction of scientific understandings than classrooms where rote memorization is valued. However, artifacts can be useful in assessing scientific understandings, as they are concrete representations of the functionality of such understandings.

Furthermore, I argue that the *Standards* provide both a tool to assess understanding and a language and frame of reference in which to communicate student achievement in science classrooms. This assessment tool may aide teachers as they try to implement the vision of the *Standards* in the classroom by providing feedback on student progress. Different students will demonstrate understanding in different ways, and different students will achieve different degrees of depth and breadth of understanding depending on interest, ability, and context (NRC, 1996, Chap 5). *Standards* can provide a frame of reference for teachers and educational researchers to communicate student achievement to parents, policy makers and other educational stakeholders.

In the next chapter, I outline the process used in this study to identify and describe students' scientific understandings using the *National Science Education Standards* (NRC 1996)

CHAPTER 3

RESEARCH DESIGN

In this chapter, I describe the design of this study. I begin by giving an overview of the study. I then describe the context in which the research was situated before providing more detail about the data collection and analysis.

Overview

This study was an investigation of the robustness of scientific understandings acquired by students engaged in extended inquiry in project-based science classrooms.

The guiding question for this study was:

What scientific understandings do students develop in a learning environment in which they pursue long-term investigation(s) of a substantial question about stream ecology and produce artifacts that represent their understandings?

In the definition of understanding presented in chapter two, scientific understanding was defined as being a multi-dimensional continuum including elements of both structure and function both within and between the various dimensions of understanding. The functional component of scientific understanding was critical to this study. If students did not use their scientific knowledge in appropriate contexts, then the students did not demonstrate scientific understandings. An important context in which students were expected to demonstrate their understandings was in the construction of

artifacts. Therefore, I employed content analysis of major student artifacts that were thought to represent the understandings students developed in this environment. The content analysis of the artifacts was guided by the following clusters of questions:

- What opportunities were provided by the different artifacts for students to express their scientific understandings?
- What conceptual understandings (breadth and depth as mapped on to the *Standards*) did students demonstrate in the artifacts?
- What patterns of understandings were demonstrated by the students? How did their levels of demonstrated understanding change across the artifacts?

These questions helped to frame a five-step analysis process. The five steps include:

1. The delineation of the curriculum and mapping it onto the National Science Education Standards.
2. The identification of opportunities (and expectations) to demonstrate understanding in the selected artifacts and a pre/post test instrument.
3. Analysis of student understandings in each of the measures.
4. Comparison of course content and artifact specifications to actual demonstrations of understanding across the artifacts.
5. The examination of student understandings across time.

To show the process by which these questions were addressed, I first describe the nature of the learning environment to substantiate its claim to being an active learning environment. I then attempt to represent student understandings while simultaneously comparing them to standards of proficiency described by science education reform documents such as the *National Science Education Standards* (NRC, 1996) .

Context

In this section, I describe the context in which this study occurs. The context includes brief descriptions of the school, classrooms, students, teachers and researcher and a more detailed description of the curriculum.

The School

This study was conducted in a small alternative public high school in an urban university town (population ~100,000) in the within the Great Lakes Watershed. The school had an enrollment of 350 full-time students and 100 part-time students from the district's other high schools. Admission to the school at the time of this study was determined half by a lottery system and half by first-come, first-served enrollment.

The school's culture promoted independence and responsibility among the students. The campus was open for all grades with a casual climate, i.e., teachers and other adults were called by their first names. There were no bells or intercoms and relatively little formal pressure was placed on students to attend classes. The curriculum is solid academically, though not high-powered. There were no advanced placement (AP) course offerings. There was a focus on the arts (drawing, painting, photography, music, and dance) at the school. The school did not offer intramural athletic programs. Most graduates attended college.

As part of a large research and development effort, the teachers in the science program had been working with educational researchers from a local university to develop and implement a three year, integrated, project-based science curriculum for all students called Foundations of Science (FOS) (Huebel-Drake, Finkel, Stern, & Mouradian, 1995). Foundations of Science (FOS) had been phased in as the mandatory

science curriculum at this school, replacing separate earth science, biology, and chemistry courses at the ninth (FOS-I), tenth (FOS-II), and eleventh grade levels (FOS-III)⁵. Throughout the year, students studied scientific subject matter by investigating broad questions and creating artifacts.

During the 1993-1994 school year, the FOS program was piloted in one ninth-grade class. The program was extended to the entire ninth grade the next year (1994-1995) while being piloted in the tenth grade. FOS was expanded into the entire tenth grade in 1995-96 and into the eleventh grade in 1996-97. All incoming freshmen (n≈100) were required to participate in FOS I (Foundations of Science I - The Tools of Science) during the fall of 1996.

The 1996-97 school year also saw a change in the school scheduling system from a seven x45 minute period day to a block-eight schedule with four 90-minute blocks Monday through Thursday and seven short 45 minute periods on Friday. With this scheduling system, FOS-I met for long (90 minute) periods on Monday and Wednesday, and for short (45 minute) periods on Friday. If the school calendar had only a four-day week, then the Friday schedule was adjusted to mirror the missing long-period day.

The Classroom

The building was a former elementary school so classrooms were relatively small. The FOS classroom was furnished with two-student lab tables and chairs instead of desks. Lab counters ran the length of two sides of the classroom and provide space for the sinks and gas jet fixtures. A large window took up the rear wall while the front wall

⁵ A separate (non-project-based) physics course remained in the school curriculum.

included the blackboard, projection screen, teacher's desk and demonstration table.

Power cables and the computer network lines had been dropped from the ceiling to three points in the center of the classroom. The students' lab tables were arranged in clusters around these power and network points. Student lab teams shared *Hach*[®] and *Lamotte*[®] water testing kits, microscopes, and other basic science equipment.

The Students

All incoming freshmen plus a few transfer students in other grade levels who had not completed the course ($n_s=99$, 41 females) divided into four sections/blocks of approximately 25 students each (Table 3.1.)

Table 3.1: FOS I (ninth grade) student subjects by teacher and class.

Class Block	Teacher	Number of Students	Female Students	"Specials" [†] (CARE, ICL)
1	Ms. K	25	10 (40%)	3 (12%)
3	Ms. L	27	14 (52%)	5 (19%)
5	Ms. K.	22	7 (32%)	12 (56%)
7	Ms. L.	25	10 (40%)	None ID
Total		99	41 (41%)	20 (20%)

[†] "specials" refers to those students who had been identified by the school as having special needs and were provided additional resources by the school.

All students agreed to participate in the study (See Appendix A for consent letter; HSR: Enhancing the Teaching of Science Projects, 11/94). The students had a range of racial, academic, and socioeconomic characteristics although the majority of students were white, middle- to upper middle-class.

The Teachers

Two teachers each taught two sections of Foundations of Science I (FOS I) during the semester of this study. The two teachers had different levels of both general teaching experience and experience with the FOS program.

Ms. L.⁶ did her student teaching at this school, and is an alumna of the local university's education program with a major in Biology. She joined the school's staff in 1990 and was one of the original architects of the FOS-I curriculum. She had taught FOS-I for three years before this study and prior to that, she had three years teaching experience in this science department.

Ms. K. was newer to the FOS program, starting in 1995. She came to the school with one year of teaching experience in chemistry and physical science at another high school in the same school district. She had also taught biology, chemistry and physics at a small university in the UK.

A National Science Foundation grant supported the common planning time for these two teachers. Ms. L. and Ms. K. routinely planned the FOS-I class together so that all FOS-I students would be engaged in the same activities at approximately the same time. They shared teaching notes and handouts. They were supported in planning and implementation by two other experienced science teachers and by university personnel.

Two student teachers from the university's teacher education program were also present in the classrooms. Ms. M was placed in Ms. K.'s classes and was an active teacher in these classes. Her background was in biology and physical science. Ms A, with a background in biology and environmental science, was placed with one of the other science teachers but often assisted with Ms. L.'s classes, especially during field trips.

⁶ Pseudonyms are used for the teachers and student names.

The Researcher

During the summer of 1996, I met with Ms. L. and Ms. K., the two FOS-I teachers, to review the curriculum design for the Traver Creek Project and to assist them with revisions they wished to make. Specifically, we revised the physical assessment part of the creek curriculum to facilitate its integration with the biological assessments. We discussed the sequence of instructional activities and outlined a tentative calendar for the semester. Throughout the semester, I continued to attend the planning sessions (~3 hours per week).

During the semester, I was in daily attendance for two of the FOS-I blocks (one class for each teacher). I participated in the class both as a researcher with full disclosure to the students as to the purpose of this study and as a content specialist (expert on streams). I interacted with the students if they had questions about the content or any of the tools that they were using.

The Classroom Culture and Curriculum

Although the physical appearance of the classroom would be familiar to science educators, the classroom culture and curriculum of FOS differed from a traditional science course in several ways. First, the pedagogical approach was based upon the Project-based science model (Blumenfeld, et al., 1991; Krajcik et al, 2000 Marx, et al., 1997; Krajcik, et al. 1999) and each of the courses were arranged around investigations of relevant “driving questions”. For instance, the question in FOS-I during the fall semester, “Is Traver Creek Ecologically Balanced?” provided students an opportunity to explore the various biological, physical, and chemical aspects of their creek and to investigate connections between each of these factors (Huebel-Drake, et al., 1995). The second

semester explored the driving question: "Is our climate changing? Does it matter?" which provided students with an opportunity to explore weather, global change and paleoclimatic effects.

Second, virtually all work in FOS-I was done in groups of two to four students, facilitating collaboration and communication. The teachers made the assignments to teams to collect data on certain sections of the creek; but for other activities, students were free to choose their own team members.

Third, the FOS-I curriculum was highly integrated with and supported by computer technology. Students used computers as tools to gather information through telecommunications and probeware, analyze data, express results graphically and/or pictorially, create scientific models, and write reports. The classroom was equipped with Apple Powerbook laptop computers, four Power Macs, an AppleTalk network connection to a server and laser printers, and an ISDN network that connected each machine to the Internet. A laptop computer was available for every two students and students could check out laptops overnight.

Several computer applications were available on the PowerBooks. Claris®Works⁷ software was routinely used by the students and teachers for word processing, drawing and spreadsheet functionality. FOS students also had access to a specially designed software-suite of investigation tools. During the fall of 1996, two pieces from this suite were used in the classroom: Riverbank and Model-It 3.0b (Jackson, Stratford, Krajcik & Soloway, 1996a, 1996b; Metcalf, Krajcik, & Soloway, 2000).

⁷ A precursor of AppleWorks

Riverbank, a data-storage program designed to serve as a standardized tool for collecting and assessing watershed monitoring data, allowed students to input and store their data from site surveys, benthic analyses and physical/chemical water quality tests and then share that data via e-mail and the World Wide Web.

Model-It 3.0b (Metcalf, 1999) provided a dynamic modeling environment designed specifically for learners unfamiliar with dynamic modeling, mathematical or symbol manipulation. In this application, models consisted of objects (“things” in the system being modeled e.g. the stream or benthic macroinvertebrates), factors (measurable attributes of objects, e.g. pH, temperature, diversity), and relationships between factors (as pH decreases, diversity decreases). Objects were represented visually with photo-realistic or graphical images. Factors were defined qualitatively with text (high to low) or quantitatively (pH range from 1-14). Relationships were defined with text (as temperature increases, DO decreases), graphically or numerically. This software employed an approach of guided learner adaptable scaffolding (learner-centered designed), in which the learner controlled the fading of help features and the activation of more sophisticated options with guidance and support provided by the software (Jackson, Krajcik & Soloway, 2001) and by others in the classroom.

Instruction on computers and software use was integrated into the general science instruction, and the teachers relied on assistance from student helpers, student teachers and research assistants for technical and classroom assistance.

This study focused on the ninth grade, fall semester FOS-I project on Traver Creek. At the core of the Traver Creek project was the driving question: “Is Traver Creek ecologically healthy?” This project was chosen because (1) the content was

interdisciplinary, combining content from earth science, biology and chemistry; (2) the project had the potential of meeting a number of science standards and state objectives; (3) studying the creek was a meaningful, relevant activity as the creek was within walking distance of the school and flowed into a nearby river that was the primary source of the community's drinking water; and (4) the project provided an opportunity to build classroom-community connections because the local watershed council had established an "Adopt-A-Creek" program and served as an organizer and clearinghouse for water quality information between local citizens and government agencies throughout the watershed.

The first class day was spent on organizational activities and the completing of a pre-test instrument (described later in the chapter). On the second class day, students took a walking field trip to the creek. On the way, the teachers pointed out features of the watershed such as the storm drain, railroad tracks, parks, dam, power station, and other land uses. Students were provided a worksheet to support their observations and note taking. The next class day, the teachers provided more information about the area history and clarified the assignment for students to write up. Students were given an orientation to the computer technology in the classroom and class time to work on their essays, and later in the week, class time for peer critique on the essays. These essays on students' first impressions of the creek were the first artifact selected for content analysis.

The next major activity was mapping the catchment area (Stapp & Mitchell, 1995 - Activity 1.4, p 127.) using USGS topographic maps of the area. This activity had students identifying the boundaries of the watershed, the headwaters, mouth, and tributaries of the creek, major land uses in the catchment area, the size of the catchment

and the average gradient of the stream. This information was later incorporated into their reports.

On day seven, after these orientation activities, groups of 3-5 students were assigned a unique 30-foot (~10 meters)⁸ section of the creek. Between the four class sections of FOS-1, approximately 700 feet (216 m) of the stream were surveyed and studied. Over the next three weeks, students conducted physical, biological and chemical assessments of their sections of the stream.

Physical assessments focused on habitat components, including substrate, current, bank attributes, vegetation, and the topography of that section. The biological assessment involved a qualitative sampling using D-frame kick nets of the section of the stream in order to collect the greatest diversity of benthic macro-invertebrates. The benthics were preserved in 70% alcohol and taken back to the classroom where they were identified to the level of biological order and classified as belonging to one of three pollution tolerance groups (“taxa groups”) described in the students’ text (Mitchell & Stapp, 1994).

The chemical assessments followed the protocol for the nine water quality parameters detailed in the *Field Manual for Water Quality Monitoring* (Mitchell & Stapp, 1994): percent saturation of dissolved oxygen (DO), fecal coliform bacteria counts (fc), pH, biochemical oxygen demand (BOD), temperature (°T), total phosphates (tPO₄), nitrates (NO_x), turbidity, and total solids (TS).

⁸Although metric units are the preferred unit of measurement in American science classes (NRC, 1996), English units were used in the creek project due to the data reporting requirements of the watershed council.

After completing the physical and biological assessments, students were asked to make a presentation to the class about their 30-foot section of the creek. These presentations were to include a summary of the physical nature of their section and a report on the benthics found in their section. Students were encouraged to consider how conditions up and down stream of their section might affect their findings. Students were also encouraged to ask questions during the presentations and to provide peer feedback on the presentations. Following the presentations, the groups were asked to incorporate the feedback they received and to write a report on their investigation that would be shared with the watershed council. The reports were written by the student teams over a period of two weeks and were the second artifact selected for analysis.

In early November, when the reports had been completed, students were introduced to modeling and the Model-It 3.0b software (Jackson, 1995; Jackson, et al., 1996a, 1996b) using the non-content scenario: "What affects my grade in science class?" Students worked in pairs during the practice session (one 90 min. period). Once they had become familiar with the software, students were instructed to create a model relating the physical and biological factors in the creek and teams were given one long (90 min.) and one short (45 min.) class period to build their models. This set of models was the third artifact selected for analysis.

Following the construction of the models, the focus of instruction moved away from the physical and biological properties of the creek to water chemistry. For the rest of the month, concepts such as physical and chemical changes, mixtures and solutions, acids and bases, periodicity and the periodic table, naming elements and compounds and

balancing chemical equations were introduced, explored, and related to the chemical tests students had conducted at the creek earlier in the term (Appendix B).

During the summer planning sessions, the teachers had decided to have students complete small research projects as part of the creek curriculum. But as the semester moved along, the student research projects were dropped because of time constraints.

In December, the focus of instruction shifted back to the biological domain as the class addressed topics on bacteria, photosynthesis and respiration. The latter two concepts were tied to students' prior findings about dissolved oxygen and its important role in the creek. Students were given a mini-assignment to develop a representation demonstrating the inter-relationship of photosynthesis and respiration in the creek.

After the winter break, a couple of class periods were spent reviewing what students had done over the course of the semester and tying together all their findings to determine the overall quality of the creek. Students were given their final assignment, "build a model that demonstrates in-depth understanding of a stream ecosystem." Before beginning construction of their final models, students critiqued some models from the first cycle of construction and were introduced to additional features of the Model-It software. Teams of students spent one week constructing these models, which were the fourth artifact, selected for analysis. Each of the four artifacts is described in more detail in the section on data sources.

Weather days shortened the instructional time between the winter break and the end of the semester. The teachers had to rearrange their planned activities and the public presentations of the final models were dropped. During the school's final exam period, students completed the post-test instrument.

Over the course of the semester, students conducted physical, biological and chemical assessments of the creek. They collaborated to make sense of their data and to hypothesize different relationships between stream factors. They used their findings from the assessments and data analysis to build computer models of the aquatic ecosystem and to craft written reports to be shared with the local watershed council.

The next section describes the approach used to capture and describe students' scientific understandings demonstrated in this environment.

Research approach

This study describes students' scientific understandings as represented in artifacts constructed during the Traver Creek project. The descriptions are based on content analysis of multiple data sources, specifically the multiple artifacts produced by each student. The *Standards* (NRC, 1996) were used throughout the study as the "target" for student understandings in terms of both breadth and depth/quality of understandings. The approach followed a five step process: (1) The delineation of the curriculum and mapping it onto the National Science Education Standards. (2) The identification of opportunities (and expectations) to demonstrate understanding in each of the artifacts and the pre/post instrument. (3) Analysis of student understandings demonstrated in each of the assessments. (4) Comparison of course content and artifact specifications to actual demonstrations of understanding across the artifacts. And, (5) the examination of student understandings across time. Each of these steps is described below.

Delineation of the curriculum and mapping course content to standards

The identification of science content and process skills addressed during the creek project began during the summer meetings with the teachers. The teachers described the different activities that they had used in previous years and shared handouts, quizzes, and other curricular materials. We made some revisions as described above and devised a calendar for the project. From these meetings and materials, I was able to start mapping the project content onto the *Standards*.

Through out the semester, I made classroom observations during each of the class periods for two sections, one section taught by each teacher. Field notes were used to capture the major events of the day and to note particular episodes or questions for closer attention. A video camera was placed in one corner of the classroom and was used to collect videotape on all classroom activities as a back up to the field notes. Observations focused on the teacher as she introduced the project, introduced lessons, led discussions and organized the students into working groups. Observations focused on students during class discussions, group work and student presentations. Field notes and video were also collected during the class excursions to the field sites at the creek. The classroom observations in the two classrooms were used to describe the content coverage of the project and teachers' verbal expectations and instructions on the artifact construction. These in turn were matched to relevant *Standards* (Appendix B).

For example, as students began to study water chemistry in early November, they were introduced to the scientific concepts of solutions and mixtures. These concepts were important because some of the water chemistry tests they conducted involved substances in solution (i.e. dissolved oxygen, nitrates, phosphates, and total solids) while

turbidity is affected by suspensions (mixtures). In the classroom students investigated various substances ranging from orange juice, chicken bouillon, and beans and rice to chemical solutions made from copper sulfate or sodium chloride dissolved in hydrogen hydroxide. Students could manipulate these substances with filters and by evaporating small samples. These manipulations led to the next major content idea, physical vs. chemical changes. For although mixtures and solutions can be separated by physical means (i.e. dissolved oxygen can be forced out of solution by raising the temperature of the water), in water chemistry, it is often necessary to cause a chemical reaction in which the substance of interest is chemically bound to another substance so it can be measured. For example, the students used a modified Winker test (reduction-oxidation reaction) in which iodine is released in proportion to the amount of oxygen. The resulting solution, which is colored, was titrated to determine how much iodine was released, an indication of the amount of oxygen that the water held prior to the reactions.

This part of the curriculum was basic physical science. References to solutions, mixtures, physical changes and chemical changes are common in science textbooks. However, “solutions,” “mixtures” and “physical changes” are not found in the text of the *Standards*.⁹ Chemical reactions on the other hand, have a whole sub-section of the *Standards* in Content Standard B – The Physical Sciences. Consequently, mapping course content required some interpretation of the standards.

⁹ The word “solution” does appear in Content Standard E relating to technological designs. The word “mixture” does appear in standard C.1.1 in reference to the compounds found in cytoplasm.

This part of the curriculum fit into two major substandards: B.2. Structure And Properties Of Matter and B.3. Chemical Reactions, which are each, again divided. Within substandard B.2 "Structure And Properties Of Matter" are six finer divisions:

B.2.1. Atoms interact with one another by transferring or sharing electrons that are furthest from the nucleus. These outer electrons govern the chemical properties of the element.

B.2.2. An element is composed of a single type of atom. When elements are listed in order according to the number of protons (called the atomic number), repeating patterns of physical and chemical properties identify families of elements with similar properties. This "Periodic Table" is a consequence of the repeating pattern of outermost electrons and their permitted energies.

B.2.3. Bonds between atoms are created when electrons are paired up by being transferred or shared. A substance composed of a single kind of atom is called an element. The atoms may be bonded together into molecules or crystalline solids. A compound is formed when two or more kinds of atoms bind together chemically.

B.2.4. The physical properties of compounds reflect the nature of the interactions among its molecules. These interactions are determined by the structure of the molecule, including the constituent atoms and the distances and angles between them.

B.2.5. Solids, liquids, and gases differ in the distances and angles between molecules or atoms and therefore the energy that binds them together. In solids the structure is nearly rigid; in liquids molecules or atoms move around each other but do not move apart; and in gases molecules or atoms move almost independently of each other and are mostly far apart.

B.2.6. Carbon atoms can bond to one another in chains, rings, and branching networks to form a variety of structures, including synthetic polymers, oils, and the large molecules essential to life.
(NRC, 1996, Chap 6)

Standard B.2.4 was chosen for the concepts of mixtures and solutions because of its emphasis on the physical properties of matter. Standard B.2.4 was also used for content about physical changes that involve changes of state.

The idea of chemical reactions was mapped onto the division in Standard B.3, and in particular, B.3.1 – Chemical reactions occur all around us and B.3.3 – Oxidation/reduction reactions. Students' investigations of mixtures and solutions also mapped onto the Earth and Space Science – Geochemical Cycles (Content Standard D, NRC 1996) as students explored mixtures and solutions in the context of water, carbon and nutrient cycles (Appendix B).

In addition to the classroom observations, all handouts (assignments, notes, quizzes, etc.) distributed during class were also collected. These documents were also used to characterize the curriculum content and the teachers' criteria for artifact construction. This information from the classroom observations and instructional handouts was used to map out the course content and create an analysis tool, the Matrix of Understanding Standards (Appendix B).

This matrix was divided into sections differentiating the conceptual understandings from other scientific understandings, but basically followed the structural organization of the *Standards*. For example, the conceptual understanding standards are presented in the following order:

CONTENT STANDARD B: Physical Science

CONTENT STANDARD C: Life Science

CONTENT STANDARD D: Earth and Space Science

CONTENT STANDARD F: Science in Personal and Social Perspectives

Content Standards A, E and G are not represented in the conceptual understanding section as their standards apply to the strategic and epistemological understanding dimensions. Table 3.2 shows a small piece of the matrix from Appendix B for discussion purposes.

Table 3.2: Excerpt from Appendix B: Matrix of Understanding Standards.

CONCEPTUAL UNDERSTANDING STANDARDS Content Standard B: Physical Science									
B.1. STRUCTURE OF ATOMS	Traver Creek Curr.	OTL	E/O	Pre Test	Essay	Report	Model 1	Model 2	Post test
B.1.1. Matter is made of minute particles called atoms, and atoms are composed of even smaller components. These components have measurable properties, such as mass and electrical charge. Each	11/15 structure of atoms 11/20 electron shells 11/18 video on atom Bohr models		E						
			O	2.25					2.54
B.1.2. The atom's nucleus is composed of protons and neutrons, which are much more massive than electrons. When an element has atoms that differ in the number of neutrons, these atoms are called	chemical vs. physical changes 11/18 simple replacement reactions. - Spot plate lab		E						
			O						
B.1.3. The nuclear forces that hold the nucleus of an atom together, at nuclear distances, are usually stronger than the electric forces that would make it fly apart. Nuclear reactions convert a fraction of the			E						
			O						
B.1.4. Radioactive isotopes are unstable and undergo spontaneous nuclear reactions, emitting particles and/or wavelike radiation. The decay of any one nucleus cannot be predicted, but a large group of			E						
			O						

OTL means "opportunity to learn," for standards addressed by the curriculum. E/O means "Expected in assessment" and "Observed in assessment."

In the matrix (Table 3.2), the first column is a unique three-character code for each standard used throughout this study as a reference for the *Standards*. The code for each standard was not given in the official document but was derived by using the *Standards'* major designations (Content Standards A-G) and then sequentially numbering the sub-standards below each designation. The second column in the matrix includes the full text of each of the standards for grades 9-12 (NRC 1996). The third column includes content that was observed during the Traver Creek project. The fourth, labeled OTL for "Opportunities to Learn," served a database function to keep a record of the number of standards addressed in the curriculum. Column five, labeled E/O for "Expected and Observed" was used to organize the analysis of which standards were expected to be assessed in a particular artifact or instrument and whether or not it was

observed in that assessment. The remaining columns represent each of the six assessments used in the study. These assessments are described in the next section.

The understandings matrix was used as a major organizational tool for the content analysis of the different measures. After mapping all expected assessments of standards, the standards for which there was no "Opportunity to Learn" nor expected assessments were deleted from the matrix, reducing the number of standards under analysis from 90 to 45. This abbreviated matrix was used to create a nested database with the content standards, each of the six measures (four artifacts and the pre/post instrument) and within each measure, the individual students who completed that assessment. The nested database was used to record which individual's demonstrated understandings for a standard in each assessment and an evaluation of the level or quality of the understanding.

Identification of content expectations in the selected measures

Purposeful sampling was used to select which artifacts were collected for analysis. Throughout the semester, the ninth grade students produced a plethora of artifacts assigned by their teachers. Some of these required relatively little student effort such as fill-in-the-blank homework assignments, summaries of reading assignments, quizzes, activity handouts, etc. and potentially represent more knowledge recall than understanding. Other artifacts, such as the written reports and computer models, required a greater time investment, were constructed over a number of days, and often underwent cycles of revision. These were the assessments where questions and tasks could be discussed, clarified, and even appropriately modified through discussion with the teacher and/or the student's peers. Therefore, because of the costs in time and effort in the

analysis, as well as the difficulty in collecting artifacts from approximately 100 students, I restricted the content analysis of student artifacts to those major artifacts that entailed more effort on the part of the students and were supported by assessment criteria set by the teachers: descriptive essays, scientific reports, and two computer models. For each of these four artifacts, students were provided classroom time for construction and revision over multiple class periods. These artifacts were a substantial part of the assessment by which the classroom teachers determined student achievement and assigned course grades. Each of these major artifacts is described below. Table 3.3 summarizes the data sources among the artifacts and pre/post test instruments that were used to determine students' understandings.

Table 3.3: Data sources used to determine student scientific understandings

DATA SOURCE							
	Class room Observ	Pre-Tests	Essays	Reports	Model 1	Model 2	Post - Tests
Time (week)	1-19	1	2-3	9	11	18	19
Group Size	whole class	individual	individual	sm grps (3-5 S's)	dyads	dyads	individual
Number of artifacts (n_a)		103	45	23	48	46	97
Number of students (n_s)		103	45	84	84	85	97 ¹⁰

The use of student generated artifacts for content analysis was a relatively non-reactive and unobtrusive form of data collection (Jones, 1985). It used data generated as a by-product of other endeavors and/or data collected for other purposes; in this case data

¹⁰ There were 96 students who completed both pre and post tests, three students dropped the class after the first month and so their data were eliminated from the analysis.

produced for the purpose of teacher assessment of student outcomes. In addition, the selected artifacts were constructed at different points during the semester, interspersed with periods of instruction and investigation. Thus, they may reveal a developmental sequence of understandings.

The analysis of artifacts and the pre/post tests examined the opportunities and constraints provided by the task structure and media of the different artifacts for students to express their scientific understandings. This analysis sought the answer to the first research question: "What opportunities are provided by the different artifacts for students to express their scientific understandings?" To address this question and to guide the analysis, four subquestions were created.

- How much of the course content can be potentially represented by each artifact?
- How much of the course content is represented by the set of artifacts?
- Which parts of the curriculum are not represented by the artifacts?
- Are some constructs over-represented and others under-represented by the set of artifacts?

The analysis of artifact expectations and opportunities included close examination of written materials such as project guidelines and handouts, assessment criteria presented by the teachers, teacher explanations in class, and the capabilities and features of the technology employed in the artifact construction (e.g. ClarisWorks spreadsheets and charts for the report and features of Model-It for the creation of the computer models.).

The analysis of artifact expectations used the matrix of understanding standards (Appendix B). The nature of each artifact was examined and the standards in the matrix

that it explicitly and implicitly addressed (E for expected standards) were identified. An example of this process is provided below in the essay section.

Descriptive Essays

The descriptive essays were based on students' first visit to the creek. These essays represent students' initial perceptions of the creek and their connections of this experience to prior knowledge as well as demonstrations of epistemological and strategic understandings in presenting and supporting a hypothesis. Students were supported in the essay creation by the provision of note taking sheets during the initial site visit (Note-taking assignment at Traver Creek) and a handout specifying the criteria for the essays (Appendix C).

In the instructions, students were asked to write a six to seven paragraph essay. In the first paragraph, they should describe the location of Traver Creek and their purpose for visiting it. They were also asked to make a thesis statement about the creek, defined as a judgment about the creek and the surrounding area (Handout text – Appendix C). In the body of the essay, students were to expand upon three areas of observation: the general surroundings, the creek's banks and the creek itself, providing details in each section to support their thesis.

There were numerous opportunities for demonstrating understandings in this assignment. Specifically, the essay assignment provided opportunities to demonstrate conceptual understandings about the use of resources (F.3.1) and natural systems (F.3.3) related to land use and human impacts such as pollution on the watershed (C.4.5 and F.6.5). Students could also write about interactions among earth systems (D.3.3) such as the effects of weather and erosion (F.5.3) on organisms and habitats in the creek. To a

lesser extent, they might convey understandings about the organization of the biological organisms of the creek (C.3.5), the effect of sun and shading (D.1.1) and how their own personal actions might impact the creek or how the creek impressed them (F.1.3). Other conceptual scientific understandings could be expressed in the essays even though they were not expressly specified by the task. An example of such an understanding would be if students included concepts of photosynthesis and respiration (C.5.2) in the description of how shading might affect the creek.

Students had in-class time for peer critique of their essays and to make revisions. Photocopies of the essays were made for research purposes. These essays represent artifacts created by individuals (Table 3.3).

Scientific Reports

The scientific reports on the benthic and physical assessments were assigned after the data was collected and analyzed by student groups toward the middle of the project. Students were grouped heterogeneously by the teacher and assigned to specific 30-foot sections of the creek. Ms. L. and Ms. K. informed the students that their data, in the form of reports, would be sent to the local Watershed Council.

Students were given a detailed handout (Appendix D) describing the parts of their reports. This handout was reviewed during class time on October 21st and additional examples were given. The details include very specific instructions and check lists. For example, for the report introduction, students were directed as follows:

INTRODUCTION (About 2 paragraphs)

The introduction should provide a context for the topic under study. The introduction provides the background necessary to understand the rest of the report. In addition the introduction should provide a concise

statement of the problem. That is, tell precisely what questions you are trying to answer.

Suggestions for what you should have in the introduction:

____ What question were you trying to answer about Traver Creek?

____ A description of what benthic organisms are and why they are used as indicators of health for the creek.

____ A description of physical forces that influence the benthic community. (Class handout 10/21/96 – Appendix D.)

Just in this short piece of the assignment, it is possible to identify multiple matches to the standards. The descriptions “of what benthic organisms are” map onto standard C.3.5 (biological classification)¹¹. Why benthics are used as indicators of a healthy creek maps onto standard C.5.5 (niches and environmental tolerances). The description of the physical forces that influence the benthic community maps onto standards D.3.3 (interactions in earth systems) and C.5.5 (niches). The remaining standards assessed in this artifact and the specific opportunities provided by the remaining measures are discussed in the findings in Chapter Four.

Students were first directed to write up the titles, introductions and discussion sections as individual homework assignments and to bring those pieces to work on together as a team during the next class period. Second, students were asked to write the methods and results sections as a team in class. Teams then divided the different sections of the report and assigned members to write the different pieces and a second member to revise each piece.

Final copies of the group reports were collected and photocopies made for research analysis.

Computer Models 1 and 2

During the Traver Creek project, there were two cycles of computer modeling using Model-It 3.0b. The first model chosen for content analysis represented the integration of subject matter knowledge across the first nine weeks of the project. The second model represented content integration across the entire project. In both cycles of modeling, most students worked in pairs.

During the first model building cycle (Nov. 4-8), students were given one long (90 min.) and one short (45 min.) class period to build a model of the physical and biological factors in the study creek and to test their models. The handout (Appendix E) specified that ...

Your model must convincingly demonstrate your understanding of the physical factors you've chosen to model and how they relate to the benthics. Make sure that you fill out a plan, describe each object and factor, explain each relationship, test your model as you go along, and evaluate it at the end. (Emphasis in original - Class Handout 11/4/1996 - Appendix E)

In building the model, students had available a selection of clip art images and digital photos of their study site on the classroom server as attested by further instructions in the class handout.

Once you have some ideas, you're ready to start building a model. On the server, in the FOS1 folder, there are files with pictures of different sections of the stream, and of different kinds of benthics, for you to use in your model. Make sure you drag the picture files onto the desktop of your

¹¹ Benthic Macro-Invertebrates or BMI's include orders of insects (mayflies, stoneflies, odonates, hemipterans, diptera, etc.), crustaceans (isopods, amphipods, crayfish), molluscs (snails, clams, limpets), and annelids (tubifex, aquatic worms, leeches).

powerbook before opening them, so you don't tie up the server. (Emphasis in original - Class Handout 11/4/1996 – Appendix E)

The second modeling cycle occurred at the end of the semester. For this assignment, students were asked to build a model that demonstrated in-depth understanding of a stream ecosystem and that included physical, chemical, and biological factors of the stream. The assignment sheet (Appendix F) was handed out on January 8th and final models were due January 17th. Students had access to the same folder of images and clip art on the class server as for the first assignment.

Students saved their completed computer models to the classroom server. These files were then copied and collected for analysis. Printouts of each model, which include students' plans, goals, descriptions of factors and objects, explanations of relationships, and the student's evaluations of their models were made. Only the final versions of the models were analyzed.

Pre/Post Test Instruments

A pre-/post-test instrument was constructed to examine students' scientific understandings before and after the semester long project. The pre/post test instrument (Appendices G & H) consisted of 31 questions across the Traver Creek curriculum. Some of these questions included multiple parts for a total of 66 items. The instrument included Likert scale rankings, multiple-choice items, agree/disagree statements with explanations, constructed responses, hypothetical actions based on scenarios and a concept map. The instrument items were drawn from a prior study involving stream investigations (Talsma, 1992) and teacher generated questions from quizzes and exams used during the creek project in previous years.

The test instrument was constructed in collaboration with the teachers, Ms. L and Ms. K, to represent content across the semester. The test items were then mapped onto the *Standards*. The *Standards* measured by the pre/post instrument were included in the standards matrix (Appendix B).

Most short response questions addressed only one standard with two or more questions addressing that standard. The longer constructed response questions often addressed two or more standards within a single question. For example, in response to question 15:

15. What would happen if extra phosphates enter the creek? Describe this process in terms of photosynthesis, cellular respiration, Dissolved oxygen (DO) and Biochemical Oxygen Demand (BOD). Include any relevant chemical equations in your answer. (pre-post instrument, Appendices G & H)

A complete answer would contain understandings that would map onto multiple conceptual standards as in the following example:

Extra phosphates might come from human wastes, animal wastes, fertilizers, human disturbance of land & vegetation, or draining of wetlands^(F.5.2). Phosphates are a plant nutrient (fertilizer)^(C.5.5). In Aquatic systems, phosphates are often limited. Extra phosphates in the creek will cause algae to grow. Rapid or exponential growth is also known as an algal bloom.^(F.2.1) Algae carry out photosynthesis (fix energy in the form of carbon in the presence of sunlight) ($6\text{CO}_2 + 6\text{H}_2\text{O} \implies \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$.) More algae carrying out Ps means more oxygen will be produced during the day (increased dissolved oxygen or DO).^(C.5.2) Algae (like most living things) carry out cellular respiration ($\text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2 \implies 6\text{CO}_2 + 6\text{H}_2\text{O}$). At night, use of O_2 exceeds production and DO drops^(C.5.2). After the phosphates are used up, the alga dies and begins to decompose. Decomposition by bacteria requires DO (= BOD) and so DO levels drop further. This process uses aerobic respiration and energy is released^(C.5.1). Other living things (e.g. benthics) may be affected by the changing DO levels. E.g., Taxa 1 organisms such as mayflies cannot survive in low DO waters^(C.5.5). (Answer generated by content specialist, Sept. 1996)

Figure 3.1 shows how the statements above, annotated with the *Standards'* codes, match

Figure 3.1: Example of mapping standards onto a response to Pre/Post Test Question 15.

<p align="center">Conceptual Understanding</p> <p align="center">Standards</p>	<p align="center">Understandings represented in</p> <p align="center">response</p>
<p>C.5.1 - All matter tends toward more disorganized states. With death and the cessation of energy input, living things rapidly disintegrate.</p>	<p>After the phosphates are used up, the algae dies and begins to decompose. Decomposition by bacteria requires DO (= increasing BOD) and so DO levels drop further. This process uses aerobic respiration and energy is released.</p>
<p>C.5.2: The energy for life primarily derives from the sun. Plants capture energy by absorbing light and using it to form strong (covalent) chemical bonds between the atoms of carbon-containing (organic) molecules. In addition, the energy stored in bonds between the atoms (chemical energy) can be used as sources of energy for life processes.</p>	<p>Algae carry out photosynthesis (fix energy in the form of carbon in the presence of sunlight ($6\text{CO}_2 + 6\text{H}_2\text{O} \Rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$.) More algae carrying out Ps means more oxygen will be produced during the day than is consumed (increased DO) Algae (like most living things) carry out cellular respiration using O_2 to release energy. ($\text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2 \Rightarrow 6\text{CO}_2 + 6\text{H}_2\text{O}$). At night, use of O_2 exceeds production and DO drops.</p>
<p>C.5.5 The distribution and abundance of organisms and populations in ecosystems are limited by the availability of matter and energy and the ability of the ecosystem to recycle materials.</p>	<p>Phosphates are a plant nutrient (fertilizer). In Aquatic systems, phosphates are often limited. Adding phosphates means that more algae will grow. Other living things (e.g. benthics) may be affected by the changing DO levels. i.e. Taxa 1 organisms (e.g. mayflies) cannot survive in low DO waters</p>
<p>F.2.1. Populations grow or decline through the combined effects of births and deaths, and through emigration and immigration. Populations can increase through linear or exponential growth, with effects on resource use and environmental pollution.</p>	<p>Extra phosphates in the creek will cause algae to grow. Rapid or exponential growth is also known as an algal bloom</p>
<p>F.5.2. Human activities can enhance potential for hazards. Acquisition of resources, urban growth, and waste disposal can accelerate rates of natural change.</p>	<p>Extra phosphates might come from human wastes, animal wastes, fertilizers, human disturbance of land & vegetation, or draining of wetlands</p>

up with the actual content of the *Standards*. As illustrated in this example, explicit standards specified by this question include standards C.5.1, C.5.2, C.5.5, and F.2.1.

Standard F.5.2, which deals with human sources of phosphates, was not explicitly addressed by the question, but still may be included in student's answers.

Two versions of the pre/post instrument were constructed (Form A – Appendix G and Form B – Appendix H). The two forms were randomly distributed to students on the first day of class. The two versions of the instrument were compared using t-tests ($\alpha = .05$) on both individual items and responses collapsed across standards for both the pretest and the posttest. Upon finding no statistical difference between instrument forms, forms A and B were treated as a single data source in subsequent analyses. Students who completed Form-A on the pre-test were given Form-B for the post-test. Students who completed Form-B as the pre-test completed Form-A for the post-test.

The content analysis of student responses in the individual artifacts is described in the next sections.

Analyzing Student Artifacts for Conceptual Understandings

In general, the content analysis of artifacts involved careful examination of the artifact, scanning and re-scanning the artifact to identify understandings that matched one or more of the *Standards* (NRC, 1996). The expected standards for an artifact, as noted in the *Standards* matrix (Appendix B) helped to focus the search, but identification included all understandings that could be mapped onto any of the 45 conceptual standards addressed by the curriculum. Conceptual understandings were inferred especially from student descriptions and explanations of phenomena. Once the standards were identified and noted in the matrix, another reading of the artifact was used to score the quality of the understanding. A four level (0-3) coding scheme, adapted from Stratford (1996), was

used to classify the “quality” of the understandings for Standards B, C, D and F. These levels included:

Level 3: Explanation is scientifically correct to the level used in the *National Science Education Standards* (NRC, 1996) and contains no extraneous or incorrect ideas, statements concur with expert propositions (proficient or mastery level).¹²

Level 2: Explanation is partially correct but contains some extraneous and/or incorrect information (developing).

Level 1: Explanation contains substantial errors or fundamental differences between the students' and expert's conceptions as depicted in the *Standards* (non-scientific or novice).

Level 0: Student did not provide an explanation. OR, if some response is given, it does not evidence understanding, perhaps nonsensical (no evidence)

The quality levels of understanding were then also recorded in the matrix tool. Because these four levels represent an ordinal scale, non-parametric statistics were used for subsequent analysis, which will be described in a later section. In the following paragraphs, I provide more detailed examples of the content analysis of the students' artifacts and the pre/post test instrument.

Essays and Scientific Reports

Students' essays and scientific reports were analyzed through careful reading and then mapping of text contents onto the standards in the matrix (Appendix B).

Understandings were identified when two or more ideas were connected. Thus, lists of observations that closely resemble the note taking assignment were not counted as

¹² Note that a Level 3 code does not represent the highest quality of understanding. A content specialist would naturally demonstrate higher levels of understanding. The level

understandings. But when students made connections between two or more ideas, like shade affecting the level of dissolved oxygen in the creek, these were considered evidence of understanding and mapped onto the appropriate standards (e.g. B.2.5 and D.3.3).

After all the standards were identified, the quality of understanding for each standard was determined according to the coding schema above using a conservative approach. For example, if a student wrote, "We found high levels of dissolved oxygen in our section of the creek. We think this might be because we had a lot of shade in our section." This statement would be scored as Level 1 for B.2.5 dealing with gas solubility because there is no explanation of causality and Level 2 for standard D.3.3 for identifying an interaction between earth systems (biotic shade influencing an abiotic factor.). If instead the students had written:

We found high levels of dissolved oxygen in our section of the creek. We think this might be because we had a lot of shade in our section. Trees create shade which blocks the sun from reaching the creek. Because the water is shaded from the sun, it doesn't warm up as much. And cooler waters can hold more dissolved oxygen than warm waters. (Hypothetical example)

They would reveal a more robust understanding of the relationship between trees, shade, temperature, and dissolved oxygen. The second example would map onto standard D.1.1 - Sun as source of energy - Level 3, D.3.3 - interaction of earth systems - Level 3 (shade influences water temperature and thus DO levels), and B.2.5 - gas solubility - Level 2. A

3 represents a threshold value for the quality of scientific understanding expected of high school graduates in the *National Science Education Standards* (NRC, 1996).

B.2.5 Level 3 would be recorded if a student had correctly explained why warm waters hold less dissolved gases than cool waters.

All understandings identified in the written artifacts were recorded in the standards matrix (Appendix B). Examples of student work are included in the analysis in chapter 4. Where explanatory comments were deemed necessary, they are indicated in the quotes from student work by the square parenthesis, e.g. [comment]. However, most of the misspellings and grammatical errors were not emphasized with the [sic] designation. Since the examples of student work were copied electronically from their original artifacts, uses of creative spelling should be considered the students' work.

Computer Models

In the students' computer models, conceptual understandings were identified in student's descriptions of factors, relationships between factors and explanations for those relationships similar in scope to the identification of understandings in the essays and reports. A relatively simple model created by Chase, a student in Block 3, illustrates how student understandings were identified and characterized in their models.

Chase's Model of the Effects of Forest Fires

For his final model, Chase, a male student working alone (Field notes 1/8/97 FOS-B3) decided that he wanted to create of model "to show how a forest fire would affect various characteristics of the creek." In the planning of his model, Chase stated his purpose and defined his problem and he began to plan the model by filling out the fields in the planning window of the Model-It software (Figure 3.2). The scientific

understanding demonstrated in this part of the model map onto non-conceptual understanding standards.

Chase required a little assistance in manipulating the software during the beginning of the building phase until he regained familiarity (Field notes 1/11/97 FOS-B3). He then proceeded to construct a model with three objects, six factors, and nine relationships (Figure 3.3). In the descriptions of the factors and relationships, Chase demonstrated his conceptual understanding. For example, in his creation of the factor, "stream: total solids" (Figure 3.4), Chase demonstrated his conceptual understanding of this stream factor in two ways: first in his description of the factor, "Total solids are the dissolved matireals [sic] and the suspended matireals [sic]" and in his definition of the range, quantitatively from 0 to 500 mg/L. Since both Chase's description and his defined range are scientifically accurate, and because this understanding of total solids maps onto the standard on states of matter/mixtures and solutions (B.2.5), Chase's conceptual understanding on this standard was coded at Level 3.

Chase also demonstrated his conceptual understandings in building the factors and the relationships between factors. For example, in the relationship between total solids and water quality (Figure 3.4), Chase demonstrates a high degree of understanding (Level 3) in three areas: he has created a scientifically accurate relationship (Mitchell & Stapp, 1994 Chart 9: Total Solids (TS) Test Results, p. 84.); he has provided an accurate explanation for this relationship (Mitchell & Stapp, 1994, p. 71) and he has provided an elaborate explanation by listing more than one or two causes. Chase has actually listed four reasons: the reduction of water clarity due to increasing turbidity, a decrease in photosynthesis rates caused by a reduction in sunlight penetration, the possibility that

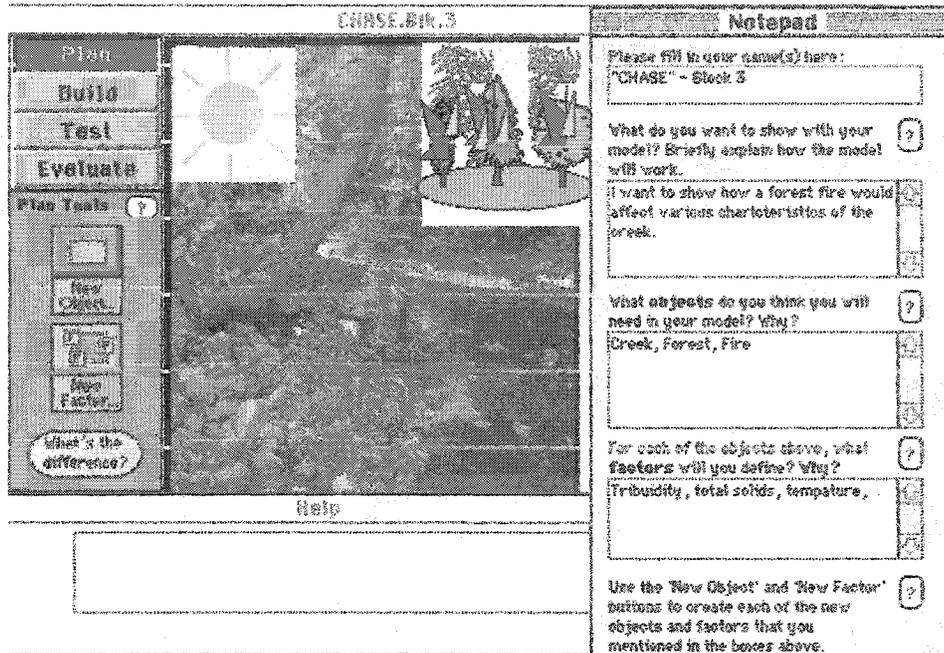


Figure 3.2: The Planning Stage for Chase's Model. The worldview on the left and the notepad, on the right, where the modeler can define the purpose of the model and plan the objects and factors that will be needed in the model.

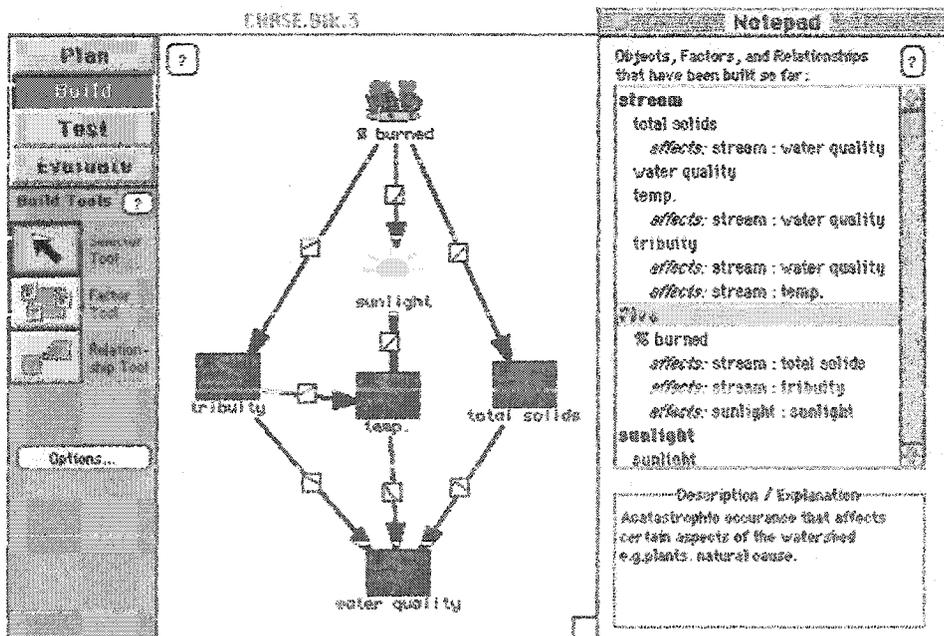


Figure 3.3: The "Build" window for Chase's Model. The factor map is on the left. The notepad on the right lists a summary of the objects, factors and relationships that have been built. Selecting one of these features (e.g. fire) causes Chase's description or explanation to be called up in the lower box of the notepad.

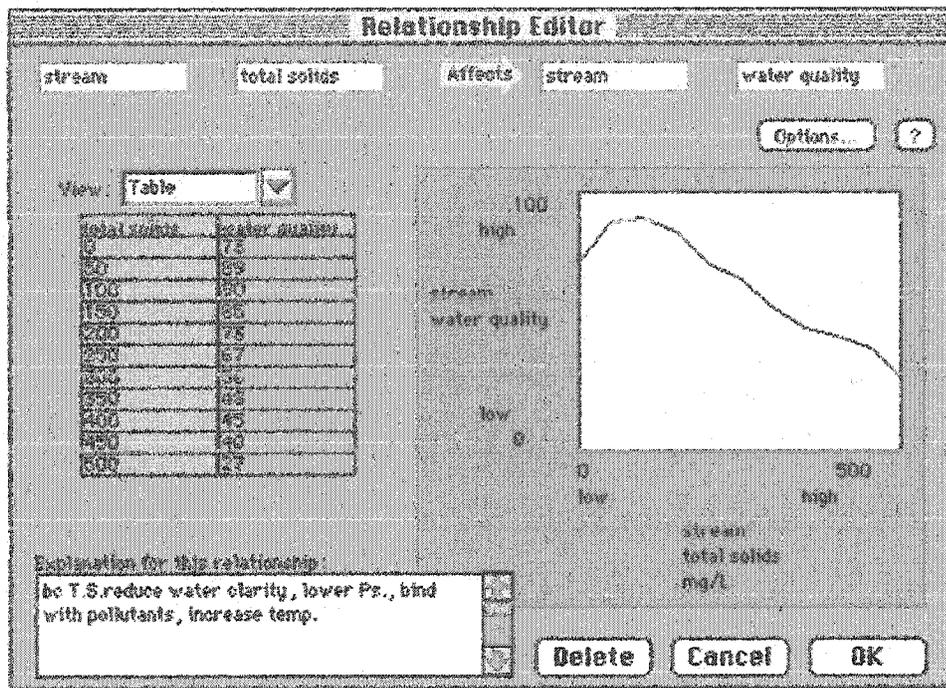


Figure 3.4: The relationship editor for the relationship Chase created between stream:total solids and stream:water quality. In this relationship, Chase has built a quantitative relationship using the table view option and then clicking on the graph. His explanation is both scientifically accurate and elaborate in that he includes more than one causal reason for the decrease in water quality (D.3.3 Interactions of Earth systems Level 3).

these materials will bind to pollutants, and an increase of stream temperature caused by the absorption of sun energy of the total solids (Figure 3.4). The content of this relationship maps onto Standard D.3.3 - Interactions of Earth Systems. Chase's demonstration of his understanding of this content standard was coded at Level 3, matching the expectations of the science standards.

In the earth and space sciences, Chase's model also addresses Standard D.1.1 - Sources of Energy. This understanding was demonstrated in three relationships: stream:tribuity [turbidity] affects stream:temp[ature], Fire:%burned affects sunlight:sunlight, and sunlight:sunlight affects stream:temp (Figure 3.5). In these three relationships, while the overall sense of the relationships is correct, there are some errors.

For example in the relationship between Fire:%burned and sunlight:sunlight, the shape of the relationship should be “increases by a little” not by “more and more.”

Chase has defined the initial amount of sunlight with no fires in the watershed at 50 and the initial amount of fire at 0%, but choosing the relationship “more and more” resets the value of sunlight to 0, a fact that he recognizes in the end when evaluating his model but one that he doesn’t correct. In addition, the banks of the stream comprise only a small part of the watershed that is burned, but this is the only area that would be shading the stream and thus affecting sunlight. Therefore, although a large amount of the watershed might burn, a much smaller percent of that affects the sunlight reaching the stream, thus sunlight increases by a little.. Another relationship, between sunlight and temperature (Figure 3.5), lacks an explanation. For these reasons, Chase’s demonstrated understandings along Standard D.1.1 were coded at Level 2.

<p>RELATIONSHIP: (Direct) stream : tributary affects stream : temp. As stream : tributary increases, stream : temp. INCREASES by A LITTLE. Explanation: waters become warmer as suspended particles absorb heat from sunlight.</p> <p>RELATIONSHIP: (Direct) Fire : % burned affects sunlight : sunlight As Fire : % burned increases, sunlight : sunlight INCREASES by MORE AND MORE. Explanation: bc sunlight is blocked by the trees when trees burn the sunlight can be absorbed by the water.</p> <p>RELATIONSHIP: (Direct) sunlight : sunlight affects stream : temp. As sunlight : sunlight increases, stream : temp. INCREASES by ABOUT THE SAME. Explanation: <Click on an object, factor, or relationship in the list above to see its description or explanation.></p>

Figure 3.5: Chases' relationships, which demonstrate his understanding of the sun as a source of energy (Standard D.1.1).

Chase's model included factors and content from the physical and chemical assessment of the creek. His model did not explicitly include biological factors, which

were part of the assignment, resulting in a lower assessment by the teachers who used the assignment criteria as a grading rubric. Never the less, Chase's overall model showed an elegance and parsimony that explained the essential effects of a forest fire on the creek.

Pre-/Post-tests

Each item on the pre/post tests was individually scored. Constrained responses (e.g. multiple choice) were scored as correct (Level 3) or incorrect/no response (Level 0). Constructed responses were scored by mapping student responses onto the standards and then coding for level of scientific understanding using the four level scale outlined above. Table 3.4 illustrates how students' responses to question 15 on the posttest were mapped onto the standards and then coded for level of understanding.

The responses by Kelly and Kiley (Table 3.4) show text that can be mapped onto five different standards, while Mack's response barely touches upon one standard (C.5.2 - photosynthesis). Kelly's responses are more elaborate and more closely match the standards than Kiley's (see Expert's responses and standards in Figure 3.1 on page 90) and so were accorded higher ratings. As with the artifacts, all understandings were recorded in the data matrix.

Table 3.4: Selected Student responses to Question 15 on the posttest.

Student Responses with standard codes affixed: C.5.1 – entropy, C.5.2 – photosynthesis, C.5.5 – niches, F.2.1 - population growth, F.5.2 – human impacts	Standard codes	Understanding Levels
Through urban run-off, the phosphates would enter the creek. ^(F.5.2) Phosphates cause excessive plant growth ^(F.2.1) , which through a number of steps depleats the oxygen level. ^(C.5.1) Photosynthesis ($H_2O + CO_2 \rightarrow C_6H_{12}O_6 + O_2$) is the process which green plants use to make energy from the sun's light. When more plant are there to photosynthesize, During the day the DO levels increase ^(C.5.2) However, during the night when there is no sun to help produce more oxygen and cellular respiration ($C_6H_{12}O_6 + O_2 \rightarrow H_2O + CO_2 + ATP$) is occurring, the DO levels go down. ^(C.5.2) More plants start dying, and the BOD goes up, because there is so much more dead material in need of being decomposed. ^(C.5.2) So the oxygen levels go down and Taxa #1 organisms also begin dying, because they need higher oxygen levels ^(C.5.5) (Kelly B3-11/45F)	C.5.1 C.5.2 C.5.5 F.2.1 F.5.2	3 3 3 2 2
If extra phosphates enter the creek then cultural eutrophication occurs ^(F.5.2) More plant would grow, ^(F.2.1) during photosynthesis they produce more DO ^(C.5.2) which in turn would attract[sic, attract] more animals. ^(C.5.5) The animals use cellular repiration which produces CO_2 ^(C.5.2) Then the cycle would continue until the phosphates dissappear. When the phosphates dissappear the plants die ^(C.5.5) When the plants die, the bacteria break them down, using up the O_2 left. ^(C.5.1) The BOD goes up but with less O_2 the animals start to die too. ^(C.5.5) (Kiley B3-03/11F)	C.5.1 C.5.2 C.5.5 F.2.1 F.5.2	2 2 2 1 1
If xtra phosphates entered the water the the D.O. could go down and lots of stuff would change. ^(C.5.2) I don't really how xtra phosphates would effect all of that. (Mack B3-09)	C.5.1 C.5.2 C.5.5 F.2.1	0 1 0 0

Statistical Analysis of Student Understandings in Artifacts

Because students' understandings were evaluated on a four level ordinal scale, non-parametric statistics were used for data analysis. An individual student's understandings within a standard were computed by taking the median score for all items that applied to that standard within a given artifact. The student's overall understanding for the standard area was determined by using their median score as an indication of central tendency.

A second educational researcher provided inter-rater reliability on the artifact scoring. After instruction in the basics of stream related scientific understandings, an introduction to the scoring guide, and a few practice artifacts, the second researcher and I independently scored a subset (~10%) of the artifacts. Pearson product-moment correlation coefficients were calculated for agreement on each standard scored in a particular measure. Inter-rater reliability on conceptual understanding standards included a fairly high range ($0.837 \leq r \leq 0.958$) or ($.70 \leq r^2 \leq .92$). With a high inter-rater reliability indicated, the primary researcher proceeded to independently score the remaining artifacts.

Comparison of course content and artifact specifications to actual demonstrations of understanding across the artifacts

The number of occurrences of understanding for each standard for each data source was determined and the frequency of demonstrated understandings at each level (L3, L2, L1 and L0) was calculated for that data source. Analysis of the data matrices (Appendix I) revealed which constructs were demonstrated with understanding by the students and at what frequency through each of the different artifacts. This tool allowed an analysis of the range of understandings demonstrated in each of the different artifacts by looking at the whole set of student responses. Examination of the frequency patterns in the matrices was used to determine if the array of measures (artifacts) was sufficient to assess student understandings of the enacted curriculum and to determine if there were patterns of demonstrated understandings (e.g. some constructs are over represented and others are missing from the assessment pieces).

Analysis of Student Understandings across Time

The analysis of individual student understandings was accomplished by using the raw data of standards and levels of understanding described above in step three, a total of 7,687 demonstrations of understanding. Step five analysis was conducted at the individual level as students demonstrated understandings of each standard throughout the project. This analysis was used to answer the third research question: "What patterns of understandings were demonstrated by the students? How did their levels of demonstrated understanding change across the artifacts and the semester?" Part of this question was a concern that students would "specialize" or consistently display understandings around the same content across the different artifacts rather than demonstrating different understandings.

Since some of the measures were completed by individuals and others by groups of two to five students (see Table 3.3), a critical assumption had to be made in the analysis of individual understandings. This assumption, based on an ideal world, was that the demonstrated understandings in an artifact could be attributed to all authors of that artifact. For example, if a report provided evidence of a Level 2 understanding about gas solubility (B.2.5), all of the students whose names were on that report were recorded as demonstrating a Level 2 understanding for standard B.2.5. The check and balance on this assumption was the pre/post tests, which were completed by individuals. By employing this assumption, it was possible to do a student-by-student, measure-by-measure, standard-by-standard analysis of demonstrated understandings across the semester curriculum. For example, a student's understanding of species competition

(Standard C.4.3) could be tracked from pre-test to essay to report to models 1 and 2 and finally to the post-test.

The Sign-test was used for the analysis of student achievement across artifacts. This basic statistical test can be used with non-parametric data such as the ordinal scale used to denote levels of understandings in this study. The Sign-test was used with matched pairs of observations and tested the null hypothesis of no difference in the matched observations.

As its name suggests, the Sign-test is based on the signs of the response differences, that is, the difference between an understanding demonstrated on one measurement compared to a prior measurement. In a pre/post test measurement, if the treatment has no effect on developing that understanding, the response difference D_i in each pair is as likely to be positive as it is to be negative and the null hypothesis can be stated as $H_0: P[+] = .5 = P[-]$. However, if students are developing understanding, then the difference between a given measurement and an earlier measurement should be positive ($H_a: P[+] > .5$). The Sign-test measures the number of positive results, and large numbers of positive differences are expected to fall outside the range of the null hypothesis for a given confidence level (in this study $\alpha = .05$) (Bhattacharyya & Johnson, 1977).

MYSTAT (Wilkinson & Bjerknes, 1989) was used to compute the Sign-test on all pairs of variables, omitting zero values. It is important to note that the pairs of variables consisted of a measure of understanding and the previous time understanding on that standard was recorded. For example, on standard C.4.3 related to biological competition, "Jane" may have demonstrated this understanding in all six measures. "Dick" may

address it in only four of the six measures (e.g. pre and post tests, report and model 1). Moreover, "Sally's" understanding of this content may only be assessed on the pre and posttest. The Sign-test for Jane would compare her understanding of competition on the pre-test to the essay, from the essay to the report, from the report to model 1, from model 1 to model 2, and from model 2 to the post test while the sign-test for Sally would only compare pre and post test scores. Thus, the Sign-test allowed a determination of whether or not scientific understandings were increasing along a particular standard over the course of the semester by looking for signs of positive change within students within particular standards. The outcomes of the Sign-tests are reported as probabilities in the data tables in Chapter Four.

In addition to the Sign -tests used to compare student achievement across measures, the Wilcoxon non-parametric sign-rank test was used to compare students' pre and posttest achievement as a class ($\alpha = .05$).

Summary of Research Design

In this chapter, I have described the context of the school and curriculum in which this study is situated. I have also described each of the student artifacts and data sources which were subjected to content analysis and how that analysis was performed using the *National Science Education Standards* (NRC, 1996). The five steps of the analysis included: (1) The delineation of the curriculum and mapping it onto the *National Science Education Standards* (NRC, 1996), (2) The identification of opportunities (and expectations) to demonstrate understanding in each of the selected measures. (3) Analysis of student understandings in each of the measures. (4) Comparison of course content and

artifact specifications to actual demonstrations of understanding across the artifacts. And
(5) the examination of student understandings across time. In the next chapter, I report
the findings of those analyses.

CHAPTER 4

FINDINGS: CONTENT, STANDARDS, AND STUDENT UNDERSTANDINGS

This chapter, describing the match between the curriculum, student understandings, and the *National Science Education Standards (NRC 1996)*, is organized based on the steps of the research design described in Chapter 3. The mapping of the Traver Creek curriculum (step 1) and the content specified in the assessed artifacts (step 2) onto the *Standards* is discussed first. Then the specific understandings disclosed in the students' artifacts (step 3) are reported and deliberated. In the final section, the results of the fourth and fifth lines of analysis, comparing understandings demonstrated across artifacts and across time, are presented and discussed.

The Match between the Standards, Course Content and Artifact specifications

There are 126 high school science standards explicated in the *National Science Education Standards* (Chapter 6, NRC, 1996). These 126 standards include ninety (90) that can be classified as addressing conceptual understandings (Standards B, C, D & F, Appendix B). The Traver Creek Project in Foundations of Science-I is an interdisciplinary science curriculum that includes content from earth science, biology and chemistry. When mapped onto the *Standards*, the creek curriculum was found to address forty-five (45) conceptual understanding standards at the high school level, half of the content explicated by the *Standards* as being important for students to understand (Table

4.1 and Appendix B). These were distributed across the four major content standards of B-Physical Science, C-Life Science, D-Earth and Space Science and F- Science in Personal and Social Perspectives as summarized in Table 4.1. The creek curriculum also explicitly and implicitly addresses *Standards* about the nature of science and the processes of doing science.

Although the curriculum content is relatively balanced between earth science, biology, and chemistry (see description in Chapter 3), this balance is not evident when looking at the number of standards addressed. When mapped onto the *Standards*, the Traver Creek Curriculum appears heavily weighted toward the Life Sciences (19 in Standard C) and Environmental Sciences (11 in Standard F) (Table 4.1).

Table 4.1: The Traver Creek Project curriculum addresses National Science Education Standards (NRC, 1996) across the four science content areas of conceptual understanding.

Major NRC standards addressed by the Creek Project	Number of NSES in the project	Number on Pre/post Tests	Number assessed in artifacts
Conceptual Understandings	45	17	23
Standard B: Physical Science	11	4	2
Standard C: Life Science	19	8	10
Standard D: Earth and Space Science	4	2	4
Standard F: Science in Personal and Social Perspectives	11	3	7

The apparent imbalance in *Standards* is due, in part, to how well the content is delineated among the different substandards in each section. For example, in the Life Sciences, standard C.4 addresses the interdependence of organisms and there are five objectives that differentiate the various kinds of interactions (Appendix B). All five substandards match content addressed in the creek project during the three weeks that were focused on the creek as an ecosystem. In contrast, over three weeks of classroom time were also devoted to introductory chemistry (atoms, compounds, simple reactions,

etc.). Yet only two of the five sub-standards under B.3 Chemical Reactions applied to this project (Appendix B). Likewise in Standard D on Earth and Space Science, weather phenomena, which were addressed in the context that weather impacts a watershed through flooding, drought, seasonal changes, etc., were only tangentially addressed in three standards (D.1.1, D.2.1 and D.3.3)¹³ (Appendix B).

The curricular standards that were assessed in the four artifacts and the pre/post tests formed a subset of the standards addressed by the project. The standards addressed by the different measurement methods are examined next.

Standards assessed in the Artifacts.

Of the 45 conceptual standards addressed in the creek project, twenty-three (23) were assessed in the four major artifacts, e.g., essays, reports and models.

As noted in the example in Chapter 3 (page 72), the analysis of the essay assignment (Appendix C) showed that students had the opportunity to address ten conceptual standards (Table 4.2). Specifically, the essay assignment provided affordances to demonstrate conceptual understandings about the use of resources (F.3.1) and natural systems (F.3.3) related to land use and human impacts such as pollution on the watershed (C.4.5 and F.6.5). In addition, other conceptual scientific understandings could readily be expressed in the essays even though they were not expressly provided for by the task. An example of such an understanding would be if students included concepts of photosynthesis and respiration (C.5.2) in the description of how shading might affect.

¹³ Standards at the middle school level (grades 5-8) do address weather phenomena in more explicit forms (NRC, 1996)

Table 4.2: National Science Education Standards (NRC, 1996) assessed in the artifacts and pre/post test throughout the creek curriculum. Codes inside parenthesis e.g. (C.3.5) represent standards that are afforded by the measure but not explicitly required.

Major NRC standards addressed by the Creek Project	Essay	Report	Model 1	Model 2	Pre/Post
Conceptual Understandings	10	6	7	21	17
Standard B: Physical Science B.1.1 Atomic structure B.2.2 Structure and properties of matter B.2.5 Physical states B.3.1 Where reactions occur B.3.3 Red/ox and acid/base Rx's				(B.2.5) (B.3.3)	B.1.1 B.2.2 B.2.5 B.3.1
Standard C: Life Science C.3.5 Classification, taxa groups C.4.1 Biochemical cycles C.4.2 Food chains C.4.3 Interactions, competition C.4.4 Population potentials C.4.5 Human ecology C.5.1 Entropy C.5.2 Photo and molecular synthesis C.5.4 Adaptations C.5.5 Niche concept C.5.6 Energy flow and conservation	(C.3.5) C.4.5 (C.5.2)	C.3.5 C.4.2 C.5.5	 (C.4.2) (C.4.3) (C.4.5) C.5.5	(C.3.5) (C.4.1) (C.4.2) (C.4.3) (C.4.4) (C.4.5) (C.5.1) (C.5.2) (C.5.5) (C.5.6)	C.3.5 C.4.3 C.4.5 C.5.1 C.5.2 C.5.4 C.5.5 C.5.6
Standard D: Earth and Space Science D.1.1 Energy sources D.2.1 Chemical cycles D.2.2 Chemical cycles - carbon D.3.3 Interactions in earth systems	(D.1.1) (D.3.3)	 D.3.3	(D.1.1) (D.2.1) D.3.3	(D.1.1) (D.2.1) (D.2.2) (D.3.3)	 D.2.1 D.3.3
Standard F: Science in Personal and Social Perspectives F.1.3 Choice & environmental consequences F.2.1 Population growth patterns F.3.1 Human use of resources F.3.3 Use of natural systems F.5.2 Hazardous consequences F.5.3 Natural hazards F.6.5 Human effects on habitat	(F.1.3) F.3.1 F.3.3 (F.5.3) F.6.5	 F.5.2 (F.5.3)	 	(F.2.1) (F.3.1) (F.3.3) (F.5.2) (F.6.5)	F.2.1 F.5.2 F.5.3

the creek. Students could also write about interactions among earth systems (D.3.3) such as the effects of weather and erosion (F.5.3) on organisms and habitats in the creek. To a lesser extent they might convey understandings about the organization of the biological

organisms of the creek (C.3.5), the effect of sun and shading (D.1.1) and how their own personal actions might impact the creek or how the creek impressed them (F.1.3)

So of the ten conceptual standards identified, four standards (C.4.5, F.3.1, F.3.3, F.6.5) are explicitly expected by the essay assignment and six standards (C.3.5, C.5.2, D.1.1, D.3.3, F.1.3 and F.5.3) are implied by the assignment.

The scientific report assignment (Appendix D) provided explicit opportunities for students to represent understandings on five conceptual standards distributed among standards C, D and F (Table 4.2 and Appendix B). In the Life Sciences standards, opportunities were found that mapped onto biological classification (C.3.5), the interdependence of organisms/foodchains (C.4.2) and the niche concept (C.5.5) (See analysis in Chapter 3, p 75). In Earth and Space Sciences, opportunities were provided for understandings on the interactions of earth systems (D.3.3) and the concepts around “watershed.” Ideas of human land uses would map onto Standard F (F.5.2). Like the essay assignment, students were not restricted to the five content standards but could demonstrate additional understandings.

The first model building cycle (week 11) provided opportunities to address seven standards as students built models of the physical and biological factors in the study creek. The modeling assignment asked students to model a section of the creek to show how physical factors affected the benthic macro-invertebrates (Appendix E). A model that would evidence these conceptual understandings must include one or more physical factors and show a relationship between that factor and the benthics (object). The interactions of physical and biological components of the system map onto standards D.3.3 (earth systems) and C.5.5 (niche concept). If students included physical factors

such as the sun affecting the temperature of the water, their understandings might also map onto other earth science standards (i.e. D.1.1 and D.2.1). If they elaborated on the benthic macro-invertebrates, they might also represent other Life Science understandings such as those under *C.4 - The interdependence of organisms*. The conceptual understandings that were assessed in this assignment include seven standards in Standards C & D (Table 4.2 and Appendix B) although students needed to only address two to meet the requirements of the assignment (Appendix E).

For the second modeling cycle assignment, students were asked to build a model that demonstrated an “in-depth understanding” of a stream ecosystem and that included physical, chemical, and biological factors of the stream (Appendix F). This modeling assignment was much more open in terms of which conceptual understandings students might include and lacked the specificity found in some of the other assignments. As such, students might address a number of different conceptual understandings that would map onto the Physical Sciences (Standard B), Life Sciences (Standard C), Earth Sciences (Standard D), and Environmental Sciences (Standard F). However, there were a few standards, such as B.1.1 on the structure of atoms, B.2.1 on chemical reactions, and B.2.2 on the Periodic Table, which include content that could not be reasonably represented within the modeling environment. These standards are examples of those few that were not supported by this assignment. The other standards not supported by the second modeling assignment included C.5.4 (adaptations), C.6.2 (behavioral adaptations), F.2.3 (population carrying capacity), and F.5.4 (risk assessment) (Appendix B).

Although students were offered the opportunity to include content from 21 different standards in their models (Table 4.2), it was not expected that they would do so.

Rather, the assignment was designed for students to select content from the biological, physical, and chemical factors of stream phenomena, representing a subset of content standards in any one model.

Standards assessed in the Pre/Posttests

The pre/post test instruments addressed seventeen (17) conceptual understanding standards distributed across the four major content areas (Table 4.2). It was the only measurement to explicitly assess conceptual understandings of chemistry related to the structures of atoms (B.1.1), the Periodic Table (B.2.2), and chemical reactions (B.3.1). The instruments also assessed understandings about states of matter (B.2.5), a standard that was also assessed in the artifacts. In addition to these standards in the physical science, the instrument assessed eight standards in the life sciences, and three each in earth science and environmental science (Table 4.2).

Summary of the Assessed Standards

All together, the four major artifacts provided opportunities for students to demonstrate understandings on 23 conceptual standards while the pre/post instrument only assessed understandings on 17 standards (Table 4.2). In this analysis, the four artifacts offered the potential for a more complete assessment of understandings across the curriculum than the pre/post tests alone (51% vs. 38% of the curriculum). The four major artifacts do not provide support for students to demonstrate their conceptual understandings along several standards, including B.1.1 (structure of atoms), B.2.1 (interactions of atoms), B.2.2 (periodic table), C.6.2 (behavioral responses/adaptations), F.2.3 (population limits to growth), and F.5.4 (environmental risk assessment). Together,

the artifacts plus the pre/post instrument give a more complete assessment picture with evaluation on 27 of the 45 standards.

All four artifacts assessed only one common content standard (D.3.3 - Interactions of earth systems). Five standards (C.3.5, C.4.2, C.4.5, C.5.5, D.1.1) were assessed by three of the four artifacts and seven (C.4.3, D.2.1, F.3.1, F.3.3, F.5.2, F.5.3, F.6.5) were assessed by two of the four artifacts (Table 4.2 and Appendix B). The remaining fourteen standards assessed in the artifacts were supported by a single artifact, most often in the second modeling assignment.

From the distribution of standards across the artifacts, there does not appear to be an over-representation of specific content across the different artifacts. (Standard D.3.3 on the interactions of earth systems incorporates a broad array of potential conceptual understandings). There does appear to be an under-representation of project content in the artifacts, especially representation of content in Standards B (Physical Sciences) and F (Science in Personal and Social Perspectives).

The content in the Physical Sciences (Standard B) represents several weeks of classroom instruction (November 13- December 6) so the lack of artifacts that would assess students' conceptual understandings in this area is a potential weakness of this approach. However, this type of analysis identifies this weakness in assessment and steps can be taken to correct the deficit. An artifact that required students to represent the chemical reactions in one of the chemistry tests and the land uses that would impact that test would address this gap in the assessed standards. Students did complete some smaller artifacts including quizzes and mini-lab write-ups that provided the classroom

teachers with some assessment of students' chemical knowledge. However, these smaller artifacts were not selected for the in-depth analysis required by this study.

In contrast, the other three standards not assessed by the major artifacts represent minor content in the creek project. For instance, content along C.6.2 (behavioral responses/adaptations) was only address briefly in two class periods (September 20 & 30). Environmental risk assessment (F.5.4) was addressed in assigned readings and in films but was not emphasized by either teacher.

Demonstrations of Student Understandings within Artifacts

In this section, I examine the evidence of scientific understandings that students demonstrated in their artifacts found during step three of the research design. Examples of demonstrated understandings taken from students' artifacts are used to illustrate the analysis. Following the same outline as the analysis of standards assessed in the artifacts, demonstrations of scientific understandings will first be examined in the four artifacts: essays, reports, model 1 and model 2; and then in the pre/post test instrument.

Essays

I think that Traver Creek was a dirty creek. This is because of the fact that there was no plant life directly in the water and it was filled with trash. I would not want to swim in this creek because I think it might be dangerous. (Essay by Abner, Block 1).

The descriptive essays that the students wrote during the second week of the project were one to two pages long. In these essays, students tended to list a number of unconnected facts and observations of the creek with varying degrees of accuracy. Many of the essays read like disjointed notes rather than a descriptive essay designed to inform

future FOS students about the environments of the creek as specified by the assignment (Appendix C).

Forty-five essays ($n_a=45$; $n_s=45$)¹⁴ were analyzed for scientific understandings. Conceptual understandings were found in those statements that connected different observations, provided explanations, interpreted observations in light of other knowledge, or hypothesized about possible causes and effects.

The most frequent conceptual understanding expressed in these essays dealt with the classification of organisms (C.3.5), with twelve relevant statements (Figure 4.1). An example of this type of understanding was found in Casie's essay:

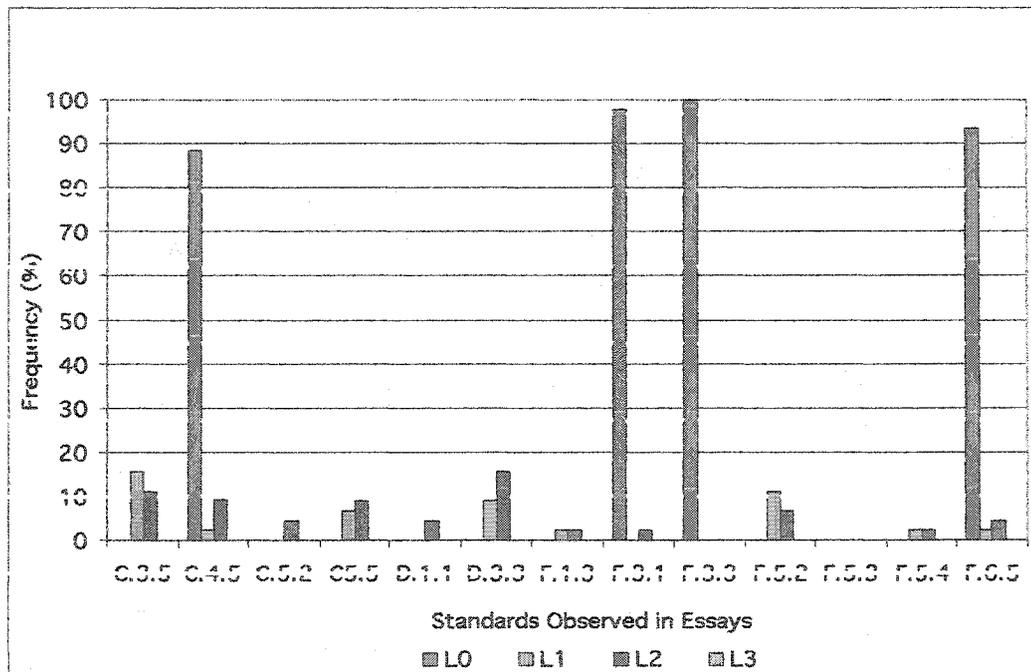
There weren't very many living things visible from where I was sitting. I saw the algae, and also some bugs. The bugs were mainly water striders, with some little gnats hanging around too. Once I looked on my shoulder and saw an odd looking green bug perched there. It had a praying mantis style body, only it wasn't praying and it had the largest antennae I've ever seen on a bug. It also had big red eyes that seemed to stare at me. It was kind of kooky looking.

I personally didn't see anything larger, but if I was to make a guess, I'd say that there were probably fish in the water, and one of the other students said they saw crayfish, which sounds pretty realistic to me. There was also probably a lot of living bacteria in the water, but nobody can see that anyway. (Casie - Block 3; C.3.5 - Level 2)

In this extract, Casie identifies several major divisions of organisms in the environs of the creek including algae, bugs (insects), fish, crayfish (Crustaceans), and bacteria. However, the level of classification is not very sophisticated so that while water striders are true bugs (HEMIPTERA: Gerridae), the gnats (DIPTERA) and the other "bug", possibly a katydid (ORTHOPTERA: Tettigoniidae) would be more scientifically

¹⁴ n_a = Number of artifacts, n_s = Number of students. See Table 3.3

classified as insects. Thus, the level of understanding for standard C.3.5 in this essay was coded at Level 2 (developing literacy).



C.3.5 – Classification, C.4.5 – Human ecology, C.5.2 – Photosynthesis, C.5.5 – Niches, D.1.1 Energy sources, D.3.3 – interactions of earth systems, F.1.3 human choice, F.3.1 use of natural resources, F.3.3 use of natural systems, F.5.2 hazardous consequences, F.5.3 – natural hazards, F.6.5 Human effects on habitat.

Figure 4.1: Conceptual Understandings Demonstrated in Student Essays. Level 0 (L0 = red) results are included only for expected standards (See Table 4.2).

Other conceptual understandings demonstrated in the life sciences were understandings about human impacts on the creek environment (C.4.5) found in five essays, and understandings about how the physical environment impacts the distribution of organisms (C.5.5) in seven essays. In the case of C.4.5, a student’s example statement would be:

*Because of the bank’s sharp angle leading to the water, pollutants such as waste from picnics in the park and waste products from the surrounding buildings all easily wash into the creek. ...
The area is not as dirty as I thought it would be but it needs some improvement. Smoke from buildings and pollutants from the cars on the roads also contribute to the condition of the creek. I didn’t think the*

Water was as dirty as the surrounding lands. I founds lots of trash laying around. (Joel, Block 3- C.4.5 Level 2)

In this extract, while Joel makes an explicit link between specific human activities (e.g. picnics) and pollution entering the creek, the negative impact of these activities on the creek are implied. Since there was no further description, the above statements were coded Level 2.

For standard C.5.5, dealing with the concept of environmental limits and the niche concept, an example of one of the higher demonstrations of conceptual understanding was found in Leia's essay:

The water is very clean and there not a lot of algae or living organisms because of loss of oxygen. The water is not as cool as it should be, which is probably also a factor. While watching the river for an hour the only living thing in the water I saw was a crayfish (with its left clipper broken off). (Leia, Block 3 - C.5.5 Level 2.)

In this case, the student was attributing the lack of visible [animal] life in the creek to the lack of oxygen and the temperature of the water, demonstrating a developing or Level 2 understanding. She did not explicitly link together the warm water temperatures and low oxygen levels in a relationship so understandings about gas solubility (B.2.5) were not inferred.

The second most frequent conceptual understanding demonstrated in the essays had to do with interactions between earth systems (standard D.3.3). Various statements that mapped onto this standard occurred in 11 essays. In one such instance, a student wrote:

The creek itself doesn't have much current or make very much noise, although there are some pretty large rocks and gravel under the water. It

might be true that the reason why there's so little movement is because the creek is only one, to one and a half-feet deep. But the lack of movement could also be causing the water to be warmer, which leads to reduced oxygen. Reduced oxygen can make it more difficult for some organisms to survive. It can also affect the metabolism of fish, insects, and amphibians making it harder for them to survive under warmer conditions. Some organisms may adapt to the higher temperatures and become dependent on it. For many insects the higher temperatures causes them to finish their metamorphosis earlier, which affects the predators who depend on certain insects at a particular time of the year. The temperature is very critical to all the plants and animals and other organisms around the creek. (Anita, Block 3 - B25-2, C55-2, D11-2, D33-2, F65 - 2)

In this extract from Anita's essay, she makes a number of connections demonstrating rich conceptual understandings along several standards. For standard D.3.3 (interactions of earth systems), she includes the chain from the shallow creek and low current to warmer water to reduced oxygen to fewer organisms, demonstrating an understanding of how the abiotic and biotic environments may interact. In this same quote, it is also possible to identify conceptual understandings that map unto the niche concept, standard C.5.5 (the niche concept, e.g. "Reduced oxygen can make it more difficult for some organisms to survive." and the temperature effects). Another applicable standard is B.2.5, interpreted in this project as including the concepts of gas solubility. In Anita's essay, understanding of these relationships is demonstrated when she writes, "...causing the water to be warmer, which leads to reduced oxygen." In this extract, she has the qualitative nature of the relationship right although she doesn't elaborate on the causality - e.g. because oxygen or other gases are less soluble at higher temperatures - so her understanding, using the conservative rule, was coded at Level 2 for Standard B.2.5.

A fifth conceptual understanding, occurring in seven (15%) of the essays, mapped onto standard F.5.2, which deals with the impact of human activity on natural change. In

the extract below, Jack lists a number of human artifacts present in the water that he identifies as stabilizing the environment.

The creek itself is very small, and is a good example of how civilization has almost taken over nature. ... There are quite a few things in the water. These include cement blocks, drain pipes, and a piece of rubber. This human influence here, rather than polluting, is actually helping to stabilize the environment. (Jack, Block 1- F.5.2 - Level 2)

Jack has made a reasonable conclusion in this specific creek environment. The blocks and bricks in the creek work as substrate for colonization by benthic macro-invertebrates. In addition, some of the drainpipes that have been eroded out of the bank, function as an artificial log or berm, diverting the current and forming a pool, creating additional stream habitat.

Even though students were given class time for peer critiques and revision of their essays, most of the students followed the format of the note taking sheet in completing the essay assignment. In fact, apart from some of those that were excerpted above, many of the essays read as a transcription of the note taking sheet rather than a descriptive essay designed to inform future FOS 1 students about the environments of Traver Creek in 1996. Thus, while the essays revealed a little bit about students prior knowledge of stream ecology, they provided little evidence of scientific understandings.

There are different explanations of why students exhibited so little identifiable understandings in their essays, even at the lower levels of L1 and L2. One reason may simply be the writing abilities of these ninth grade students. However, more intriguing in the context of science education, may be the interaction between conceptual understandings and the dimensions of epistemological and strategic understandings. Although, these other two dimensions were not a focus of this study, their presence or

absence was noted in the initial analysis of the student artifacts. The assignment directions explicitly called for thesis statements. Yet, over one third (38%) of the students did not include an identifiable thesis statement. Only a few (15%) of the essays were interpreted as students having a more “constructed way of knowing” in their understanding of the purpose of their visit to Traver Creek and the resulting essay. An example can be found in Leia’s essay, where she situated the purpose of the activity in the larger semester project. There was also a tendency for students to relate what they “know” rather than what they “observed.” E.g. they “know” that parking lots may have negative influence on the creek so the Kroger Shopping Center has a negative impact on the study site in Plymouth Park even though it is several hundred meters downstream. The interplay of the other dimensions of scientific understandings with expressions of conceptual understandings requires further study.

The analysis of the essay assignment would also suggest the need for clarification of teacher expectations of an essay in contrast to simple notes and perhaps some additional scaffolds. For example, the whole class could have critiqued an example essay and provided feedback to its author before doing the individual peer critiques. They could also have read some essays from previous years and discussed which information included in the essays was useful for their own study of the creek and what information was lacking that they would want to know. Such an analysis could help them improve their own writing. In addition, high quality essays from previous years could be held up as models for the students in this FOS I class.

In summary, 33 of the 45 essays included statements that mapped onto the conceptual understanding standards. Fifty-eight statements, mapping unto thirteen

standards, were identified. Twenty-seven statements (47%) were coded at Level 1 (nonscientific) and 31 (53%) at Level 2 (developing literacy) understandings; none were coded at level 3 (attaining the standard) (Figure 4.1). Students did not act upon or demonstrate understandings on three standards (C.4.5, F.3.3, F.5.3,) expected by the assignment. Each of these missing standards has to do with the interaction between humans and the environment. A few students demonstrated understandings on four standards (C.5.2, C.5.5, F.5.2, and F.5.4) not explicitly specified in the assignment (Table 4.2, Figure 4.1). Twelve essays (26%) did not provide any evidence of conceptual understandings along the national standards. The average number of conceptual understanding standards addressed in the essays was 1.29.

Stream Water Quality Reports

Midway through the semester, after the students had completed the physical and biological assessments of the creek and the chemical testing, they wrote their reports. In the reports ($n_a=23$, $n_s=84$)¹⁵, the students provided evidence that they had knowledge of the stream and reported considerable data, but there was little sense making of the data. A total of 175 statements were classified as conceptual understandings, which mapped onto nineteen standards (average of 8 conceptual understanding standards per report).

Most of the conceptual understandings were expressed in the introduction to the report. For example, in the introduction written by Ezra, Leah, Gene, and Kiley, Block 3 (Figure

¹⁵ n_a = Number of artifacts, n_s = Number of students. See Table 3.3

4.2), there were eleven conceptual understanding standards identified of the thirteen total in the report.

The Relationship Between the Water Quality and the Benthic Macro Invertebrates of Traver Creek
Section 420-450, October 30, 1996

It is commonly accepted that Benthic Macro Invertebrates are a reliable way of testing water quality (Benthic, meaning bottom dwelling, macro, meaning microscopic or visible to the naked eye and invertebrate, meaning without a backbone.) Therefore, one way they are helpful to us is to tell us about the overall health and condition of a body of water. By looking at the sensitivity and the number of a certain organism in a given stream river, or lake, one should be able to determine the pollution level.

The Benthic community is affected by a number of physical forces: pollution, nutrition levels and run-off. Pollution by toxic chemicals is common in many bodies of water. It kills, maims and sickens animal and plant life. Nutrient load can affect the food chain in many ways: in addition to natural nutrients supplied by soil and vegetation, man-made fertilizers can cause water plants to bloom drastically and this can lead to changes in the oxygen level of the water. Many intolerant Benthics are not accustomed to that. Run-off from surfaces surrounding within the watershed of the creek lead to different flow rates. After a heavy rain in an area with a large amount of run-off, the bodies of water swell, and their flow rate increases. This can cause some Macro Invertebrates to be swept away in the current. Only a few are adapted to living under the conditions of fast moving water. Light levels also affect the Benthic Macro Invertebrates. Cutting down trees along the bank of a stream can affect the amount of light reaching the water, affecting the growth of the plants and algae and thus the whole food chain. Fifth, temperature. In some places thermal pollution caused by power plants can be a problem. Any rise in temperature might result in a change in life cycle, or the conditions might be too hot for some Macro Invertebrates.

Our study was based on our study of the Macro Invertebrates we found, as well as a physical assessment of the creek

Ezra, Leah, Gene, and Kiley (elok) Introduction to report – Ms L Block 3.

Figure 4.2 Sample of student work - Introduction from scientific report written by Ezra, Leah, Gene and Kiley.

Standard C.5.5 (niche concept) was noted in the group's writing about the different conditions affecting the sensitivity and number of certain organisms (L3).

Standard C.5.2 (photosynthesis & growth), at L2, was identified in the references to

“man-made fertilizers can cause water plants to bloom drastically and this can lead to changes in the oxygen level of the water.” And D.3.3 (interactions of earth systems)- L2 was identified in the references to water runoff, light levels and shading by vegetation, and their effects on the benthic macro-invertebrates. Some of these same standards were also demonstrated in the discussion section, for example standard C.5.5 - where the students discussed interactions of the habitat and their findings of certain benthic macro-invertebrates. In general, the students’ discussion sections were much weaker in making connections and demonstrating understandings than their introductions.

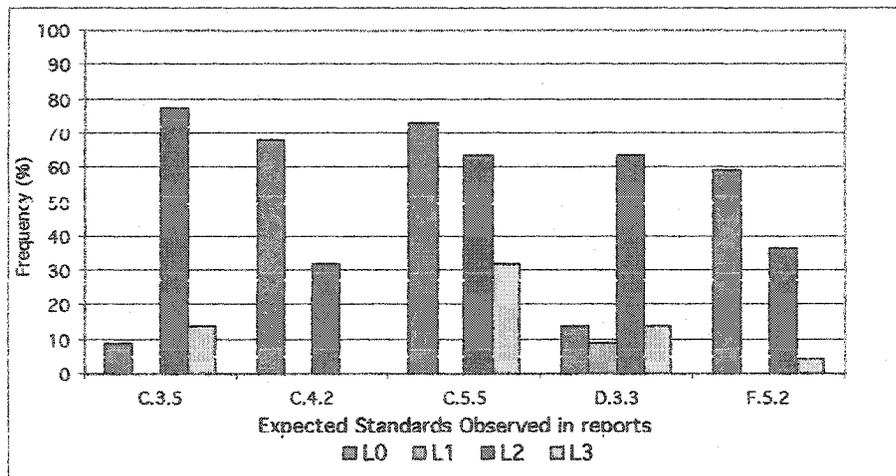


Figure 4.3: Expected conceptual understandings demonstrated in water quality reports ($n_a=23$, $n_s=84$). C.3.5 –classification, C.4.2 – food chains, C.5.5 – niches, D.3.3 – interactions of earth systems, F.5.2 - human caused hazards.

Twenty-one reports (95%) addressed standard C.5.5 (the niche concept and environmental limits of organisms). Seven reports (32%) addressed this standard at Level 3 understanding while most were classified as Level 2 understandings (Figure 4.3).

An example of a Level 3 demonstrated understanding was found in Chase's introduction to the report¹⁶,

*One physical factor that affects the existence of BMIs [Benthic Macro-Invertebrates] in the creek is discharge. Discharge is how much water flows that flows past a given point at a given time. BMI's use water flow to move wastes downstream and to gather nutrients from upstream. If the flow is slow, then there is less food, which will diminish the population or kill them off. IF the waste cannot be discharged downstream, the water will become polluted not only by their own waste but by human waste which would become trapped by the slow moving water...
Chase Block 3, Report page 1. - C.5.5 L3.*

Here Chase accurately identified two limiting factors: food availability and waste accumulation, so his understanding of environmental tolerances was coded at Level 3.

The second most common conceptual understanding was the identification and classification of organisms (C.3.5). Twenty reports (91%) included some demonstration of this understanding - most at Level 2 - a general classification of the benthic macro-invertebrates and/or vegetation, usually with common names. This was the only other Life Science standard in which a few students expressed a Level 3 understanding. One example of Level 3 understandings is from the report by Doris, Kayla, Esteban, and Richard, where they wrote:

We found four types of benthic macroinvertebrates (eight benthics total) in our section of the creek. They were: Leech, Riffle Beetle, Sow Bug, and Caddisfly. Leeches are predators and have a taxa # of 3. This means they

¹⁶ One group (Lilith, Taylor, Chase and Riku) in Block 3 apparently had trouble working together. This group submitted separate Introduction and Discussion sections but common Methods and Result sections. This text passage is from the introduction by Chase. For this report, the understandings demonstrated in the individual sections were attributed only to that author; understandings in the methods and results sections were attributed to all four students.

are tolerant to many pollutants. They are in the group Hirudinea. The Riffle Beetle has a Taxa # of 1 as does the Caddisfly. The Riffle Beetle is part of the group Hydrophilidae. They are collectors. They feed on decomposing organic matter. The Caddisfly is of the group tricopter [sic]. It is a Filterer-Collector. It strains minute particles out of the water. The Sow Bug has a Taxa # of 2. The Sow Bug is an isopod and a collector.

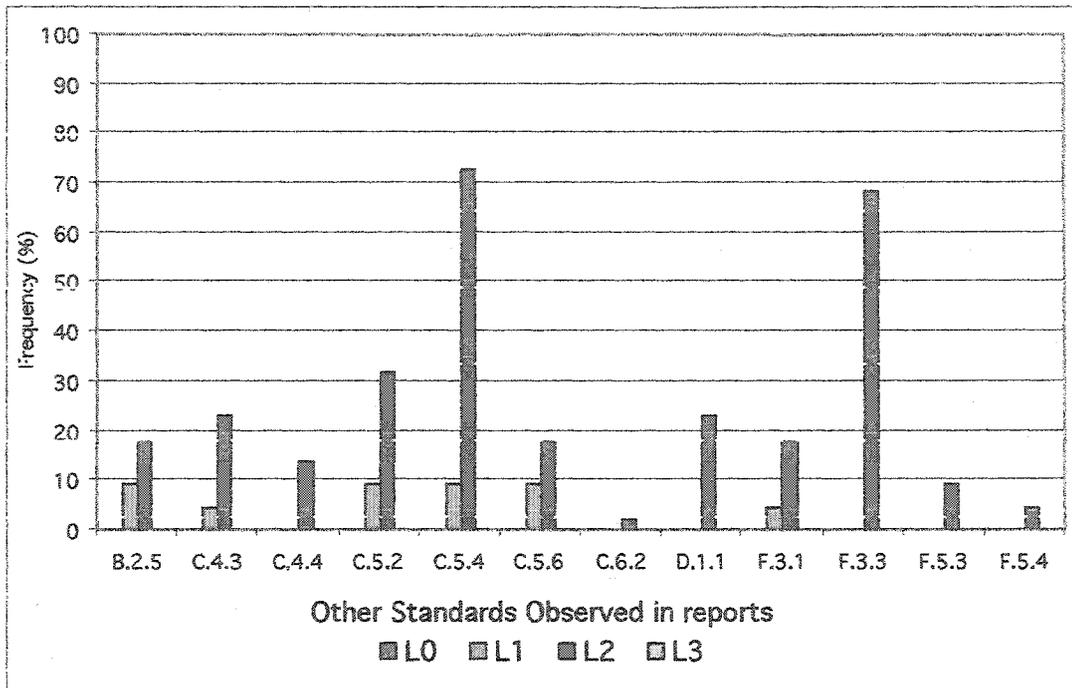
(Doris, Kayla, Esteban, Richard, Block 5, Report, p 4.)

In this passage, students identify their benthic macro-invertebrates by both common and scientific names. (The riffle beetle, if the common name is correct, belongs in the family Elmidae, not Hydrophilidae). They have also identified the functional group of each benthic, demonstrating an understanding along standard C.5.4, and the pollution tolerance index (PTI) taxon number, a demonstration of understanding along standard C.5.5. This group of students provided classification information than most. Most reports did not use scientific names.

Understandings about watersheds and understandings of interactions among earth systems (D.3.3) were addressed by the majority of the reports, with all but three reports demonstrating some evidence of understanding of these concepts (Figure 4.3). One example is from Chase's report in Ms. L's block 3:

Water temperature also is an important factor that affects the benthic population in the water. Temperature is how cold and warm the stream is. Temperature affects the life cycles of benthics. At the top [headwaters] of a stream the temperatures are much more predictable because there is less runoff than in an urban creek like Traver. There are factors which affect the temperature of the creek. The top [headwaters] stays constant but the bottom [lower reaches] of the stream varies depending on how much thermal runoff there is from other sources like power plants, houses, and shady/open spots. ...

(Chase, Block 3, Report, page 2. D.1.1 L2, D.3.3 L2)



B.2.5 – gas solubility, C.4.3 – competition, C.4.4 - population potentials, C.5.2 – photosynthesis, C.5.4 – adaptations, C.5.6 – energy in ecosystems, C.6.2 – behavioral adaptations, D.1.1 sun energy, F.3.1 human choices, F.3.3 – use of natural systems, F.5.3 – natural hazards, F.5.4 – human impacts.

Figure 4.4: Additional conceptual understandings demonstrated in water quality reports ($n_a=23$, $n_s=84$).

In addition to the assessed standards, some students also demonstrated understandings that mapped onto other standards (Figure 4.4). The most frequent one of these (found in 18 reports, 82%) was standard C.5.4, dealing with functional adaptations, usually in terms of functional feeding groups (e.g., grazers, predators, collectors), as illustrated by the passage from Doris, Kayla, Esteban, Richard given above. Standard F.3.3 - involving understandings around the use of human resources, including the use of indicator species - was demonstrated in over two-thirds of the reports. Other standards, which appeared in fewer than half of the reports, include, in order of decreasing frequency, F.3.1, C.5.6, C.4.4, and B.2.5 (Figure 4.4).

In general, students tended to report their data but made very little effort to interpret it or to connect one piece of data to another. In other words, like their essays,

they exhibited quite a bit of knowledge about their stream sections but very little understanding (connected knowing). The reports provided thin evidence of conceptual understandings, at least with enough depth or elaboration to determine whether the student had a Level 3 (scientifically literate) understanding versus a partially correct (Level 2) understanding.

For the five conceptual standards expected by the assignment (C.3.5, C.4.2, C.5.5, D.3.3, and F.5.2; Table 4.2), understanding was observed in one or more reports as well as understandings about watersheds in general (Figure 4.3). There were also a few isolated demonstrations of conceptual understanding on an additional twelve standards by the class as a whole (Figure 4.4). However, in individual artifacts, there were only a few understandings demonstrated. These demonstrations of conceptual understandings were generally expressed at a Level 2 (developing literacy) (Figures 4.3 & 4.4). Eight reports (36%) demonstrated a Level 3 understanding on at least one standard, but no report demonstrated an overall (or median) Level 3 conceptual understanding.

An investigation report has the potential to facilitate the negotiation of meaning and the construction of conceptual understanding (Keys, 1994). Considering the reports' potential for synthesis of the findings from the physical and biological assessments, the number and quality of conceptual understandings actually demonstrated was disappointing. There are a number of different hypotheses that might explain this weak demonstration of understandings. Two result from interactions with other frames of understanding.

First, perhaps students simply do not know how to write scientific reports. In this case, their lack of strategic understandings interferes with their ability to demonstrate

conceptual understandings. Good writing begins with a careful definition and interpretation of the problem, defining the problem first in terms of aim or purpose, and of audience (Maimon, et al., 1981). In the report assignments, students were expected to supply their thesis or problem statements and also address their reports to members of the local watershed council. If this hypothesis of interactions between strategic and conceptual understandings was supported by evidence of weak strategic understandings, then implications for instruction include more modeling and instruction on report writing.

A second explanation arises from a potential interaction of epistemology and conceptual understanding. This interaction may arise if reports are perceived as a “school” activity rather than a “science” activity, which is partly a function of audience (teacher vs. community). In this case, a visit by a representative of the watershed council and a presentation of the student’s findings to a member of the watershed council might help students perceive the greater purpose of their reports.

Creek Models - Cycle 1: Learning how to model ecological phenomena

In the first round of modeling in early November, students both learned how to use the Model-It software and then created a model of stream phenomena of their choice, with the provision that the model had to include physical and biological factors. Forty-eight models were collected off the class server: 12 built by individual students, 16 by male pairs, 12 by female pairs, and 8 by mixed pairs ($n_a=48$, $n_g=84$).

In these models, students created an average of 4 objects (range 1-15), which are the “physical” entities in the system, such as the stream itself, 5.4 factors (range 0-13) the measurable attributes of the object such as dissolved oxygen, temperature and substrate size and 5.6 relationships between factors (range 0-19) (Table 4.3).

The low factor to object ratio of 2.05 (Table 4.3) indicates that many students created a large number of objects (up to 15 in two models) and then defined only a single factor for each object, or no factors at all for some objects.

Table 4.3: General features of student models after model cycle 1 ($n_a=48$; $n_s=84$).

Model Attribute	n_a	Range	Mean	Stdev
Plan	39			
Number of Objects (O) in a Model	48	1-15	4.04	3.34
Number of Factors (F) in a Model	48	0-13	5.44	2.47
Number of Relationships (R) in a Model	48	0-19	5.56	4.02
Factor : Object ratio (F/O)	48	0-6	2.05	1.51
Relationship Integration (R/F)	47	0-3	0.97	0.52
Shortest Factor Path	47	0-6	2.43	1.31
Longest Factor Path	47	0-7	3.47	1.60

The largest fraction of the defined factors dealt with factors of the physical environment of the creek such as substrate (25 models), discharge (19 models) and bank vegetation (14 models). Twenty-four percent of the factors were defined around the stream organisms, primarily the numbers of benthic macro-invertebrates. Sixteen percent related to the Water Quality Index (WQI) and the nine parameters defined in the students' textbook (Mitchell, & Stapp, 1994). Of these, twenty-three cases related to stream temperature and twelve to dissolved oxygen. Other classes of factors defined in the students' models include biological processes, weather, human activity, and pollution.

Students connected their factors together by building relationships between factors. Most of these chains of factors and relationships consisted of three factors connected by two relationships. In nine models, students constructed factor paths of at least five factors and four relationships. The most common type of relationship was between two physical assessment factors (23%) followed by a physical factor affecting a benthic factor (18%) and a physical factor affecting one of the nine parameters of the WQI (9%). For instance, Casie, Lilith, and Taylor (Figure 4.5) built a model that showed

one physical factor affecting another (as discharge increases, substrate increases about the same) and of two physical factors affecting a biological factor (as substrate increases, benthics increase about the same, and as temperature increases benthics decrease about the same). This model also has one biological factor (algae) affecting another (benthics), and organic matter affecting benthics (Figure 4.5).

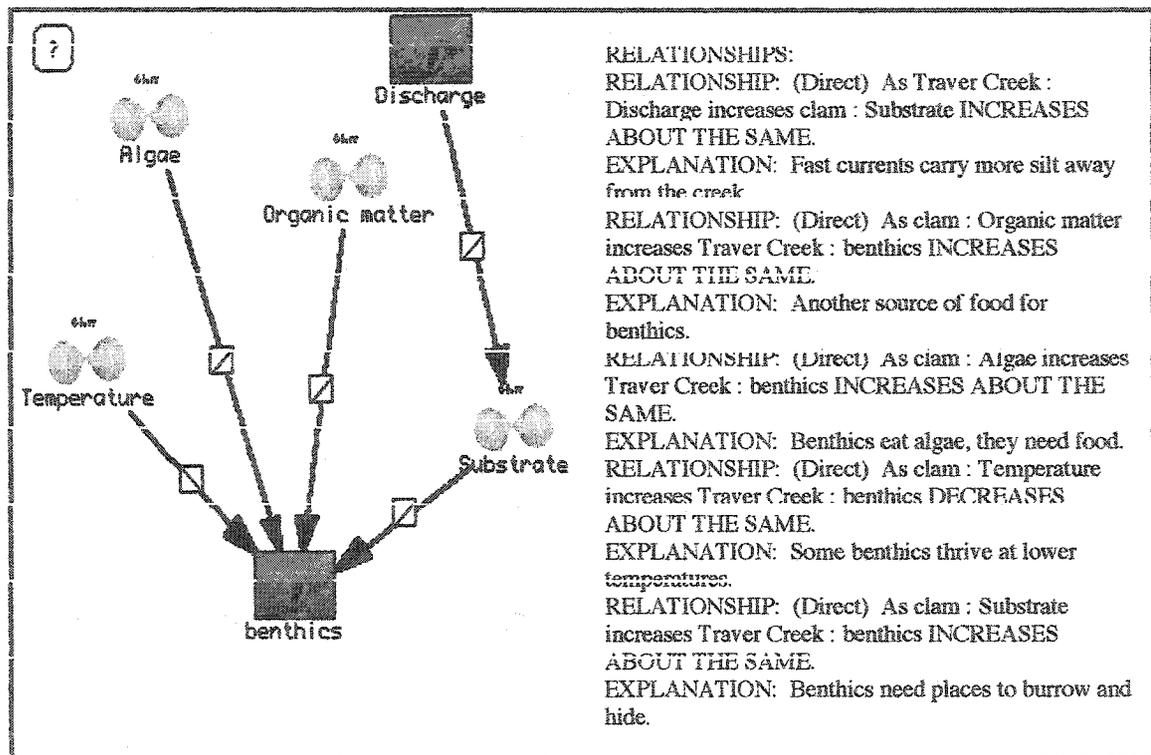


Figure 4.5: Factor and Relationship Map from Casie, Lilith, and Taylor (Block 3, model 1) showing factors and constructed relationships. In this model, the students have created two objects - Traver creek and clams, six factors, and five relationships. The longest factor path, from discharge to benthics, is 3.

In building their factors and relationships, students demonstrated conceptual understandings that mapped onto sixteen standards, six assessed standards (Figure 4.6a) and ten standards not explicitly assessed by the assignment (Figure 4.6b). Individual models addressed an average of 2.5 conceptual standards with the most common

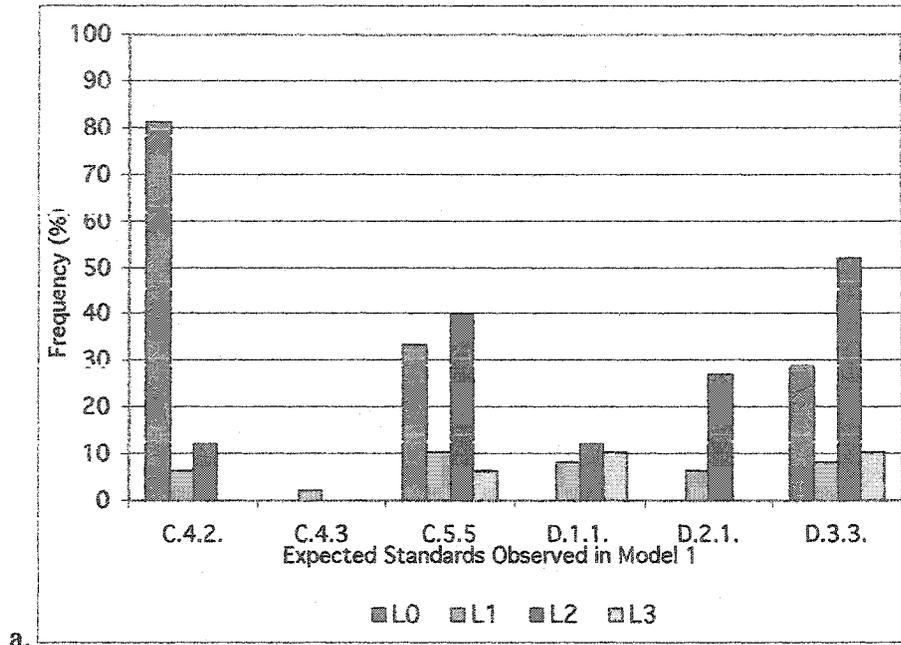
standards being Environmental tolerances and the niche concept (C.5.5), addressed in 32 models and Interactions of Earth Systems (D.3.3) addressed in 34 models (Figure 4.6a).

There were eight models (17%) where students' median level of conceptual understanding across all content included in their model was demonstrated at Level 3 - scientifically literate at the level of the *Standards*. Another 69% demonstrated a Level 2 overall conceptual understanding in their first model.

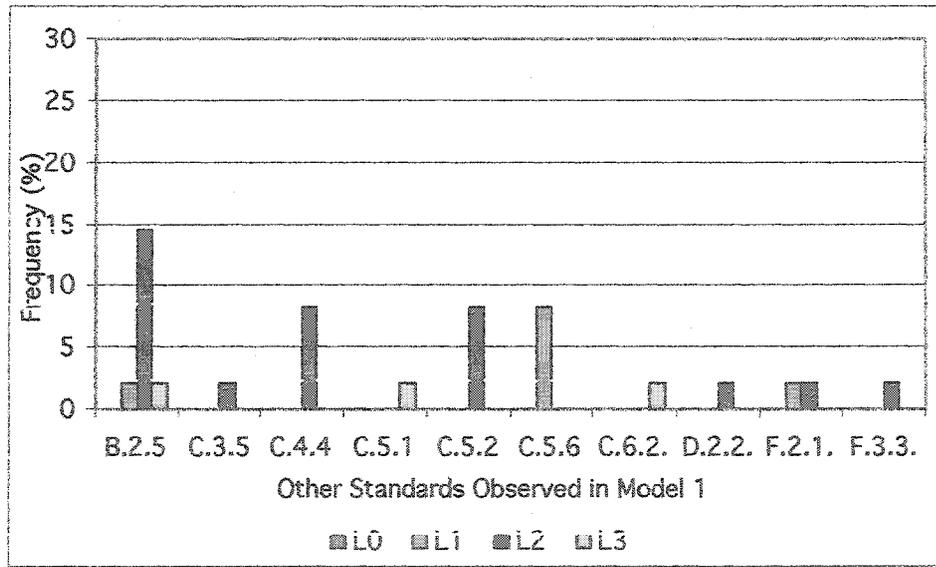
Examining these demonstrated conceptual understandings more closely by major content strands; it was found that students generally demonstrated more competent understandings of the Earth Sciences (Standard D) with three-quarters of the students scoring at Level 2 or higher (Figure 4.6). The Life Sciences (C), Physical Sciences (B) and Environmental Sciences (F) respectively followed earth science.

Looking at one content area at a time, it was found that students only addressed one standard in the Physical Sciences (B.2.5 - states of matter) (Figure 4.6b). This standard was addressed in nine models in the context of oxygen solubility in water being temperature dependent.

In the Life Sciences (Standard C), students' understandings mapped onto nine standards distributed across 4 sub-content areas (Figure 4.6). The most common demonstrated understanding mapped onto standard C.5.5, which involved understandings about the environmental tolerances, requirements and limits of organisms. Along this standard, half of the models demonstrated a Level 2 understanding and another three achieved a Level 3 understanding (Figure 4.6a). The second most common standard in the life sciences was C.4.2 (energy flow between organisms). This standard was addressed in nine models (19%) at levels 2 and 1. The other seven standards



a. C.4.2 – food chains, C.4.3 – competition, C.5.5 – niches, D.1.1 sun energy, D.2.1- chemical cycles, D.3.3 – interactions in earth systems.



b. B.2.5 – gas solubility C.3.5 –classification, C.4.4 - population potentials, C.5.2 – photosynthesis, C.5.6 – energy in ecosystems, C.6.2 – behavioral adaptations, D.2.2 – carbon cycle, F.2.1 – population growth, F.3.3 – use of natural systems,

Figure 4.6: Conceptual understandings demonstrated in Model 1. ($n_2=48$, $n_5=84$). Chart a shows those standards expected in Model 1. Chart b shows additional content that was observed in the models.

(C.3.5, C.4.3, C.4.4, C.5.1, C.5.2, C.5.6, and C.6.2) involved content addressed by fewer than 4 models (Figure 4.6b) although students tended to demonstrate level 2 or 3 understandings of these concepts when they included them in their models. With the exception of C.4.3 (organism interrelationships), these standards were not explicitly provided for by the assignment.

In the Earth Sciences (Content Standard D), the demonstrated conceptual understandings in the models mapped onto three standards (Figure 4.6a). The most common understanding in this content area relates to standard D.3.3 - Interactions among earth systems. This is a very broad standard and 34 models included conceptual content that mapped within this standard. For instance, Keshia and Glynis created the following relationship in their model (Figure 4.7):

*RELATIONSHIP: (Direct) As 540-570 : substrate increases Caddis fly Larva : quantity INCREASES ABOUT THE SAME.
EXPLANATION: When silt replaces cobble, rocks, and boulders, the caddis fly larva loses places in which it can hide from predators and take refuge from too fast water flow
(Keshia & Glynis, Block 3, Model 1. D.3.3 - L3. Figure 4.7)*

In this relationship, the two students have illustrated a relationship between the physical environment (stream substrate) and biological organisms (caddisfly larva). Their explanation for this relationship was based on their understanding of caddisfly habitat in the rocky or cobble part of the stream. On the strength of this relationship, the students' understanding of standard D.3.3 was rated at Level 3. In the next relationship they constructed in their model, Keshia and Glynis created a similar relationship between stream substrate and leeches. In this relationship they wrote:

RELATIONSHIP: (Direct) As 540-570 : substrate increases Leech : quantity DECREASES ABOUT THE SAME.

EXPLANATION: Leeches live better in muddy situations. (Keshia & Glynis, Block 3, Model 1. C.5.5 - L2, Figure 4.7.)

The explanation for this relationship was not well elaborated, without any indication of why the students believe leeches to live better in muddy situations. These two relationships, as well as the relationships between the other creek factors and the benthics (Figure 4.7) were used to classify the students' understandings on standard C.5.5 - the niche concept, at Level 2 (the median value).

The second most common understanding in the earth sciences involved the thermal effects of the sun on the stream, found in 15 models and mapping onto standard D.1.1 - Sources of energy. Keshia and Glynis's model (Figure 4.7) also demonstrates this understanding in the relationship between stream discharge patterns and water temperature. In their notepad, they wrote:

RELATIONSHIP: (Direct)As 540-570 : discharge patterns increases 540-570 : water temperature DECREASES ABOUT THE SAME. Description: As the water moves faster, the water becomes colder because the water isn't staying in one place and the sun's rays cannot constantly reach it. (Keshia & Glynis, Block 3, Model 1. D.1.1 - L2 Figure 4.7.)

In their explanation, Keshia and Glynis link temperature to sunlight. However, they misattribute cooling to discharge rate where a better explanation would be to the shading of the creek. Thus, their level of understanding for standard D.1.1 was coded as level 2.

The third standard (D.2.2), on geochemical cycles, was found in only a single model created by Ewen and Juma in block 3 (Figure 4.6b).

The three remaining conceptual understanding standards found in this set of students' models fall under Standard F - Science in Personal and Social Perspectives

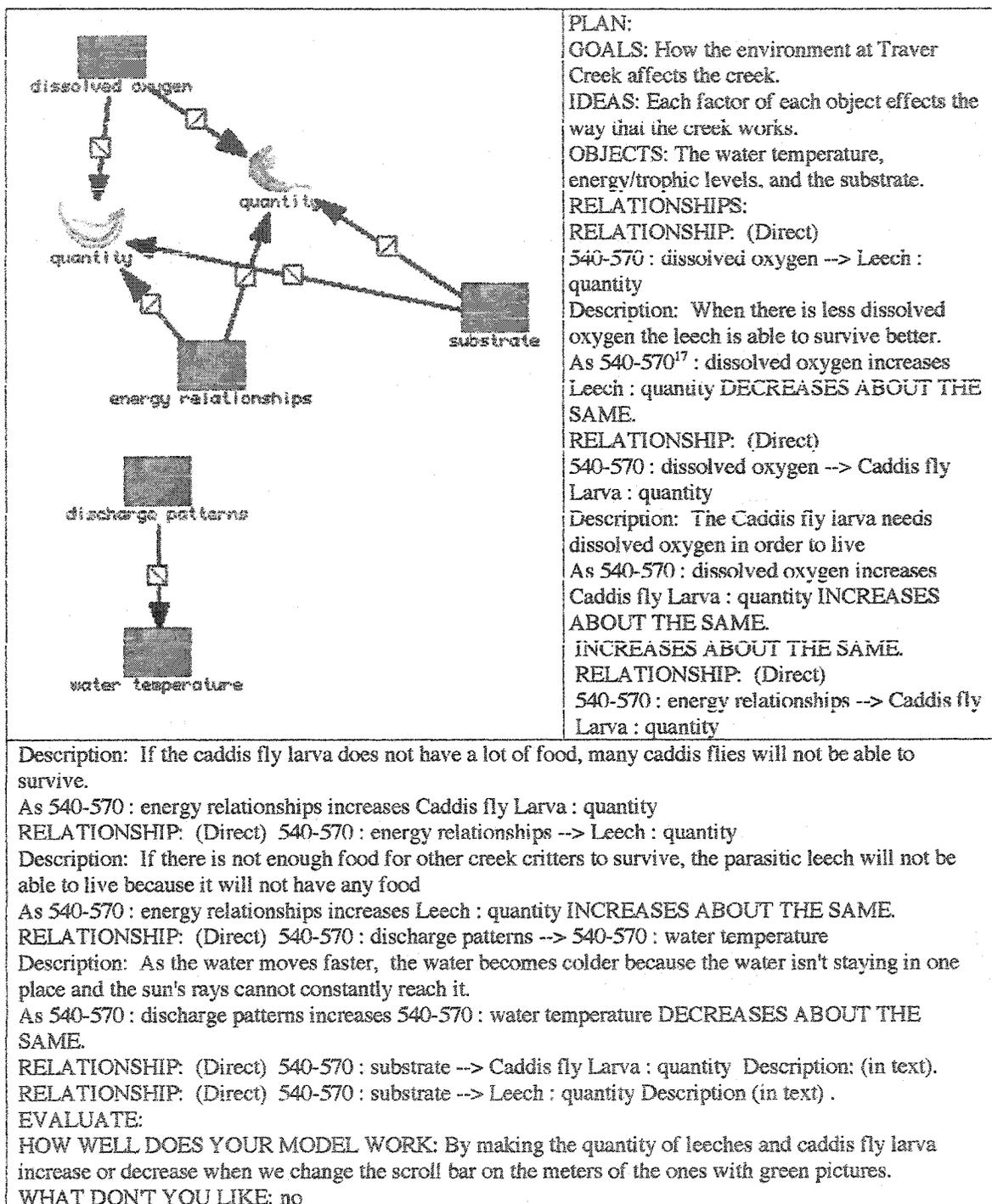


Figure 4.7: Keshia and Glynis's model of Traver Creek.

¹⁷ Note: 540-570 is the designation of their creek section in feet above the Broadway Bridge

(Environmental Science). Within this content area, there were two models that included population factors (F.2.1) (Barnard & Perry, Block 1 and Ezra & Gene, Block 3), one case where the model addressed the use of natural systems (F.3.3) (Tim, Block 7) and two models with content which mapped onto F.5.3 - progressive changes (Mallory & Carol, Block 3, Barth & Osborn, Block 7). The median level of understandings expressed in these few cases was Level 2 (Figure 4.6b).

The determinations of understanding in the models were limited by the information that students provided (or failed to provide). Many explanation fields were left empty so inferences about actual understandings were limited. Even so, it is interesting to note that relatively few demonstrations of understanding were coded at Level 1 (non-literate) (Figure 4.6). This observation would imply that students are including content in their models that they understand and are not including weaker understandings. Thus models as assessment tools for student understandings may allow student to demonstrate "strengths" in their understanding in contrast to other forms of assessment that may target students' weaknesses.

Individual models tended to include a small subset of the intended and additional content (mean of 2.5 standards). However, a small number of conceptual understandings should not be interpreted as weak understandings. Modeling combines an interaction between strategic, epistemological and conceptual understanding. Part of the nature of science is a value placed on parsimony in explanations and models. Thus, a more scientifically valued model may actually have less content in the model than an artifact that attempts to demonstrate a broad range of understandings. The parsimony and

elegance of the models need to be analyzed as part of students' strategic understandings (Standard E.1.3 implementation).

Creek Models - Cycle 2: The Final Artifact

For the second set of models created at the end of the semester, students were asked to build a model that demonstrated "in-depth understanding of a stream ecosystem and that included physical, chemical, and biological factors of the stream" (Appendix F). This modeling assignment was much more open in terms of which conceptual understandings students might include compared to the first modeling assignment.

Forty-six models representing the work of 85 students were collected from the classroom server and analyzed. Nineteen models were created by male pairs, fifteen by female pairs, seven by mixed gender pairs, and five by individuals working alone. In this set of models, students created, on average 5.42 objects (range 1-12), 10.38 factors (range 5-22), and 13.82 relationships (range 5-36) in each model ($n_a=46$; $n_s=85$) (Table 4.4). These simple statistics show the greater complexity of the second set of models when compared to the first set (Table 4.3).

Table 4.4: General features of student models in model cycle 2 ($n_a=46$; $n_s=85$).

Model Attribute	n_a	Range	Mean	Stdev
Plan	39			
Number of Objects (O) in a Model	46	1-12	5.37	2.71
Number of Factors (F) in a Model	46	5-22	10.28	3.85
Number of Relationships (R) in a Model	46	5-36	13.72	6.96
Factor : Object ratio (F/O)	46	.89-8	2.21	1.52
Relationship Integration (R/F)	46	.73-2.31	1.29	.33
Shortest Factor Path	46	2-5	2.91	.89
Longest Factor Path	46	3-8	5.17	1.54
Evaluation	31			

The largest fraction of the defined factors (32%) was ascribed to the biological organisms in the creek (e.g. population number for taxa 1, birth rate, etc.) (Table 4.5). The water quality index and its related nine parameters accounted for the second largest cluster of factors (29%). In this set of models, all nine water quality parameters (Mitchell & Stapp, 1994) were accounted for in different models. Thirty-six models included the factor of dissolved oxygen. Temperature was the second most frequent water quality parameter, included in sixteen models. Nitrates and total solids were the least common factors in this cluster, appearing in just seven models. The Water Quality Index (WQI) or general stream health was a factor in thirteen models.

Table 4.5: Content areas of factors in student's computer models cycle 2 ($n_a=46$; $n_s=85$).

Content area of Model Factor	Number of Factors	Frequency (%)
WQI and 9 water quality tests (DO, °T)	139	29
Physical Assessment factors	54	11
Benthic Organisms and Algae	147	32
Biological processes (Ps, Rs, growth, etc)	13	3
Weather	33	7
Pollution	55	12
Land Uses/Human Activity	31	7

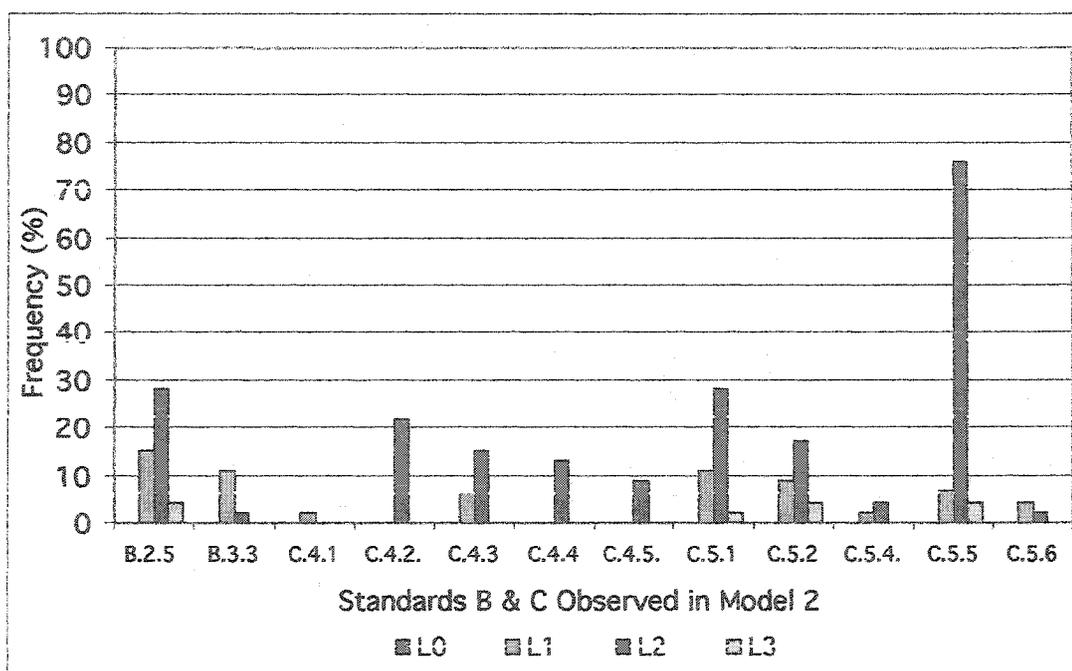
The physical assessment factors in model two accounted for only 11% of the defined factors (Table 4.5) in contrast to the first set of models where they account for 39% of the factors. The most frequently represented physical assessment factors were discharge (15 models) and bank erosion/stability (11 models). Outside influences, such as weather, pollution and land uses/human activity, accounted for 26% of all factors in the second set of models (only 14% of factors in Model 1).

Twenty three percent (23%) of the relationships involved some factor outside of the creek causing a change on a factor inside the creek and 65% involved within creek

relationships. The most common within stream relationship (23%) was that of some creek factor (e.g. dissolved oxygen (45 cases) or pH (21 cases)) affecting the benthic organisms. The high number of relationships between the physical and chemical factors of the creek and the biological organisms can also be seen in analysis of the *National Science Education Standards* addressed with 87% of the models addressing standard C.5.5 (environmental limits, niches) (Figure 4.8a).

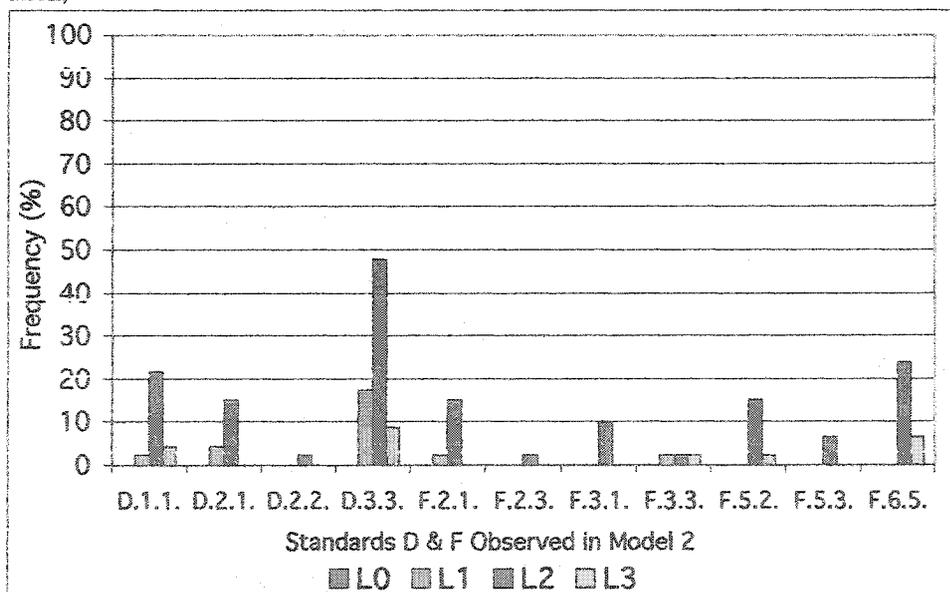
The content expected in the models (Table 4.2) mapped onto an average of five standards in the conceptual understandings frame (standards B, C, D & F) per model or ~11% of the curriculum. However, the set of models included understandings that mapped over 23 of the standards (51% of the total curriculum). Students did not demonstrate understandings on five of the intended standards (B.3.1, C.1.5, C.3.5, F.1.3, & F.3.3; Table 4.2) but they demonstrated understandings on one standard (F.2.3 - Natural hazards) that was not expected in this assignment. Ten models (22%) addressed all four science areas (B, C, D, F) and 21 (46%) addressed three of the four areas. Ninety-one percent of the students demonstrated overall (median) understandings of level 2 or higher with a few models (4.4%) indicating a more robust Level 3 understanding across the entire model. However, conceptual understandings were not demonstrated evenly across models with modelers demonstrating more robust understandings in some standards and weaker understandings in other standards.

Examining the demonstrated conceptual understanding more closely by major content strands, it was found that students generally demonstrated more competent understandings in the Life Sciences (Standard C), with almost 90% demonstrating understandings at level 2 or higher. In the Earth Sciences (Standard D), almost two-



a.

B.2.5 Physical states, B.3.3 Red/ox and acid/base Rx's, C.4.1 Biochemical cycles, C.4.2 Food chains, C.4.3 Interactions, competition, C.4.4 Population potentials, C.4.5 Human ecology, C.5.1 Entropy, C.5.2 Photo and molecular synthesis, C.5.4 Adaptations, C.5.5 Niche concept, C.5.6 Energy flow and conservation,



b.

D.1.1 Energy sources, D.2.1 Chemical cycles, D.2.2 Chemical cycles - carbon, D.3.3 Interactions in earth systems, F.2.1 Population growth patterns, F.3.1 Human use of resources, F.3.3 Use of natural systems, F.5.2 Hazardous consequences, F.5.3 Natural hazards, F.6.5 Human effects on habitat

Figure 4.8: Conceptual understandings demonstrated in Model 2 ($n_a=46$; $n_s=85$). Chart a shows Std B: Physical Science and Std C: Life Sciences Chart b shows Std D: Earth and Space Science and Std F: Environmental Sciences.

thirds of the students showed overall understandings of Level 2 or higher. This was followed closely by understandings in Standard F (Environmental sciences) with over half of the student's models demonstrating level 2 or higher conceptual understandings.

Another way to look at the data is to examine the number of models that demonstrated Level 1 (non-scientific) conceptual understandings (Figure 4.8). The low frequency of Level 1 understandings indicates content students attempted to include in their models although they may not have understood this content. As seen in Figure 4.8, fewer than 10% of the conceptual understandings were assessed at Level 1. The highest proportion of low understandings occurred in Standard B- The Physical Sciences, where 24% of the models showed an overall non-scientific understanding of Physical Science content (Level 1)(Figure 4.8).

In the next sections, the conceptual understandings demonstrated in the second set of models will be examined for each of the four major content standards using excerpts from the students' models to illustrate their understandings.

Standard B: Physical Science

Twenty-six models (57%) included content in chemistry and the physical sciences. This content mapped onto two standards, standards B.2.5 (states of matter) and B.3.3 (Chemical reactions) (Figure 4.8a). Standard B.2.5, dealing with solids, liquids and gases, was addressed in the context of gas solubility, specifically dissolved oxygen's relative solubility at different temperatures and was included in 22 models (48%). For example, in the model created by Anita and Inez, they describe a relationship between creek temperature and creek oxygen, which maps onto standard B.2.5.

*RELATIONSHIP: (Direct) Creek : Temperature affects Creek : Oxygen:
As Creek : Temperature increases, Creek : Oxygen DECREASES by
ABOUT THE SAME.*

*Explanation: Oxygen decreases in high temperatures because gases such
as oxygen dissolve more easily in cooler water.*

(Anita and Inez, Block 3, Model 2 B.5.2 L2)

Standard B.3.3, which included content involving chemical reactions, acids and bases, was addressed in the context of stream pH and acid rain. This standard was addressed in only 6 (13%) of the models (figure 4.8a). Some misunderstandings were evident around this content; for example, confusion about what end of the pH scale indicates acid conditions. Joel and Patrick explained how acid rain affects pH level as follows:

*RELATIONSHIP: (Direct) Weather : Acid Rain affects Weather : PH
level: As Weather : Acid Rain increases, Weather : PH level
DECREASES by ABOUT THE SAME. This relationship takes effect
immediately.*

*Explanation: When the rain falls that has Acid in it, the PH level of the
creek goes up*

(Joel and Patrick, Block 3, Model 2, B.3.3, L2)

In the relationship (graphic) they created, they have the correct relationship between increasing acid input and decreasing pH level. However, in their explanation they have acid rain increasing the pH level of the creek (becoming less acid), when the effect should be to decrease the pH level (become more acid). This relationship was coded at level 2, partially accurate because of the difference in the accuracy of the relationship and the explanation.

Also within this content standard, fell the interesting notion that salt (NaCl) content influences pH levels of the creek (It doesn't). This was evident in a model created by Edmund and Stephan in Block 5 shown below.

*RELATIONSHIP: (Direct) Salt : Salt Content affects Acidity : Ph level:
As Salt : Salt Content increases, Acidity : Ph level INCREASES by
ABOUT THE SAME. This relationship takes effect immediately.
Explanation: the salt will effect the balance of acid in the water because it
is a type of pollution when unbalanced
(Edmund and Stephan, Block 5, Model 2, B.3.3 = L1)*

An explanation like this illustrates a tendency for overgeneralization, namely that any given pollution source will affect any chosen water quality parameter. This relationship was coded as Level 1 - non-scientific for standard B.3.3, as were four other attempts at this standard.

Standard C: Life Science

All but two of the models (95.6%) included some biology content that mapped onto Standard C. This content fell into two sub-standard areas: C.4 - The interdependence of organisms and C.5 - Matter, energy & organization in living systems (Figure 4.8a). Twenty-one models (45.6%) included some content about interdependence from the C.4 standard. In C.4, the most common content understanding expressed understandings about food chains - C.4.2 (10 models) and competition - C.4.3 (10 models). These understandings were demonstrated at Level 2 (Figure 4.8a).

Most of the models (93.5%) included some content in area C.5 dealing with the interactions of the biotic and abiotic worlds. One understanding that appeared in the greatest number of models relates to standard C.5.5, the niche concept, which includes the understanding that biological organisms have certain tolerances or limits to environmental factors. A common representation of this understanding was the notion that benthic macro-invertebrates have different sensitivities to oxygen levels. One

example of a relationship that illustrates this understanding was in a model created by

Ayesha and Noelle in Block 1:

RELATIONSHIP: (Direct) As The creek : D.O. increases, benthics : amount of taxa 1 INCREASES by ABOUT THE SAME.

Explanation: Species that can not tolerate low levels of dissolved oxygen, taxa #1's will be replaced by a few kinds of pollution-tolerant organisms, taxa#2's, and #3's

(Ayesha and Noelle, Block 1, Model 2, C.5.5 L3).

Another example comes from a model created by Heather and Elnora (Figure 4.9). In

their model, they wrote:

FACTOR: dissolved oxygen

Range: low / medium / high

Description: The dissolved oxygen level affects the health of the stream.

One of the largest factors of A HEALTHY STREAM is the amount oxygen in the water. Because the Taxa1 benthics are so sensitive to changes, if the oxygen level decreases, a lot of the Taxa 1 benthics die out. If the Taxa 1 group dies out, the taxa 2 and 3 groups are more likely to survive because they do not have to compete with the Taxa1.

RELATIONSHIP: (Direct) STREAM : dissolved oxygen affects taxa 1 : population. This relationship takes effect immediately.

Explanation: When the dissolved oxygen increases, it allows the more sensitive benthics (Taxa 1) to thrive and grow, therefor increasing the population.

RELATIONSHIP: (Direct) STREAM : dissolved oxygen affects taxa 3 : dissolved oxygen. This relationship takes effect immediately.

Explanation: The DO affects taxa three because when DO level goes up taxa one population goes up which means there is competition with taxa three, so when DO goes up taxa three goes down. (Heather and Elnora, Block 7, Model 2)

The Heather and Elnora model shows content understanding not only about how environmental limits affect certain benthic groups in the DO effects on Taxa 1 (C.5.5) but also about how competition between different taxa groups affects the relative abundance

of those different groups (understandings which map onto standards C.4.3 & C.4.4). Other environmental factors that were included in the models, which influenced the abundance of certain groups of organisms, were temperature, pH, turbidity, and phosphates (for algae).

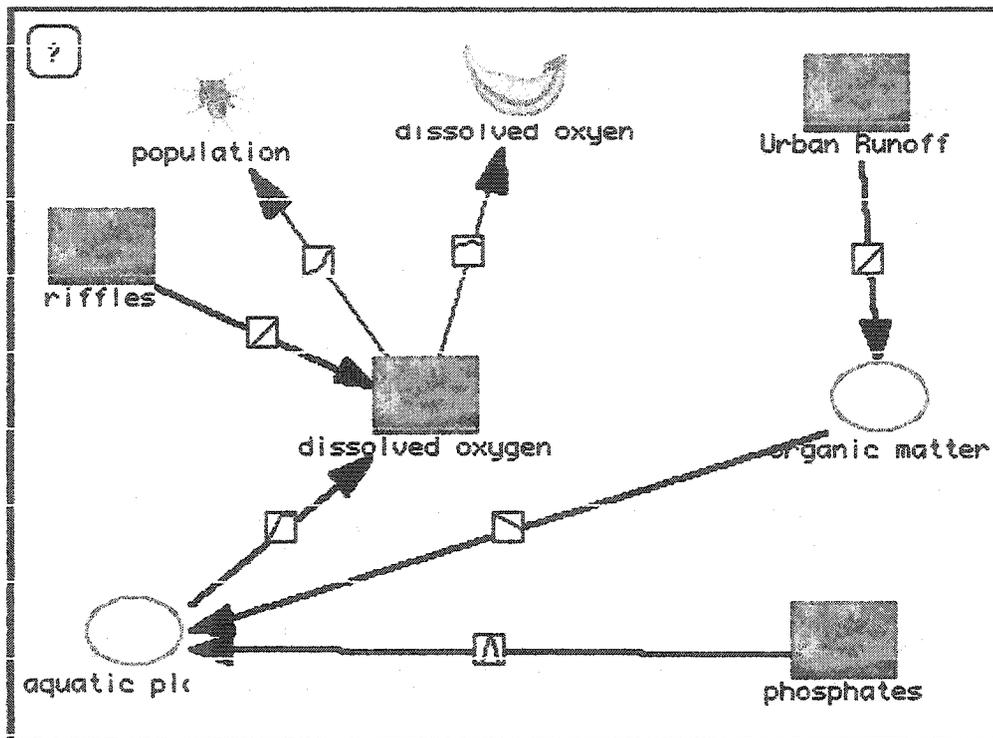


Figure 4.9: Heather and Elnora's model 2 factor and relationship map showing impacts on dissolved oxygen and the impacts of oxygen on the taxa groups.

In addition, within the life science conceptual understandings, students frequently included understandings that mapped onto the concept of decomposition (Standard C.5.1). These conceptual understandings were found in 19 models (41%) (Figure 4.8a). Heather and Elnora's model (Figure 4.9) provides an example of this type of understanding.

FACTOR: organic matter.

Range: low / medium / high

Description: Organic matter, when being decomposed lowers the dissolved oxygen. The organic matter comes from urban runoff, pet wastes, lawn fertilizers, agricultural runoff, fecal matter, and plant matter.

RELATIONSHIP: (Direct) As STREAM : phosphates increases, DISSOLVED OXYGEN : aquatic plants INCREASES by A BELL-SHAPED CURVE. This relationship takes effect slowly.

Explanation: When more phosphates enter the water plant growth occurs, called blooms. When these plants die, the breaking down of organic matter uses oxygen. So these blooms can decrease the oxygen after a certain point of time.

(Heather and Elnora, Block 7, Model 2, C.5.1 L3)

In their model, Heather and Elnora demonstrate understandings that decomposition is an aerobic process, consuming oxygen and thus lowering overall dissolved oxygen levels of the stream. This in turn affects the numbers of taxa 1 and taxa 3 benthics as seen in the earlier excerpt from their model. This chain of factors and relationships also represents understandings of how the earth systems (Standard D.3.3) interact with biological systems. The next section explores students' understandings in the earth sciences

Standard D: Earth Science

Thirty-six (78%) of the models included content that would fall within the earth science standard. Most frequent were conceptual understandings within standard D.3.3 - interactions among earth systems. This broad standard was addressed by 74% of the models (57% at level 2 or higher) (Figure 4.8b). One example is Heather and Elnora's model (Figure 4.9), where they show how stream:riffles affect stream:dissolved oxygen which in turn affects the numbers of taxa 1's and taxa 3's. They describe the first relationship in this chain as:

RELATIONSHIP: (Direct) As STREAM : riffles increases, STREAM : dissolved oxygen INCREASES by ABOUT THE SAME.

This relationship takes effect slowly.

Explanation: Riffles mix oxygen into the water. If the number of riffles in a part of the stream because of things like a changing substrate or flooding there will be more D.O.

(Heather and Elnora, Block 7, Model 2, D.3.3 – L3, figure 4.9)

Other types of earth science conceptual understandings included in models were

D.1.1 sources of energy in 13 models, D.2.1 on geochemical cycles in 9 models and D.2.2 on the movement between reservoirs in 1 model. These were addressed by most students at a Level 2 understanding (Figure 4.8b). An example of understandings that map onto standard D.1.1 was found in Annie and Mackenzie's model (Figure 4.10) where they tried to "show how Land Use and Practice (New Apartment Building) would affect the water quality." In this model, they created a relationship between Creek:Turbidity and Creek:Temperature and explained the relationship as:

RELATIONSHIP: (Direct) As Creek : Turbidity increases, Creek : Temperature INCREASES by A LITTLE.

This relationship takes effect immediately.

Explanation: As the turbidity of the creek increases, so does the temperature. This happens because the suspended solids absorb the suns heat.

(Annie & Mackenzie, Block 7, Model 2, D.1.1 L 3)

In this explanation, they identify the sun as a source of heat energy in the stream system.

Annie and Mackenzie's model also includes examples of understandings of earth science standards D.2.1 and D.3.3. For standard D.2.1 (geochemical and nutrient cycles), they have constructed relationships to show how sewage and fertilizers affects stream nitrate and phosphate levels, which go on to affect water quality, "but a little is necessary for plants to grow" (Figure 4.10). The relationships between erosion, creek temperature and turbidity and water quality addresses Standard D.3.3 (interaction of earth systems).

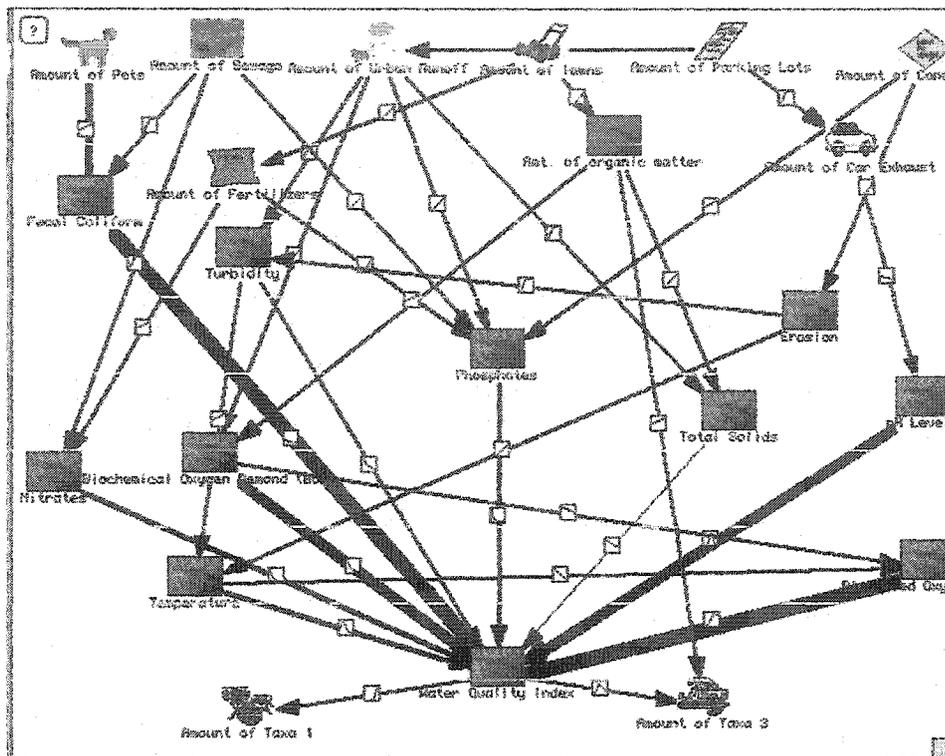


Figure 4.10: Annie and Mackenzie's model 2 factor and relationship map showing how development of new apartment buildings affects the stream.

This same model clearly shows outside factors affecting factors of the creek. These outside factors reveal conceptual understandings that map onto Standard F and are explored in the next section.

Standard F: Science in Personal and Social Perspectives

Twenty-six models (57%) included content that would fall within the environmental science standards (Standard F). Conceptual understandings were found that mapped onto seven of the Standard F objectives (Figure 4.8b) and most were demonstrated at the second level of understanding. Fourteen models included content that mapped onto standard F.6.5, which included human impacts on other species. Annie and Mackenzie's model (Figure 4.10) provided an example of this conceptual understanding. In their model, they included factor and relationship chains from human

impacts such as lawns, pets, and parking lots to the populations of Taxa 1 and Taxa 3 benthic macro-invertebrates. In their description of urban runoff, Heather and Elnora's model (Figure 4.9) provides another example of conceptual understandings in standard F.6.5.

FACTOR: Urban Runoff

Initial Value: medium Range: low / medium / high

Description: Urban runoff is introduced to the stream through heavy rains and stormsewers. Runoff includes things such as toxins and heavy metals along with silt and lawn chemicals and cleaning solutions and oils and detergents. These things increase phosphate and nitrate levels. Illegal sanitary connections can increase fecal coliform levels which makes the risk of disease bearing pathogens more likely to be in the water.

(Heather and Elnora, Block 7, Model 2, F.5.6 L3)

This factor was connected via relationships to organic matter, then to aquatic plants, and eventually to Taxa 1 and Taxa 3 populations (Figure 4.9). Kiley and Leah's model (Figure 4.11) also demonstrated a high level (L3) on this standard (F.6.5). Their understandings were exhibited in their description of the factor "amount of lawns and gardens" and in the relationship affecting the amount of trees

Eight of the models used "population objects" in their models, which mapped onto standard F.2.1. When a student chooses to create a population object, such as Taxa 1, instead of a "normal" or "background" object, Model-It 3.0b automatically defines three factors for that object: rate of growth, rate of decay and count. The default definitions set the initial rate at 0.5 - population doubling every 2-time steps.¹⁸ Students then have the option of redefining the factors and their ranges to more appropriate levels.

An example of a model that used population factors extensively was the one created by Cliff, Sean, and Jack in Block 1 (Figure 4.12). In this model, they did not re-define any of the population factors but they created relationships between the population factors and other factors. So, their level of understanding for F.2.1 (Population change from effects of births and deaths) was coded as Level 2.

The content in eight models also mapped upon standard F.5.2 - Human activities causing accelerated rates of change. Kiley and Leah in Block 3 created a model that illustrated conceptual understandings within this standard. In their model (figure 4.11), Kiley and Leah created their key independent factor, Development: number of houses (represented by the small group of people). As development increases, it affects discharge and fertilizers in the creek, which in turn has a negative impact on the overall health of Traver Creek.

Kiley and Leah's model also demonstrated understandings on standards F.3.1 and F.5.3, which were demonstrated in just a few models (Figure 4.8b). Standard F.3.1, the use of resources, is illustrated in at least three points in their model. The first point is in their relationship between development: number of houses and houses: roads and driveways, which they explain as:

RELATIONSHIP: (Direct)

As development : number of houses increases, houses : roads and driveways INCREASES by ABOUT THE SAME.

Explanation: As more and more houses go up, families move in and need roads and driveways to get into town and stuff. Therefore, the amount of driveways increase by the same amount as the houses, because one house

¹⁸ A rate of 0.5 is an extreme rate. Population rates work on the same principle as interest rates. Rapid human population growth is often in the .03-.04 range, doubling the population size every 15-25 years).

needs one driveway, and not too many roads are needed per house
(Kiley & Leah, Block 3, Model 2, Figure 4.11)

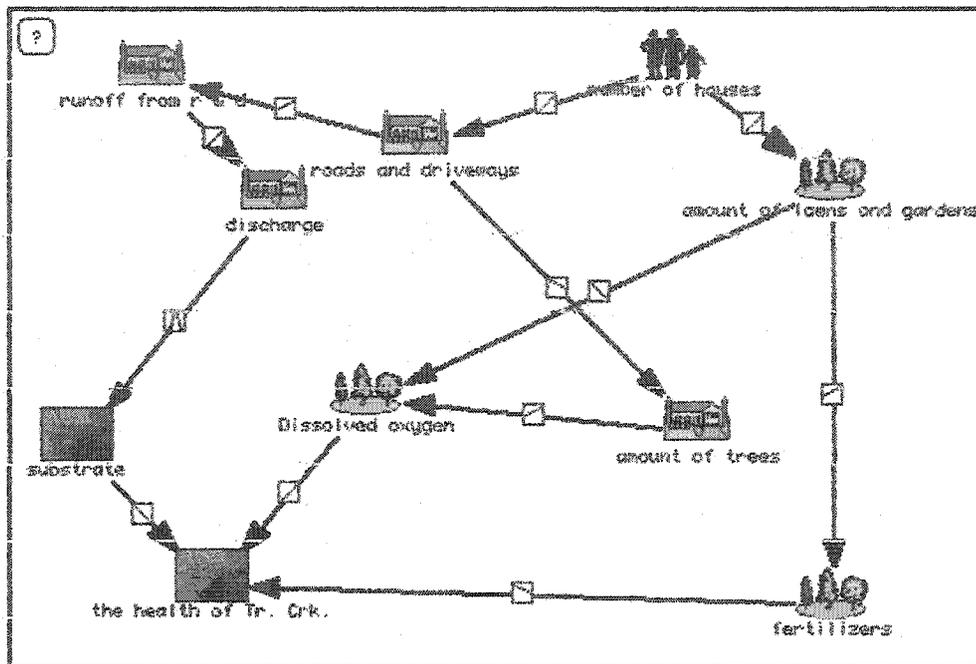


Figure 4.11: Kiley and Leah's model on how buildings will affect the creek. Cycle 2, Block 3.

The second point where they demonstrate this understanding is in the relationship between roads, driveways, and runoff, which they explain as:

RELATIONSHIP: (Direct)

As houses : roads and driveways increases, houses : runoff from r & d INCREASES by A LITTLE.

Explanation: With more roads and driveways, more and more salt will have to be put on them, as well as other chemicals from the car exhaust and other miscellaneous. These will eventually run with rain water into the creek. A little because one road will only create a small effect on the amount of runoff. One road will lessen the capacity to absorb rain water, which, with the road will now run into the creek.

(Kiley & Leah, Block 3, Model 2, Figure 4.11)

Lastly, the third point is where development affects the amount of lawns and gardens:

RELATIONSHIP: (Direct)

As development : number of houses increases, lawns and gardens : amount of lawns and gardens **INCREASES** by **ABOUT THE SAME**.

Explanation: As more families move into the houses, they will probably want lawns, and might create gardens. We can assume that the deficiency of lawns and gardens for people living in apartments will be made up for by the extra gardens and lawns by the people who live in houses around the creek

(Kiley & Leah, Block 3, Model 2, Figure 4.11 F.3.1 L2)

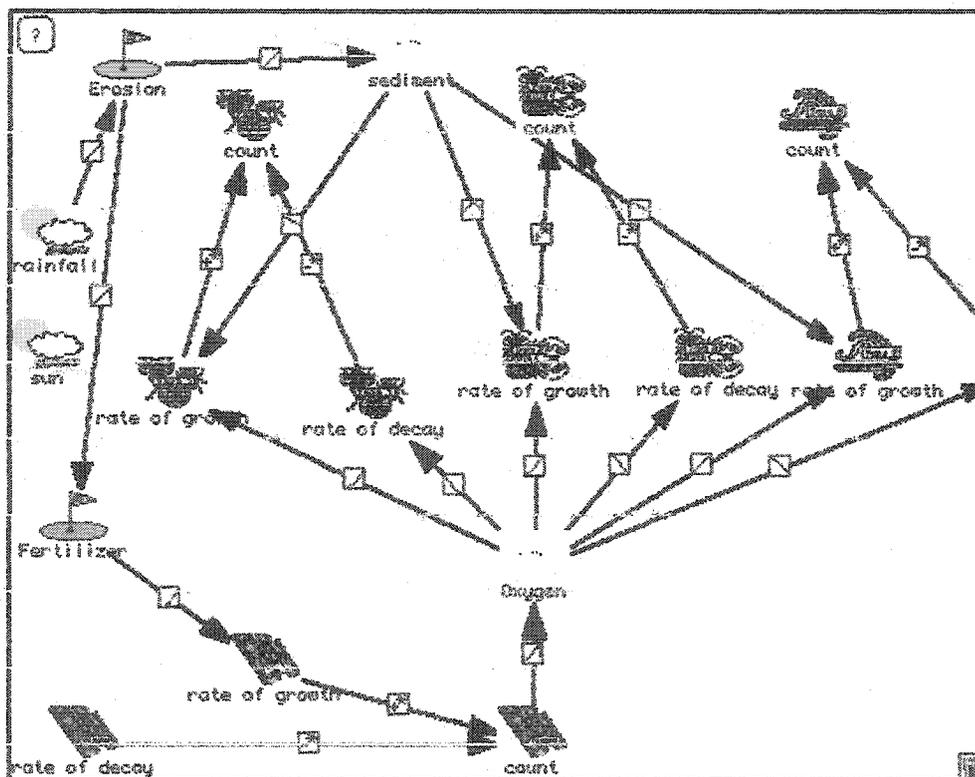


Figure 4.12: Model created by Cliff, Sean and Jack, Cycle 2, Block 1 showing the effects of a golf course on the creek.

Standard F.5.3 - progressive changes in the environment that can affect society, was demonstrated in this model through these same relationships and in the relationship between houses:discharge affecting Traver Creek:substrate. In their explanation, Kiley and Leah make specific reference to increased siltiness of the creek substrate, which directly corresponds to the concepts of sedimentation and erosion contained in this standard (F.5.3 = L2).

The second model assignment provided opportunities for students to demonstrate understandings on 23 standards (51% of the Traver Creek Curriculum). However, students' models demonstrated understandings on a much smaller fraction of the possible content, mapping onto an average of five standards per model (11% of the curriculum content). The restricted content expressed in the individual models is not a negative result since scientists value models whose simplicity generates the greatest explanatory power.

In the models, students generally included content they knew and understood. Ninety-one percent of the students demonstrated overall (median) understandings of level 2 or higher with a few models (4.4%) indicating a more robust Level 3 understanding across the entire model. However, conceptual understandings were not demonstrated evenly across models with modelers demonstrating more robust understandings in some standards and weaker understandings in other standards.

This concludes the analysis of student understandings in the four artifacts that were assessed. In the artifacts, students addressed most of the expected standards and several that were not explicitly part of the assessments. For example, the first model assignment was mapped onto seven conceptual standards but observations of student understandings mapped onto sixteen standards. At the same time, the average number of standards that students addressed in an individual artifact was much lower than the expected number in all artifacts except the report. Of the 45 conceptual standards addressed by the curriculum and 23 assessed, individual students demonstrated understandings of an average of 9.3 conceptual standards (range 0-17). Therefore, while the entire set of artifacts provided feedback on student understandings across 51% of the

curriculum, individual students were only demonstrating, on average, their competencies on 21% of the curriculum.

Understandings demonstrated on the Pre and Post Tests

On the pretest, most students demonstrated limited scientific understandings. That is, while some students may have demonstrated high levels (Level 3) of understandings on one or two standards in the conceptual dimensions, none demonstrated this level of understanding across the test (determined by their median scores). However, on the post-test, students showed significant gains ($p=.05-.000$) in their median scores across conceptual understandings when compared to their scores on the pretest (Figure 4.13 - CU). On the post-test, 26% of the students attained a median Level 3 understanding.

On the pretest, students showed higher levels of understanding in the Physical (Standard B) and Earth Sciences (Standard D), where 30.2% and 16.7% respectively of the students showed Level 3 understandings (Figure 4.13). At the beginning of the project, no student demonstrated Level 3 understandings across the Life (Standard C) or Environmental Sciences (Standard F) questions (Figure 4.13).

In the post-test, students showed significant gains in each of the major science domains ($p=.05-.000$). The largest gains were in the Physical Sciences (B) and Earth & Space Science (D) (Figure 4.13). The percentage of students demonstrating understandings at Level 0 - no understanding (red bar), markedly decreased in each of the content areas. Each of the major science content areas and their associated substandards are examined more closely below.

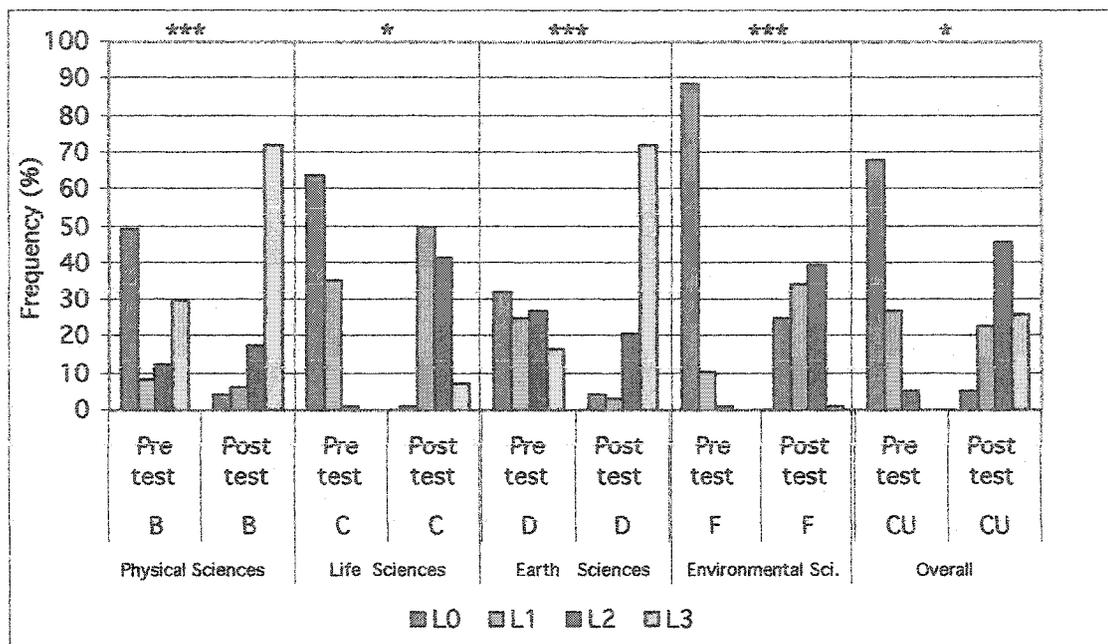
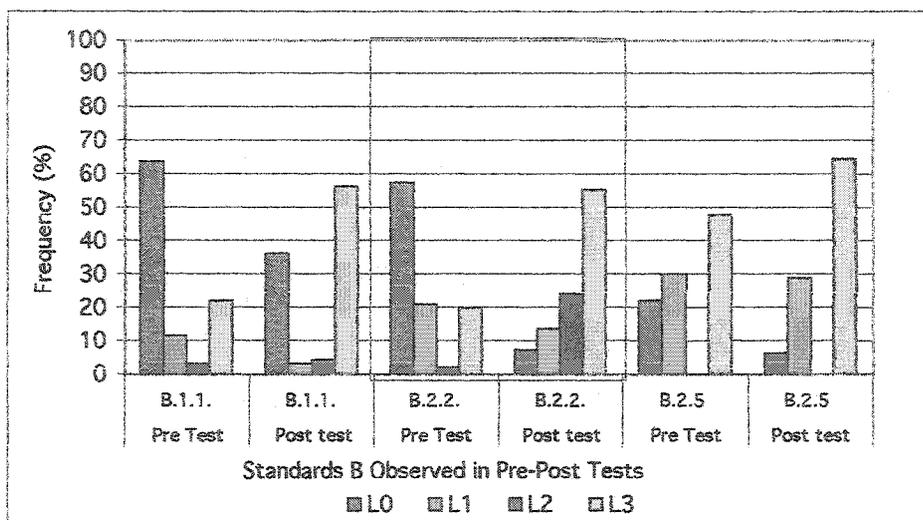


Figure 4.13: Comparison of Students' levels of conceptual understandings on the Pre- and post-test across each of the major content areas ($n_s=96$). Wilcoxon non-parametric sign-rank tests were used to compare students' pre and posttest achievement (* $p \leq .05$, ** $p \leq .01$, * $p \leq .001$).**

The pre and post-tests included questions in the Physical Science standards about the structure of the atoms (B.1.1. Q 16 in part), use of the periodic table in naming compounds (B.2.2. Q's 16, 23, 26); physical and chemical changes (B.2.5. Q's 20,25) and chemical reactions (B.3.1. Q 24, 26c)(See Appendix G & H). Students demonstrated significant gains ($p \leq 0.001$) in understanding on each of these standards (Figure 4.14).



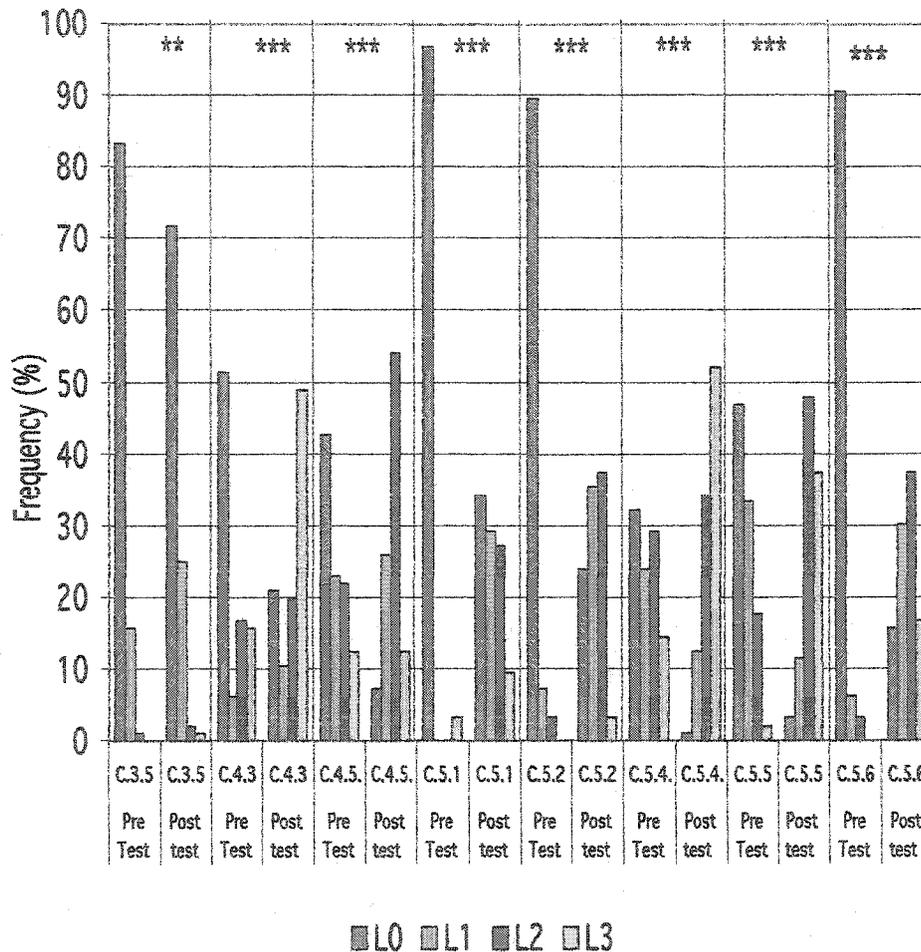
B.1.1: Structure of atoms, B.2.2: Periodic Table, B.2.5: Physical & Chemical changes.

Figure 4.14: Comparison of Students' levels of understandings on the Pre/post-test across the Physical Sciences - Standard B ($n_s=96$). ($p \leq .01$, *** $p \leq .001$).**

In the Life Sciences (Standard C), all standards showed significant gains ($p = 0.05$ to $.000$) (Figure 4.15). Some of the largest gains were in students' understandings of adaptations (C.5.4 - Question 22) with an additional 37.5% of the students reaching the level of scientific literacy (Level 3) on the post-test. This was followed by understandings of environmental tolerances (C.5.5 - Questions 10, 11, 14, 15 S1a) with 37.5% of the students demonstrating scientific literacy at the level of the standards on the post-test compared to only 20% on the pre-test. Also in the understanding of organism interrelationships (C.4.3 - Questions S1a), an additional 33.4% of the students achieved scientific literacy (Level 3 understandings) on the post-test compared to the pre-test (See also Appendix I).

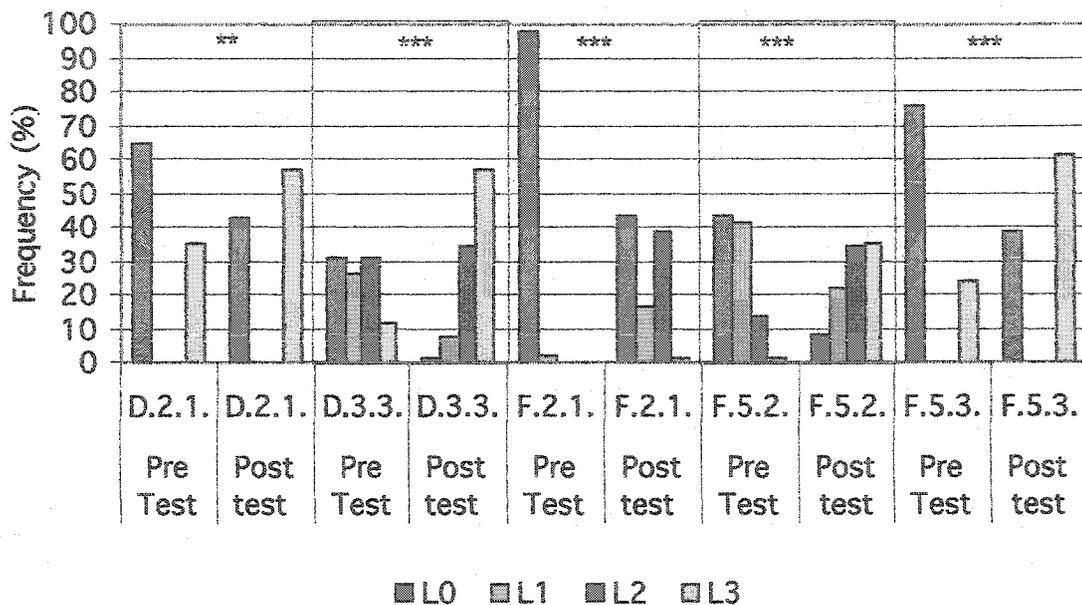
In the area of Earth and Space Sciences, student demonstrated significant gains on standards D.2.2 (chemical cycles) and D.3.3 (interactions of earth systems) ($p = .01 - .000$) (Figure 4.16). On standard D.3.3 (interactions of earth systems), assessed in

questions 17b and 27, 57.3% of the students demonstrated a Level 3-scientific literacy understanding on the posttest compared to only 11.5% on the pre-test.



C.3.5 – Classification, C.4.3 – Competition, C.4.5 – Human Ecology, C.5.1 – Entropy, C.5.2 – Photosynthesis, C.5.4 – Adaptations, C.5.5 – Niches, C.5.6 – energy flow.

Figure 4.15: Comparison of Students' levels of understandings on the Pre- and post-test across the Life Sciences Standard C (n_s=96). (* p ≤ .05, ** p ≤ .01, * p ≤ .001).**



D.2.1 – chemical cycles, D.3.3 – interactions of earth systems, F.2.1 – population growth, F.5.2 – environmental consequences, F.5.3 – natural hazards.

Figure 4.16: Comparison of Students' levels of understandings on the Pre-/post-test across the Earth and Space Sciences (Standard D) and the Environmental Sciences (Standard F) ($n_s=96$). ($p \leq .01$, *** $p \leq .001$)**

Students also demonstrated more robust understandings about watersheds, assessed on questions 1, 2 and 8 with almost 70% of the students demonstrating a literate level of understanding on the post-test.

The three standards from content Standard F - Environmental Sciences also all showed significant gains in students' understandings ($p = .000$, Figure 4.16). The largest gain in this content area was on question 12 dealing with erosion and stream substrates (F.5.3) with 61.5% of the students answering the question correctly on the post-test.

The constructed response questions (e.g. question. 15) showed considerable change between the pretest and the post-test. On the pretest, many students left these questions blank or simply wrote, "don't know." On the post-test, the quality and elaborateness of the answers varied considerably. For example on question 15, which

was used as an example of content analysis in Chapter 3 (Table 3.6), Kelly demonstrated level 3 (scientifically literate) understandings on the three life science standards (C.5.1, C.5.2, and C.5.5) and Level 2 (semi-literate) understandings on the two environmental standards (F.2.1 and F.5.2) (Table 3.6). Kiley's answer was coded at Level 2 (semi-literate) for the life science standards and Level 1 (non-scientific) for the environmental standards and Mack's responses were coded at Levels 1 and 0 (Table 3.6).

There was also some evidence that students were thinking in a "cause and effect" manner on question 15 that was not evident on the pretest. Students used phrases such that one factor would go up and another would decrease (e.g. Carol - B3, Mallory - B3, Leia - B3). A couple of students (e.g. Kwame - B3, Leah - B3) drew graphical representations showing how one factor would affect another in this question.

In the scenarios, students also wrote more accurate and elaborate answers on the post-test than on the pre-test. Here, conceptual understandings were demonstrated in the hypothesis forming in Part A. At the beginning of the semester, for the first scenario, involving an invasive species (zebra mussels or ruffe) and a decrease in native species, 37 students hypothesized that this was the result of a biotic interaction (C.4.3) and 26 identified it as an environmental consequence (C.5.5). On the posttest, 69 students hypothesized a biotic interaction, 12 an environmental consequence and 7 incorporated both possibilities. The level of understanding also changed significantly ($p = .001$) with 15% of the students demonstrating a level 3 (scientifically literate) understanding of biotic interactions (C.4.3) on the pretest and 49% at Level 3 on the post-test (Figure 4.15).

The concept map question on the pre- and post-test instruments also revealed information about the students' changing conceptual understandings. On the pretest, only 6 of the 96 students added concepts to the partial map given in the question. The mean number of added concepts was 0.2 (sd = 0.9). The largest number of concepts added to the pre-test concept map was five, added by two students (Doug and Clay).

In contrast, on the post-test at the end of the unit, over half of the students (55%) made additions to the given map. Students added an average of 4.7 concepts (sd =4.9) to the map. These improvements were made to the concept map even though the only "mapping" students had done during the semester was in the Model-It program. For students who added to the map, most added between 6 and 10 concepts. Lynnette added 18 concepts to her map. Paired t-tests showed that the difference between number of concepts added to the pre and the post-tests was highly significant ($p = .000$).

Students also did a better job of linking¹⁹ their concepts on the post-test when compared to the pretest. On the pretests, only 3 students added appropriate links (2-4) between concepts on the map. On the posttest, 55% of the students added links (mean number = 4.6, sd = 5.3) between concepts. Again, paired T-tests showed these changes to be highly significant ($p = .000$).

Likewise, student explanations about their changes to the concept maps changed significantly (Wilcoxon sign-ranked test $p = .000$) between the pre and post-tests. Since on the pretest most students did not modify the map at all, it should not be surprising that they also did not add an explanation. On the post-test however, 54 students (56%) wrote

¹⁹ In a concept map, a link is an appropriately labeled line or arrow connecting two or more concepts.

something to explain their concept maps (17 explanations were coded at conceptual understanding Level 1, 33 at Level 2 and 4 at Level 3.).

Overall, students demonstrated significant changes in their understandings as measured by the pre and posttest instruments. Students' written answers were longer and more elaborate on the post-test than the pre-test. On each of the individual standards, they also demonstrated significant gains with the highest gains being in the physical and earth sciences (Standards B & D, Figures 4.14 and 4.16).

Because of the descriptive nature of this study, the pre/post test data cannot be used to make claims about student achievement within the Creek project as compared to student achievement in other environments. It is expected that students will learn content over the course of a semester. The pre-/post-test data affords a baseline measure for student understandings at the beginning of the project and a measure of their understandings at the end of the project. As such, it provides a means for comparison for demonstrations of understandings in the four major artifacts, which were described above.

This section, which analyzed student understandings in the six different assessments, laid the foundation for the final question, "Did students' understandings change over time?" This question is partially answered by considering that the lengths of the descriptions of understandings demonstrated grew longer with each artifact that was assessed. The differences in the understandings assessed in the pre- and post-tests, as described above, laid further groundwork. In the next section, understandings demonstrated by individuals and changes over time will be described and analyzed. The

data will also be examined to see if students specialized in certain content areas or exhibited other patterns of understandings.

Demonstrations of Student Understandings across Artifacts and Time

In addition to the question of “What understandings do students demonstrate in the specific artifacts?” which was answered in the previous section, this study also questioned how students’ demonstrated levels of understanding would change across the artifacts. One concern in allowing students some choice in their artifacts was that students might present patterns of understanding such as *avoidance* or *specialization*, i.e. consistently representing the same phenomena or content across the different artifacts. The results of the Wilcoxon signed-rank analysis provide answers to these questions. In this section, the results of the fourth and fifth lines of analysis, comparing understandings demonstrated across artifacts and across time, are presented and discussed.

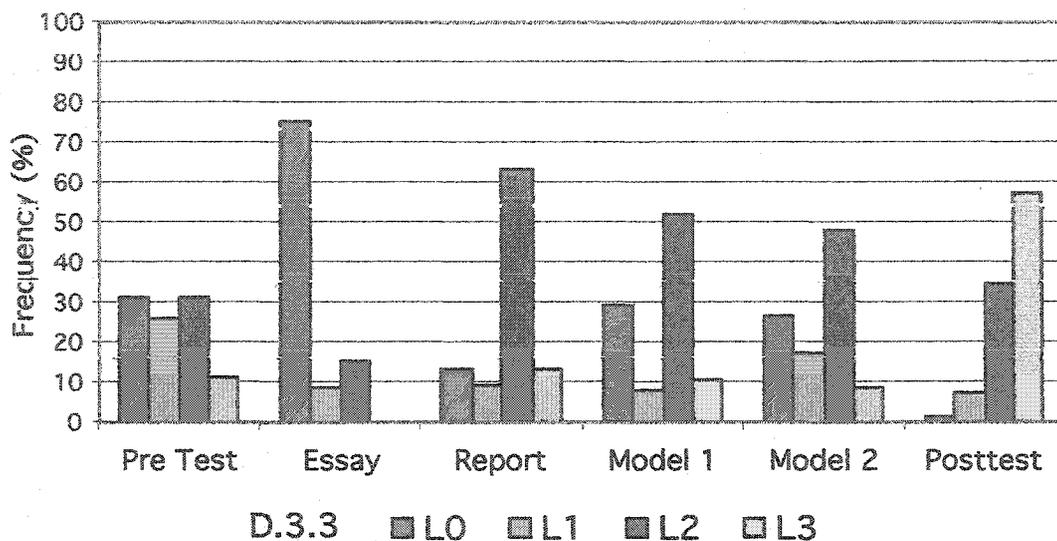
Several conceptual understanding standards were represented across the set of artifacts (Appendix B) although only one (D.3.3) was in all six assessments (Figure 4.17). The standards that were assessed in multiple measures reveal whether or not students’ understandings were stable or if (and when) they changed over the semester long creek project. However, the main finding in this step of the analysis was that even when a standard was assessed on multiple measures, students often did not display understandings in these same standards. For example, on standard D.3.3 (earth systems), which was assessed in all six measures, only 4 students demonstrated understandings on this standard in all six assessments.

Early in the semester, most students demonstrated weak conceptual understandings, a finding that was not unexpected since students had not time to engage

in the content. Most of their observations were disconnected, providing evidence that they may have picked up some knowledge of the stream, but had not yet connected these pieces of knowledge into a conceptual framework indicating an understanding of what they observed.

Four students demonstrated more connected understandings (Leia, Anita and Ezra in Block 3, Jeanette in Block 1) illustrated by the number and quality of standards that mapped onto their essays. Jeanette and Leia were also among the higher scoring students on the pretest. These facts may be the result of having studied another creek in Middle School as these students made references to Honey Creek, on the west side of the city.

Considering the next set of artifacts, both the water quality reports and the first set of computer models were created midway through the project. There were six standards that were addressed in both report and model 1 (Appendices B & I). On most of these standards, the few students that represented this content in both artifacts did so at the same level of understanding. The exception was standard D.3.3 (interactions of earth systems) where students demonstrated significantly higher levels of understanding in the models than in the reports ($n_c=36$, $p=.049^*$) (Figure 4.17).



NonPar Sign Test	Pre Test	Essay	Report	Model 1	Model 2	Post-test	pre/post comp n=96
n_c		11(8+ 1-)	5(0+ 0-)	36(13+ 4-)	45(5+13-)	62(42+13-)	96(76+ 8-)
p		0.039*	1.000	0.049*	0.096	0.000***	0.001***

Figure 4.17: Standard D.3.3: Interaction of Earth Systems across measures.
(ns $p > .05$, * $p \leq .05$ ** $p \leq .01$, *** $p \leq .001$)

Comparing the distribution of understandings demonstrated in the first set of models to the second model, there were a few changes in understanding that were evident. First, understandings mapped onto a larger range of standards in the second set of models (23 vs. 16). The difference in number of addressed standards may reflect the differential opportunities provided by the assignments as well as additional understandings developed during the intervening period. Secondly, the demonstrated understandings in each of the broad content areas were higher in the second model than the first. At the individual student and standard level, there were no significant differences between levels of understandings demonstrated in the two models (Appendix D).

The lack of apparent change within students is mostly a consequence of the low number of cases or paired comparisons (n_c values) in the pair-wise Sign-tests. Two

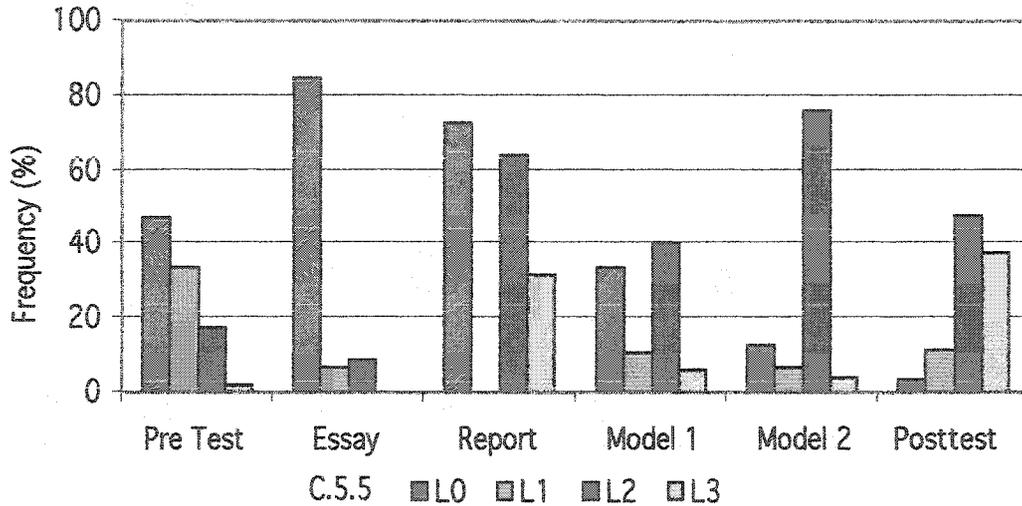
standards (C.4.2 and D.3.3) showed a tendency toward changing levels of understanding. On Standard C.4.2 (food chains), five students ($n_c=7$) showed higher levels of understanding in the second model ($p=.063$). In Standard D.3.3 (interactions of earth systems), thirteen students demonstrated higher levels of understanding in the first model than the second and five students showed a higher level of understanding in the second model ($p=.096$, $n_c = 45$ Figure 4.17, Appendix I). However, neither of these trends are statistically significant at the 0.05 level.

The mid-semester report and the second model at the end of the semester had seven content standards in common (B.2.5, C.4.4, C.5.4, C.5.5, C.5.6, D.3.3 & F.3.1). Of these, two standards showed a tendency (not significant) toward higher levels of understanding in the later artifact. In the Physical Sciences (B.2.5 – physical states) four of seven students showed higher levels in the model ($n_c = 7$, $p=.125$). In the Life Sciences (C.5.4 - adaptations), two of three students demonstrated higher levels of understanding ($n_c = 3$, $p=.500$). Again, the small number of paired cases (n_c) for these two standards means that these findings are not significant at the 0.05 levels.

Comparing demonstrated understandings in the second model and the post-test given at the end of the semester, there are more paired cases (n_c values) and several of the observed differences are significant (Appendix I). In the Physical Sciences (B.2.5), 29 students demonstrated higher understandings on the post-test, nine on Model 2 ($n_c = 41$ (29+9-), $p = .002$). In the Life Sciences, two standards showed significantly different levels of understandings on the two measures. For standard C.4.3 (interactions, competition), ten students did better on the post-test, two in the model ($n_c=15$ (10+2-), $p=.039$). For standard C.5.5 (niche concept), 31 students demonstrated higher

understandings in the post-test, 16 in their models ($n_c = 74(31+16-)$, $p=.041$ Figure 4.18).

And in the Earth Sciences (D.3.3 – earth systems), 42 students exhibited higher understandings on the post-test, 13 in their models ($n_c=62$, $p=.000$ Figure 4.17).



NonPar Sign Test	Pre Test	Essay	Report	Model 1	Model 2	Post-test	pre/post com n=96
n_c		7(5+1-)	3(3+ 0-)	46(7+ 7-)	51(9+ 8-)	74(31+ 16-)	96(79+4-)
p		0.219	0.250	1.000	1.000	0.041*	0.000***

Figure 4.18 : Standard C.5.5: The Niche Concept and Environmental Tolerances across measures. (ns $p > .05$, * $p \leq .05$ ** $p \leq .01$, *** $p \leq .001$)

In Standard C: the Life Sciences, four standards were assessed in multiple artifacts (C.3.5, C.4.3, C.5.2, and C.5.5). The niche concept (C.5.5), assessed by the reports and two modeling assignments showed the highest frequencies of demonstrated understandings in the artifacts (Figure 4.18). Of the 78 students that included this understanding in their reports, 46 used this content in their first model and 60 used it in their second model. Fifty-one (51) students included this content in both model 1 and model 2 (Figure 4.18 and Appendix I, Standard C.5.5). Nine of these students showed higher levels of understanding in their second model and eight students demonstrated

higher levels of understanding in their first model (no significance). There was a significant improvement ($p=.041$) in demonstrated understandings on the niche concept (C.5.5) between the second modeling assignment and the posttest the next week. Thirty-one (31) students demonstrated a higher level of understanding along standard C.5.5 on the post-test (Questions 10,11, 14, 15, & S1a - Appendices B & C) than on model 2.

Within the Earth & Space Science standards, Standard D, three standards appeared in multiple artifacts (D.1.1, D.2.1 and D.3.3) Understandings about the interactions of earth systems (standard D.3.3) were assessed by all six measures as discussed above. More students represented content understandings that mapped onto this standard than the other two standards in the earth science. Students showed higher levels of understanding on the essays ($n_c=11$, $p=.039$), report ($n_c=69$, $p=.000$) both models ($n_c=62$, $p=.000$) and the post-test ($n_c=96$, $p=.001$) than they did on the pretest (Figure 4.17). Comparing artifacts, the only significant difference in demonstrated levels of understanding on standard D.3.3 was between the report and Model 1 ($n_c=36$, $p=.049$). In addition, between the first and second models, there was a trend in decreasing the level of demonstrated understandings with 13 students demonstrating a lower level understanding on their second model than on their first model ($n_c=45$, $p=.096$), often because they failed to fill in their explanation boxes on Model 2.

On the second model, 17.4% of the students demonstrated a level 1 understanding for standard D.3.3 (interactions of earth systems). This is the highest percentage for Level 1 understandings among all the content standards on any artifact. The high frequency of Level 1 understandings on standard D.3.3 may be attributed to the lack of any explanation in the student's description boxes as well as inaccurate relationships.

Students' achievement for standard D.3.3 (interactions of earth systems) on the post-test was significantly better than on each of the other assessments (Figure 4.17).

In general, students' understandings along the standards increased over the course of the creek project ($p=.05-.000$). The four exceptions to this change are found in standards C.4.2, C.4.4, D.1.1 and F.5.6 (Appendix I). For three of these standards (C.4.2, C.4.4, F.5.6) there were too few cases in the sign-rank test to find a significant pattern. For standard D.1.1 (sources of energy) there was a change from a fairly even distribution of represented understandings across the three levels of understanding on the first model, to a more dominant distribution around level 2 understandings in the second model (Appendix I). In the 12 cases between the two models, four students demonstrated a higher level of understanding in the second model and four demonstrated a lower level of understanding (no statistical change).

The small numbers of paired cases of students and standards in consecutive artifacts is evidence that students are not specializing in content. That is, a student that received positive feedback on their essay when they observed eroded culverts in the stream or the effect of shading on photosynthesis did not build upon those observations when asked to create a computer model of the stream. Nor did students emphasize the same content in preparing their stream report.

Chapter Summary

In this chapter, I answered the following questions:

- What opportunities were provided by the different artifacts for students to express their scientific understandings?

- How much of the course content could potentially be represented by each artifact?
- How much of the course content was represented by the set of artifacts?
- Which parts of the curriculum were not represented by the artifacts?
- Were some constructs over-represented and others under-represented by the set of artifacts?
- What conceptual understandings, breadth and depth as mapped on to the *Standards*, did students demonstrate in the artifacts?
- What patterns of understandings were demonstrated by the students? Did some students consistently display understandings around the same content across the different artifacts, or were they displaying different understandings?
- How did students' level of demonstrated understanding change across the artifacts?

The Traver Creek project address 45 content standards. The artifacts potentially assessed more than half of the standards. Content standards in the Life Sciences (Standard C), Earth and Space Sciences (Standard D) and Environmental Sciences (Standard F) are well represented in the artifact assessments. The artifacts were weak in assessing understandings in the Physical Sciences (Standard B) especially conceptual understandings related to chemistry.

The content analysis of each artifact informed the teachers and the researcher about how useful a specific measure would be for characterizing students' understandings relative to course content. The summary of expected and observed standards in the standards matrix (Appendix B summarized in Table 4.6) shows that as a whole, students addressed most of the expected standards and several that were not explicitly part of the assessments. For example, the first model assignment was mapped onto seven conceptual standards but observations of student understandings mapped onto sixteen

standards. However, the average number of standards that students addressed in an individual artifact was much lower than the expected number in all artifacts except the report. Of the 45 conceptual standards addressed by the curriculum and 23 assessed, individual students demonstrated understandings of an average of 9.3 conceptual standards (range 0-17) across the four artifacts. Therefore, while the entire set of artifacts provided feedback on student understandings across 51% of the curriculum, individual students were only demonstrating, on average, their competencies on 21% of the curriculum. Further consideration is needed here to determine how the “explicitness” of the task structure is correlated with students demonstrating a specific understanding. Related issues will be addressed in chapter 5.

Table 4.6: Summary of expected and observed conceptual understandings along the National Science Education Standards. Values represent number of standards in that content area.

Content area.	Essay		Report		Model 1		Model 2	
	E Expected	O Observed	E Expected	O Observed	E Expected	O Observed	E Expected	O Observed
Standard B - Physical Science	0	0	0	1	0	1	2	2
Standard C - Life Science	3	3	3	9	4	9	10	10
Standard D - Earth and Space Science	2	2	1	3	3	3	4	4
Standard F - Science in Personal and Social Perspectives	5	6	2	6	0	3	5	7
Total number of conceptual Standards	10	11	6	17	7	16	21	23
Mean number of Standards per artifact:		1.29		7.95		2.5		5.15

Additional understandings, beyond the specifications of the assessments, were also observed and mapped onto relevant standards. These occurrences reinforce the need

for an evaluator to be sensitive to unexpected outcomes when trying to characterize student understandings.

For the most part, even when a standard was assessed on multiple measures, students did not display understanding on the same standard in sequential assessments. The niche concept (standard C.5.5), assessed by the reports and two modeling assignments, shows the highest frequencies of demonstrated understandings in the artifacts (Figure 4.18). This finding on the lack of follow through would not support a hypothesis that students will specialize in content when offered some choice in artifact construction (e.g. focusing on effects of salt on a creek system). In most cases, students did not demonstrate understandings of the same content in consecutive artifacts, even when those assessments offered opportunities for students to use that content. Consequently, there was insufficient data to show changes in individuals' understanding across artifacts. The pre/post results do show significant ($p = .05 - .001$) increases on seventeen standards on the post-test. But considering the total sample of students, in general, students' understandings along the standards increased over the course of the creek project ($p = .05 - .000$).

The next chapter considers the implications of using the *Standards* and artifacts in assessment programs

CHAPTER 5

DISCUSSION

In the previous chapter, I described the types of understandings students demonstrated in the artifacts they constructed. I also provided some interpretation about specific understandings that were demonstrated. In this chapter, I consider some of the larger implications of this study for classroom instruction and as a research methodology. The kinds of scientific understandings that students demonstrated during the Traver Creek project and the implications for instruction comprise the first part of this chapter. The second section focuses on the usefulness of the *National Science Education Standards* (NRC, 1996) as an assessment tool.

What do students learn? And what are the implications for instruction and assessment?

Social constructivist orientations to learning science have led several scholarly groups to research the impact of inquiry instruction on conceptual understandings (Keys & Bryan, 2001). Most of the studies cited by these authors occurred in elementary or middle school classrooms and Keys and Bryan (2001) call for more research on inquiry-based learning to support conceptual understanding at the secondary level.

This study helps fill in the gap at the secondary level. The study was situated in a public high school where the teachers were initiators and co-designers of a science curriculum, which implements many of the same innovations and principles as

recommended in the *Standards*. The 18-week curriculum had been piloted five years previous to this study and had undergone several cycles of refinements by the teachers in collaboration with university-based researchers.

The Traver Creek Curriculum addressed a substantial number of the *Standards* explicated as being important for high school students' conceptual understandings. These forty-five *Standards* were distributed across the Physical, Biological, Earth and Environmental Sciences (Appendix B). One critique of the curriculum might be that too many standards are being addressed in the 18-week semester, with little hope of students actually mastering that much conceptual content. This is certainly a valid concern, given an emphasis in science education of "less is more" and recommendations to move away from encyclopedia textbooks. However, in reply to this concern, it is important to acknowledge that the Traver Creek curriculum is but the first project in the three year Foundations of Science Curriculum. If the Creek project is the students' only encounter with the 45 conceptual standards, then indeed, too many standards are being addressed with an expectation for students to achieve mastery in all of them. On the other hand, if students repeatedly encounter the same conceptual standards in different contexts throughout the three-year curriculum, then the burden of achievement is shifted from a single project to the entire FOS program. The latter situation is the case for many of the 45 standards addressed in the creek project. For example, many of the chemistry standards in Standard B form the core of two projects in the third year of the program. Therefore, further studies of students' achievement after three years in the FOS program are recommended to fully address this concern.

Ecological Understandings

Over the duration of the Traver Creek project, students demonstrated deeper and more elaborate conceptual understandings about the creek ecosystem. This growth in understanding was evident in both the analysis of the artifacts and especially in the evaluation of the post-tests against achievement on the pretests.

Across the collection of student artifacts, students' understandings evolved from lists of unconnected observations found in their essays to detailed cause and effect relationships in their second set of models. For example, in the second set models, students created relationships showing factors outside of the creek (like urban development) affecting factors inside the creek (nutrients, turbidity or dissolved oxygen), which in turn affect the presence specific taxa of benthic macroinvertebrates and also the creek's overall water quality (e.g. Annie and Mackenzie's model in Figure 4.10). However, even though students are building more complex models, they often failed to fill in the explanation fields so the determination of their levels of understandings was hindered by the lack of information. In concert with the growth of understandings demonstrated across the four analyzed artifacts, students showed significant growth in understandings in each of the four main content areas of this ecology project when measured on the pre-post instrument.

This finding of growth in understandings is congruent with that of other researchers. For example, Fellows (1994) study, of sixth-graders and their writing, found that these students (a) added new principles or theories to their conceptual schema, (b) organized their schema around more central concepts, and (c) moved closer to scientific understandings. Indeed, one would be surprised if students did not learn content during a

program of study. The present study characterizes demonstrated levels of understanding and shows the growth in the number of connections that students make. This study does not make any claims that students' understandings have been optimized in this instructional environment. To do the later would require much additional analysis of the classroom observations. This study also does not make comparisons between the ninth grade FOS students and ninth graders in other learning environments.

Implications for Instruction: Modeling and Feedback

In this study, it was thought that the features of Project-based Science that supported the Traver Creek curriculum would promote understanding in the students. In particular, the creation of artifacts such as the stream reports and computer models, which demonstrate the connections students make between concepts, and reveal what they understand about that subject matter, were thought to both promote and reveal students' understandings (Gardner, 1991; Lehrer, 1993; Wisnudel-Spitulnik, Stratford, Krajcik, & Soloway, 1996).

However, as was illustrated in the previous chapter, students were able to demonstrate only limited numbers of understandings in the different artifacts that they created. Although the class set of artifacts demonstrated understandings across the range of expected *Standards*, individual artifacts revealed understandings on just a fraction of the expected *Standards* in all cases except for a few stream reports. Similar lack of depth in demonstrated understandings have been noted by other researchers.

As was discussed in Chapter 2 (pp 25-26) conceptual understandings do not occur in isolation in the domain of science. Conceptual understandings influence and are in turn influenced by epistemological and strategic understandings. In their paper on

“conceptual ecologies”, Posner, et al., (1982) discussed the difficulties inherent in achieving student conceptual change when the epistemological commitments of the student differ from those of the scientific community. For instance, if students are operating within an epistemology of “school” rather than science, then their epistemological commitments about the nature of evidence, the importance of parsimony in a theory, and metaphysical beliefs such as faith in the orderliness of nature influence their willingness as learners to give up their naive conceptions in favor of more scientific views (Posner, et al., 1982).

Gallagher (2000) believes that students' inability to understand and apply new scientific information presented in class lies in four related phenomenon: (1.) It may not be evident to students that the assessments, examinations and tests they take demand anything beyond factual recall. (2.) Students have not been taught how to, nor that they should, make sense of new information. (3.) Students are not commonly taught how to, nor that they should, make connections between new information and information that they have previously learned in order to develop a deeper understanding of the subject matter. And (4.) application of scientific knowledge tends to be given little importance in science classes and on science tests used to determine grades.

Gallagher's (2000) list appears to be reflected in the early student work analyzed in this study. As has been noted, students tended to relate a list of facts or unconnected observations in their essays on the stream. In the stream reports, students were able to build some connections in the introductions to their reports but were not able to integrate their own data into their discussion and conclusion sections. In contrast, the modeling artifacts, by virtue the features provided by the software, appeared to facilitate the

building of connections. That students learned to build connections from the modeling experience was evident in the post-tests, especially in the constructed response questions where students' responses were much more elaborate, included "cause and effect" statements and even incorporated graphical representations. Thus, in the Foundations of Science curriculum, where projects like the Traver Creek curriculum are situated in community contexts, students appear to develop rich connections.

Items two and three from Gallagher (2000) have particular implications for instruction in science, even in the Foundations of Science program. For many students, learning is equated with memorizing. That learning involves personal sense-making and reconciling personal constructions with accepted canonical knowledge is an unfamiliar idea for most students (Gallagher, 2000; Wiggins, 1991) but something that the FOS teachers strive to incorporate in their projects. Furthermore, making connections among ideas is not a spontaneous response for many students. Unless connections are made evident to students, many will not identify them. However, both making-sense and making-connections are essential elements in forming understanding, which teachers have always expected students would do as part of their study. The features of Project-based Science (Blumenfeld, et al., 1991; Krajcik et al, 2000, Marx, et. al., 1997; Krajcik, et al. 1999) that the teachers incorporated into Foundations of Science, i.e. driving questions, inquiry activities and artifact creation, help support and make more explicit the "making connections" part of learning science.

One implication from this study is for teachers to take some of the inferred expectations of understanding and make them more explicit to students in their classrooms through a variety of approaches including modeling sense-making behaviors,

direct instruction and providing high-end examples of student artifacts. For example, before the essay assignment early in the project, the teachers could provide students with some essays from previous years and have students discussed which information included in the essays would be useful for their own study of the creek and what information was lacking that they would want to know. In addition, the whole class could critique an example essay and provided feedback to its author before doing the individual peer critiques. Such analyses could help students improve their own writing. In addition, high quality essays from previous years could be held up as models for the students in the FOS I class.

Likewise, students' demonstrations of understandings in their reports and models could also benefit from more scaffolding of the artifacts. In terms of the reports, good writing begins with a careful definition and interpretation of the problem, defining the problem first in terms of aim or purpose, and of audience (Maimon, et al., 1981). To move the concept of audience from teacher to the community, a representative of the watershed council could meet with the students and share how their study fits into the scope of the larger community. In addition, "high-end" examples of reports and class critiques could help students improve the writing of their team reports.

The determinations of understanding in the models were limited by the information that students provided (or failed to provide). Many explanation fields were left empty so inferences about actual understandings were limited. Here again, providing more support for student artifact development through the use of examples and classroom critiques could improve the quality of understandings demonstrated in the artifacts. In the case of the Model artifacts, students did participate in a classroom critique of a model

just before the second modeling cycle. Another researcher created this model by combining and illustrating common problems like circular feedback loops (Metcalf, 1999). This classroom critique seemed to improve the “workability” of the models (Metcalf, 1999).

Unfortunately, in the “typical” classroom many students do not accept the view that understanding is what is expected. Instead, they see memorization as the goal. Classrooms attempting to implement the *Standards* need to change students’ perceptions of expectations.

Several studies have shown that the majority of class-time is devoted to helping students acquire information deemed essential as part of the knowledge base of science (Treagust, Jacobowitz, Gallagher, & Parker, 2001; Mitman, Mergendoller, Marchman, & Packer, 1987; Gallagher & Tobin, 1987; Tobin & Gallagher, 1987). Often, only a small amount of class time is devoted to helping students make sense of the new information and make connections among the various pieces of scientific knowledge in a way that leads to understanding. Moreover, even less time is devoted to helping students learn how to apply the abstract, conceptual knowledge of science to the world they experience outside of school. There is comparatively little information about development of epistemological understandings in the classroom.

Although beyond the scope of the present study, the data set developed in this study could be used to examine the teachers’ roles in promoting understandings throughout the Traver Creek curriculum.

Implications for Instruction: Engagement in Inquiry

The types of inquiry based learning environments called for by the *Standards* are

thought to promote robust understandings, the difficulty in creating these environments and in characterizing the types of understandings these environments are thought to develop (Blumenfeld, et al., 1997) means we do not have sufficient evidence to support a claim of robust understanding. Exactly what students learn, in terms of depth and breadth of scientific understandings and proficiency in the *Standards*, has yet to be established. This lack of evidence has been identified as a major gap in the research base of the science reform movement (Anderson & Helms, 2001).

The difficulty in creating inquiry –based learning was also experienced in this study. Although the curriculum for the Traver Creek project included student research projects, these projects were dropped due to classroom time constraints. As a result, the students in this study did not have the full range of opportunities to develop their strategic understandings in a way that would interact with both their conceptual and epistemological scientific understandings.

The students in this study were able to experience some inquiry types of activities, especially in their model building. Stratford (1996), Penner, et al., (1997), and Finkel and Stewart (1994), all found that as students engage in model building, they build conceptual understandings and that students tend to construct more in-depth conceptual understanding when building more complex models. Stratford (1996) related the model building processes to the conceptual richness of their final models. He suggested that students who employed more strategies (analysis, relational reasoning, synthesis, testing, debugging, planning, explaining, questioning, and searching) produced models that were conceptually richer than students who used fewer strategies. In the present study, it was the model building that appeared to have the greatest effect on student understandings, in

particular the connection of one observation to another in cause and effect relationships.

Student engagement in authentic scientific inquiry is an important classroom activity according to the *Standards*. Students are more inquisitive about the things they know about than the things they do not. Therefore students who understand some science are also likely to ask questions about how things happen, what causes them to happen, what processes are involved, how people have come to give such explanations, and what evidence supports the explanations. Scientific understanding, particularly in the conceptual and strategic dimensions, provides learners with the conceptual and technical (or procedural) tools, which produce the power and precision that would not otherwise be possible (Lee & Brophy, 1996).

In his essay on the *Structure of the Disciplines*, Schwab (1964) described the interactions of the conceptual and strategic dimensions. Schwab (1964) contended that our conceptual structure, especially our ignorance, leads us to formulate questions, which, in turn, lead us to investigate and collect data. Those same conceptual structures influence the way data are interpreted. He goes on to explain the reverse interplay of strategic understandings on conceptual understandings and the recursive nature of the feedback between the two dimensions.

In the Traver Creek project, if time constraints limit the inclusion of both modeling activities and student investigations, then perhaps another approach to creating the computer models could be employed. Instead of creating two discrete models, students could work on an evolving model over the course of the semester. In this scenario, each time students visit the creek to collect data; they would be expected to revise their models based on their new analyses. The assignments for each revision could

be made explicit based on the data being collected. For example, after the first observation walk the led to the essays; students have an idea of the land uses around the creek. They could pick a single land use they observed (e.g. parking lots) and create a model showing their hypothesis of how a change in that level of land use would affect the over all health of the creek. On each subsequent visit for the physical, biological and chemical assessments, students would collect additional data and could revise their models appropriately. Such a scenario would more closely mimic the use of modeling in the scientific community where models are refined and revised as new information becomes available.

Implications for Assessment

As noted in the early chapters, many studies of student understandings employ an “objective” instrument such as normative or standardized assessments or other *a priori* instrumentation. Often these instruments are used as a pre/post test type of measurement and are constrained by psychometric standards of validity and reliability. In constructivist learning environments, this method of using pre-constructed pre- and post-tests may not adequately capture the range of developing understandings that occur when students pursue different lines of inquiry and/or create a variety of artifacts, especially if students demonstrate deeper understandings in certain parts of the curriculum or draw unanticipated links to other content. In contrast, this dissertation study employed both *a priori* instrumentation and alternative modes of assessment grounded in student work to examine the breadth and depth of student understandings of a complex system, the ecosystem of a creek. The student work included four substantial artifacts: essays, scientific reports, and two dynamic computer models.

There is nothing particularly new or innovative about having students construct artifacts in the classroom. Students have done extended research or lab write-ups, term papers, and other projects to fulfill class requirements for as long as there has been schooling (Madaus & Tan, 1993). Now these approaches are being extended beyond individual classrooms to provide useful tools for conducting educational research and to pose a challenge to traditional ways of mass testing (e.g. Maeroff, 1991; Mehrens, 1992).

Science fairs and art contests are project type activities with a long history in schools (Fitzpatrick & Morrison, 1971). They are frequently well organized and carefully judged. Historically, the evaluation of the products was made difficult by the fact that each student was usually doing a quite different project. Not only was it hard to compare the relative merits of the art or science projects in general, but it was difficult to decide what the general bases for evaluation should be (Fitzpatrick & Morrison, 1971)

The development of scoring rubrics has made the evaluation of student artifacts more reliable and efficient. The rubric is a standard of performance for a defined population (NRC, 1996, Chap 5). Typically, scoring rubrics are developed by the teachers for the students in the class based on the criteria explicated in the assessment (see Table 5.1). The steps in designing a scoring rubric involve defining the performance standard for the scientifically literate adult and then deciding which elements of that standard are appropriate for the particular class of students (NRC, 1996, Chap 5). The draft performance standard is refined by subsequent use in evaluating student performance and work. Finally, student performances with respect to the rubric are differentiated. Performances are rated satisfactory, exemplary, or inadequate. Differences in opinions about the rubric and judgments about the quality of students' responses are moderated by

a group of teachers until consensus is reached for the rubric (NRC, 1996, Chap 5).

Table 5.1 Examples of a General Scoring Rubrics and the Coding Scheme used in this study

General scoring rubric could be applied to assess a variety of science learning situations (Price & Hine, 1994).	Coding Scheme used in this study, adapted from Stratford (1996)
<p>Outstanding: All the criteria are met, and the product or assessment exceeds the assigned task and contains additional, unexpected, or outstanding features.</p> <p>Good: The product or assessment completely or substantially meets the criteria</p> <p>Fair: The product or assessment meets some of the criteria and does not contain gross errors or crucial omissions.</p> <p>Inadequate: The product does not satisfy a significant number of the criteria, does not accomplish what was asked, contains errors, or is of poor quality.</p> <p>Poor: The student did not do the task, did not complete the assignment, or shows no comprehension of the activity.</p>	<p>Level 3: Explanation is scientifically correct to the level used in the <i>National Science Education Standards</i> (NRC, 1996) and contains no extraneous or incorrect ideas, statements concur with expert propositions (proficient or mastery level)</p> <p>Level 2: Explanation is partially correct but contains some extraneous and/or incorrect information (developing).</p> <p>Level 1: Explanation contains substantial errors or fundamental differences between the students' and expert's conceptions as depicted in the <i>Standards</i> (non-scientific or novice).</p> <p>Level 0: Student did not provide an explanation. OR, if some response is given, it does not evidence understanding, perhaps nonsensical (no evidence)</p>

The coding scheme employed in this study followed the basic design principles called for by the *Standards* document (Chap 5). However, instead of evaluating the overall artifact according to a holistic scoring rubric, each instance of conceptual understanding was dissected out and evaluated (Table 5.1 and Appendices E & G). Furthermore, instead of examining the standards for the “scientifically literate adult

(NRC, 1996)” and then deciding which elements of that standard are appropriate for the ninth grade students, this study used the standards for the “scientifically literate adult “ as the “Level 3” code (Table 5.1). As a consequence, the coding scheme used to evaluate student understandings in this study was often more severe than the teachers’ evaluations.

Throughout the semester, the ninth grade students in this study produced a plethora of artifacts assigned by their teachers. Some of these required relatively little student effort such as homework assignments, summaries of reading assignments, quizzes, activity handouts, etc. and potentially represent more knowledge recall than understanding. The classroom teachers using assorted rubrics and scoring schemes scored all of these artifacts. However, because the coding scheme employed in this study required detailed analysis along each *Standard* of conceptual understanding expected (and unexpected) in each artifact, the anticipation of knowledge vs. understanding was judged not to justify the cost of scoring these smaller artifacts in terms of time and effort. Instead, purposeful sampling was used to select which artifacts were subjected to the detailed analysis, in particular those artifacts such as the written reports and computer models, which required a greater time investment, were constructed over a number of days, and, ideally, underwent cycles of revision.

Thus, the detailed analysis employed in this study is not recommended for general classroom use. However, a detailed analysis using the *Standards* could be adapted and employed for specific artifacts or student projects. As in this study, expected outcomes can be relatively easily mapped onto the standards and a scoring matrix created for the relevant standards. The problem with *a priori* mapping of *Standards* is that it will probably miss “unexpected” demonstrations of understanding, unless the educator doing

the scoring is knowledgeable about the *Standards* and able to recognize these occurrences and map and score them appropriately.

The *National Science Education Standards* recognize that developing scoring rubrics requires highly informed teachers experienced in the process (NRC, 1996, Chap 5). The content standards call for knowledge with understanding. Considerable resources must therefore be devoted to preparing teachers and others in the science education system to design and rate assessments that require students to display understanding (NRC, 1996, chap 5), such as described in this study.

Implications for Additional Research

The above discussions point to additional lines of research. For example, although beyond the scope of the present study, one line of provocative research would be to examine the teachers' (Ms L. and Ms K) roles in promoting understandings throughout the Traver Creek curriculum. The present study confined its analysis to the assessment pieces of the curriculum. However, the data set upon which this study was based includes extensive field notes and video observations on the classroom instruction. This data can be examined for instances during instruction where teachers encouraged students to extend or elaborate upon their understandings. Crawford (2000), in her case study of a high school teacher, has identified multiple roles that a teacher might assume to promote understanding: motivator, diagnostician, guide, innovator, experimenter, researcher, modeler, mentor, collaborator, and learner. These categories may provide useful lenses to examine the teaching of Ms L. and Ms K.

In addition, the implied need for additional scaffolding in the students' creation of artifacts opens the door for quasi-experimental designs where the level of scaffolding is

varied while controlling the other variables of instruction. Such research would further illuminate the effectiveness of different forms of scaffoldings.

This study also limited its analysis to the dimension of conceptual understandings. It would be interesting to explore students' demonstrations of strategic and epistemological understandings in their artifacts and to look for patterns what would indicate interactions between the understanding frames.

Because this study characterizes students' understandings, in reference to the *Standards*, as they embark on a three-year program of integrated, project-based science, it provides a foundation for additional research. Interesting questions for follow-up study are, "How persistent are the conceptual understandings developed during the creek study?" "Do student invoke these understandings to make sense of science content in subsequent projects?" and "How do understandings demonstrated in different projects, but mapping onto the same standards (near transfer) compare to the understandings demonstrated in the artifacts examined here?" The latter question would shed some light onto the critiques about the number of Standards addressed in the FOS creek curriculum.

Also, as this study uses the NSES as a frame of reference and language, comparisons of understandings in reference to the *Standards* could be conducted across time and space. The usefulness of the *Standards* as a tool will be explored in the next section.

Summary of Learning

This study showed that over the course of the semester, students demonstrated deeper and more elaborate conceptual understandings about the creek ecosystem.

However, the analysis employed here was limited to conceptual understandings and it has

been noted how the strategic and epistemological dimensions of understanding may interact with conceptual understandings. Acquiring scientific knowledge about how the world works does not necessarily lead to an understanding of how science itself works, nor does knowledge of the philosophy and sociology of science alone lead to a scientific understanding of the world (AAAS, 1993 p. 4). The challenge for educators is to weave these different aspects of science together so that they reinforce one another. Learners need to use strategic understandings of inquiry and problem solving to develop conceptual understandings. They need to use their conceptual understanding when engaging in inquiry and problem solving and their epistemological understandings to make sense of their findings. Their epistemological understandings inform their choice of strategies when involved in investigations and as they structure their conceptual understandings.

Teachers in classrooms attempting to enact the *Standards* need to be aware of students' tendencies to employ "school" epistemologies. Keeping these tendencies in mind, teachers can help students develop scientific understandings as students carry out inquiry activities and develop artifacts that represent their understandings. Artifacts have the potential to both promote and assess students' understandings if they are employed in a thoughtful and purposeful way in the classroom. The usefulness of the *Standards* as a frame of reference for assessing these artifacts is addressed in the next section.

Using the National Science Education Standards as a assessment tool

Comparing student understandings on a pre-post objective instrument is relatively unproblematic. When there is a direct match between the two assessments, changes or gain scores can be statistically analyzed. Analyzing students' understandings

demonstrated in laboratory reports and other artifacts such as dynamic computer models are more difficult. First, because we need to define what understandings are being assessed in the artifact. And second, if we want to show that learning has occurred, we either need to repeat the assessment or we need to devise a tool by which similar understandings in different artifacts can be traced.

The modes of learning called for in the *National Science Education Standards* (NRC 1996) imply markedly different roles and tasks for students in terms of designing, interpreting, explaining, and hypothesizing in science classroom (Anderson & Helms, 2001). When tasks vary across students as they do when open-ended artifacts are produced, the assessment of student understandings becomes problematic. Comparing understanding demonstrated in laboratory reports and dynamic computer models is much like “comparing apples and oranges.” Apples and oranges are both fruits (products, artifacts) of plants. While they may share some superficial features (e.g. an approximately spherical shape), meaningful comparisons require the applications of more abstract standards, e.g., sugar content, moisture content, or the recommended daily allowance of essential vitamins and minerals.

The students in this study encountered a similar problem. They studied a local creek by conducting a variety of water quality tests, collecting benthic macro-invertebrates and making a series of observations about the physical environment. In this effort, they generated a lot of data, but the numbers and observations did not have meaning beyond a description of the creek. The numbers alone did not tell the students about the water quality. So, the students in this study used “standards of water quality (Mitchell & Stapp, 1994)” in their determination of the health of the creek. These water

quality standards allowed the students to make comparisons between specific parameters measured in their creek and scientifically defined values. Students were also able to make comparisons between the different kinds of assessments, e.g. chemical testing and bioassays, to see if the different forms of assessment led to similar results. By using standards, the students were able to make an assessment of the creek without having to find a comparable creek or without studying values determined before and after an intervention. This idea of comparing observations to defined standards may also be useful for comparing understandings demonstrated in different types of student work.

Comparing observations to defined standards may also be useful for examining understandings demonstrated in laboratory reports and dynamic computer models. Comparing understandings demonstrated in laboratory reports and dynamic computer models requires the application of standards, in this study, the *National Science Education Standards* (NRC, 1996). The *Standards* claim to be criteria by which to judge quality: the quality of what students know and are able to do and criteria to judge progress toward a national vision of learning and teaching science (NRC, 1996, Ch 5). The usefulness of this tool for describing student achievement across multiple tasks and contexts depends on how well it meets the measurement criteria of validity, reliability, sensitivity and the impact of the assessment on instruction and classroom practices (See Champagne & Newell, 1992; Haney & Madaus, 1989; Kulm & Malcolm, 1991; Malcolm, 1991; Wiggins, 1993)

The methods and techniques of measuring learning that we know today represent the products of earlier eras. Over the last two centuries the predominant methods of assessing student achievement - from the oral mode, the written essay, the short answer

form, the multiple choice format, the machine-scorable answer sheet, and finally to computer-adaptive testing - have all been geared toward increasing efficiency and making the assessment systems more manageable, standardized, easy to administer, objective, reliable, comparable, and inexpensive (Madaus, 1994). However, innovative procedures and situations believed to assess high levels of competence and reasoning abilities, such as artifact assessment, have been introduced and are advocated by educational researchers and reformers as being more authentic methods of assessment (e.g. Papert, 1991; Perkins, 1992; Wiggins, 1989b; Wiggins, 1993). Critics claim that such alternative assessments have too little correspondence to national and state norms that they can be too subjective and are too inconsistent. Thus, much of the work using alternative assessments is experimental and not without controversy centering upon unresolved issues regarding the assessment of student understandings.

The fundamental concepts of reliability and validity are the key established psychometric criteria for judging the technical adequacy of measures. Psychometricians have argued the issues of validity and reliability extensively, (e.g. see Linn, et al., 1991; Messick, 1989, 1994; Burger & Burger, 1994). The Burgers (1994) believe that the alternative assessments must be held to the same stringent standards of reliability and validity as those achieved by standardized norm-referenced assessments. In an opposing position, Moss (1994) argues that current conceptions of reliability and validity in educational measurement constrain the kinds of assessment practices that are likely to find favor, and these in turn constrain educational opportunities for teachers and students.

Linn, Baker and Dunbar (1991) claim that many issues concerning the evaluation of the new forms of assessment being developed have not been sufficiently addressed.

They argue that to develop technically sound quality assessments we must first address certain criteria for evaluating such assessments and that these criteria need to be expanded beyond those traditionally associated with validity and reliability to more adequately reflect theoretical concepts. They propose a set of eight criteria that “serious validation” of alternative assessments needs to include: intended and unintended consequences of measurement use, fairness, transfer and generalizability, cognitive complexity, content quality, content coverage, meaningfulness, and cost and efficiency.

In terms of evaluating the effectiveness of the *National Science Education Standards* (NRC, 1996) as an assessment tool, I have chosen to focus on four of these criteria: validity, reliability, sensitivity and impact on instruction. These criteria are discussed in the next sections, starting with validity.

Validity

Validity has to do with the meaning of the measurement and the inferences that can be made about the results of that measurement (Smith & Glass, 1987, p. 106). Traditional notions of validity require that an instrument measure what it claims to measure. A newer notion of validity is that a valid instrument measures that which is valued (Champagne & Newell, 1992). With an agenda of “scientific literacy for all” and for learners who can go beyond the information given to apply knowledge wisely, fluently, and flexibly in interactions with novel experiences, it is important that we develop and use assessment tools that can reveal whether or not we are achieving our objectives. As the *Standards* become an accepted goal for science instruction, using the *Standards* as an assessment tool meets the criteria of measuring what is valued.

Content validity addresses the question, “Does the measure adequately sample the

range of subject matter or objectives that it is supposed to assess?" These criteria apply to the artifacts or tasks chosen for assessment rather than the standards by which the performance is assessed. A valid assessment system provides information about the particular tasks on which students succeed or fail. Science assessments purporting to measure scientific understanding must require students to apply scientific knowledge and reasoning in real world contexts as well as in situations that approximate how scientists do their work (NRC, 1996). According to the *National Science Education Standards*, the content that we wish to assess extends beyond the ability to recall a large number of scientific facts to learning science where students describe objects and events, ask questions, acquire knowledge, construct explanations of natural phenomena, test those explanations in many different ways, and communicate their ideas to others (Chap 2, NRC, 1996).

Mapping the *Standards* onto the measures used to assess understanding (as was done in step two of the analysis in this study) is one way to establish this type of content validity. The mapping analysis showed that the four major artifacts provided opportunities for students to demonstrate understandings on 23 conceptual standards while the pre/post-instrument only assessed understandings on 17 standards (Table 4.2). In this study, the four artifacts offered the potential for a more complete assessment of understandings across the curriculum than the pre/post tests (51% vs. 38% of the curriculum).

A large part of the content validity argument is based on the assumed relationship between instruction and assessment (Miller & Legg, 1993). However, this relationship needs to be established with empirical evidence, as in this study, rather than assumed.

Judgments regarding the cognitive complexity of an assessment need to start with an analysis of the task, student familiarity with the problems and ways in which students attempt to solve these problems (Linn, et al., 1991). For example, having students determine which brand of paper towels is most absorbent, one of the performance assessments devised by Shavelson and Baxter (Shavelson & Baxter, 1992a; 1992b; Shavelson, et al., 1992), may be an authentic assessment of a student's ability to do science. However, it could be an inauthentic measure if the student had been taught "the method" for determining absorbency and then followed rote, ritualized procedure. We might get a performance, but do the students exhibit real understanding? Hence, an important aspect of assessing student scientific understandings is to document the instruction that occurs, including the content or conceptual material, the epistemological framework, the types of inquiry engaged in and the opportunities and experiences made available to students for constructing their understanding (Miller & Legg, 1993). This type of documentation occurred in the first part of my analysis when I mapped the curriculum and classroom observations onto the *Standards*. When mapped onto the *Standards*, the Traver Creek curriculum was found to address forty-five (45) conceptual understanding standards at the high school level, half of the content explicated by the *Standards* as being important for students to understand (Table 4.1 and Appendix B).

Relatively few of the proponents of alternative approaches to assessment have addressed the question of content validity when evaluating their measures. Simply because the measures are derived from actual performance or relatively high-fidelity simulations of performance, it is too often assumed that they are more valid than multiple-choice tests (Linn, et al., 1991). Linn and his colleagues (1991) point out that

we cannot assume that a hands-on scientific task encourages the demonstration of problem solving skills, reasoning ability, or more sophisticated mental models of the scientific phenomenon.

In the Traver Creek project, the four artifacts that were selected for analysis do meet the criteria for content validity as established by classroom observations and task analysis. Each of these four artifacts was constructed over a number of classroom days and often underwent cycles of revision. These were the assessments where questions and tasks could be discussed, clarified, and even appropriately modified through discussion with the teacher and/or the student's peers. The expected outcomes of the four major artifacts assessments map onto the standards addressed by the creek curriculum and provided opportunities for students to demonstrate understandings on 23 of the 45 conceptual standards addressed by the curriculum

Another validity criterion is mode validity, a measure of whether the assessment adequately samples the range of ways of expressing or demonstrating understanding. Assessments limited to a single way of expressing or demonstrating understanding, e.g., logico-mathematical expressions, can be characterized as having low mode validity (White & Gunstone, 1992). Goldberg (1992) argues that while logico-mathematical thinking is certainly an important aspect of teaching and learning and essential to functioning in western society, there are many forms of expression that remain wholly untapped in terms of assessing students' understandings. Similarly, Gardner's theory of Multiple Intelligences (Gardner, 1983) demands a wider horizon if we are to honor learners' varied abilities. This would include finding ways to make music, the visual arts, dance and sports, interpersonal skills, and skills of self-reflection more substantive and

salient in classrooms, curricula and authentic assessment strategies. Increasing mode validity by allowing different forms of representation is useful because certain ideas lend themselves better to different modes, and because students may, more easily and readily, demonstrate what they know in different ways. For example, the student who is unable to explain a concept such as the water cycle may be able to draw and label a picture to show that significant understanding exists. Conversely, a student who can describe a process, such as separating a mixture, may not be able to actually carry out the process during a performance assessment (Schneider, et al. 2002).

One strength of alternative assessments, such as using artifacts, is the diversity of the methods used to assess students' understandings. The balance between words and diagram representations shifts between the techniques of multiple choice and essay tests to concept mapping, modeling and drawings. Different methods also vary in the degree to which they limit students' responses (White & Gunstone, 1992): multiple-choice tests are an extreme form of limitation with only one pre-selected correct response. This constraint was made very evident in this study when attempting to classify student understandings demonstrated on the pre/post-test instrument on a four-point (0-3) scale. In the end, the multiple-choice items were classified as Level 3 (proficient) if correct or Level 0 if incorrect. Short answer questions were a little more open, essays even more so and each present their own unique challenges to assessing students' understanding. Closed and open, logico-mathematical and creative assessments are all valuable, and we need to appreciate the diversity of methods and the different aspects of understanding that they tap (White & Gunstone, 1992).

Artifacts actually lend themselves to the use of multiple modes of expression. For

instance, authoring with images and sounds, as well as with words, opens new opportunities for learners in their "writing" (Daiute, 1992). Recent developments in technology make it possible to integrate images, sounds, and text in multimedia computer environments that also provide tools for transforming these various symbol systems. The artifacts that students created in this study include the more traditional text based school artifacts such as essays and reports. But the reports, generated on the school's laptop computers, also allowed for integration of charts and drawings. The computer models, generated with Model-It 3.0b (Jackson, Stratford, Krajcik & Soloway, 1996a, 1996b) allowed even more integration of graphics and text. Students had the options of using graphics from a library of images of the creek and its organisms or to generate their own graphics in the drawing and paint programs on their computers. Being able to "read" and "write" with several media in an integrated workspace can help learners ground their literacy and learning in familiar contexts (Daiute, 1992). Hypermedia artifacts appear to facilitate the construction and integration of multiple representations of the understanding of science concepts (Spitulnik, 1995; Spitulnik, et al., 1996).

In addressing the criterion of assessment validity, the combinations of artifact assessment and the *Standards* meet the requirements: assessment of the valued, content validity and mode validity. The criterion of assessment reliability is addressed in the next section.

Reliability

Reliability is the consistency of the judgment that follows from the use of a measure. In a traditional reliable instrument, similar instruments administered to the same child over time will result in approximately the same score, the instrument has a

high internal consistency, and all scorers of the instrument produce the same score. However, this standard of assessment reliability can be in conflict with other valued outcomes of educational assessment.

Some of the reliability issues in assessment are related to the perceived purpose of assessment. In the traditional use of reliability, an assessment instrument is simply a tool, like a thermometer, which can be used to obtain some measurement value but with negligible impact on the phenomena being measured, in which case, we would expect high reliability. In the world of science education described by the *Standards* (NRC, 1996), there are multiple purposes to assessment, including causing students to think, to make new links, to ask questions, and to build understandings. In traditional reliable assessment practices, similar assessments administered to the same child over time will result in about the same score. But if thinking processes are valued and understanding is conceived as a dynamic process, a learning child should think differently on the second assessment (Champagne & Newell, 1992). Such an assessment would seem unreliable by traditional definitions but the outcome is one that is valued by our defined target of the *Standards*.

According to Perkins (1992), an assessment should be as much a learning experience as a testing experience. Assessments should stretch the learner even as they create an occasion for a learner to display mastery and understanding. Inherently, they test for, and therefore press for, transfer and understanding (Perkins, 1992). Consequently, we would neither expect, nor value, high reliability as traditionally defined when evaluating assessments of scientific understanding.

A second reliability issue has to do with internal reliability. Internal reliability is

whether or not performance on different portions of an assessment leads to the same conclusion. This is an issue if we try to apply internal reliability when considering portfolios of artifacts or student artifacts that may have multiple representations, or themselves multiply represent student understandings. This issue was confronted in the present study where several instances of student understandings mapped onto the same standards. Often, the student explanations were coded as demonstrating different levels of understanding on the same content standard. A resolution was sought by taking the median value of understanding (ordinal scale) when multiple instances were found in one measurement. Moss (1994) argues that hermeneutic approaches to assessment can allow students substantial latitude in selecting the products by which they will be represented - a latitude that need not be constrained by concerns about quantitative measures of consistency across tasks.

Wiley and Haertel (1996) offer another means of addressing task reliability without the constraining assumption of homogeneity of tasks. As part of a comprehensive assessment development process, they suggest carefully analyzing assessment tasks to describe the capabilities required for performance, scoring tasks separately for the relevant capabilities, and examining reliability within capability across tasks to which the capability applies. While this approach supports the use of complex and authentic tasks that may naturally vary in terms of the capabilities elicited, it still requires detailed specification of measurement intents, performance records, and scoring criteria. It is this approach that was adapted in the present study, where the assessments (artifacts) employed were analyzed and mapped onto the *National Science Education Standards*.

A third reliability issue has to do with reliability in scoring, also known as reader or rater reliability. Inter-rater reliability is a historical issue. Standardized testing evolved and proliferated because school transcripts became unreliable (Wiggins, 1989a). An "A" in a subject meant only that some adult thought the student's work was excellent. However, without being tied to a defined target or standard, there is no possible way to determine what that "A" means in terms of knowledge and understanding.

In contrast to multiple-choice questions, which are easily scored, alternative assessments, such as evaluating artifacts, include subjective decisions in which rater reliability becomes an important issue. Raters who judge student performance must agree regarding what scores would be assigned to students' work within the limits of what experts call "measurement error." Do raters agree on how an assessment would be scored? Do they assign the same or nearly similar scores to a particular student's work? If the answer is no, then students scores are a measure of who does the scoring rather than the quality of the work (Herman & Winters, 1994). Inter-rater reliability can be improved by careful definition of the relevant information and the use of rubrics. In this study, inter-rater reliability had to be addressed by carefully defining how understandings may be represented in the different artifacts and by providing examples from student work of the different levels of understandings. By taking these steps, respectable levels of inter-rater reliability were achieved ($0.837 \leq r \leq 0.958$ or $.70 \leq r^2 \leq .92$).

Different forms of assessment vary in the ease with which they can be made reliable. Relational diagrams may have good reader reliability, but concept maps and interviews have greater subjectivity in their interpretation (White & Gunstone, 1992). Lower internal consistency or reader reliability of some methods does not make them less

valuable for probing student understandings. Instead, we should try to make each method as reliable as possible, by ensuring, for instance, that students are familiar with the form of the assessment and that raters are using standard procedures.

For the criteria of reliability then, the method of artifact assessments and the *Standards* used in this study rejects the traditional notion of reliability defined as consistency in the measure over time. However, the method offered here does strive to meet the criteria of internal or task reliability and scoring reliability through the careful analysis of task structure and the mapping of the assessments onto the *Standards*. The sensitivity criteria of this assessment method are addressed in the next section.

Sensitivity

Sensitivity of an assessment tool is an issue when there is a desire to track changes in understanding over time. The more incremental the changes, the more sensitive the tool need to be. As written, the *Standards* are a threshold with no sensitivity. Either students are achieving at the level of the *Standards* or they are not. In order to increase the sensitivity of the *Standards*, this study employed a four level coding scale based on the prior work of educational researchers (Carey, Evans, Honda, Jay, & Unger, 1989; Grosslight, Unger, Jay, & Smith, 1991; Songer & Linn, 1991; Stratford, 1996). However, in this four level scale with Level 0 representing no evidence, Level 1 a non-scientific understanding, Level 2 a partial understanding and Level 3 as proficient at the standards, the intervals between levels on this ordinal scale are grossly unequal. On any standard, a student might initially demonstrate a Level 2 or partial understanding. On each succeeding measure, they might show increasing understandings, but never reach the Proficient level. While they are learning and developing scientific understanding, the

assessment tool of the *Standards*, even as adapted with a quality scale in this study, were not sensitive enough to capture these changes.

Likewise, proficiency in the standards, Level 3, does not represent the highest levels of understandings, such as those achieved by experts in a domain. In a project-based classroom where students pursue different investigations and create different types of artifacts, students may achieve understandings on individual *Standards* far beyond those articulated in the document. Again, the form of *Standards* assessment used in this study would not capture those higher levels of achievement.

An alternative scale would be to assess students according to the grade levels differentiated in the standards document. Students' demonstrations of understanding could be characterized as below 4th grade, at 4th grade but not yet 8th grade, at 8th and not yet 12th grade or at and above 12th grade (Level 3 or proficient). Such a scale may also seem less abstract to parents and practitioners when discussing student achievement (e.g. Sally shows a 12th grade understanding about the interaction of earth systems (D.3.3) but only a 4th grade level about the properties of matter (B.2.1))

However, a grade-level scale brings into relief a second sensitivity issue, that of specificity of content across the standards. As noted in Chapter 4, although the Traver Creek curriculum was relatively balanced between earth science, biology, and chemistry, this balance was not evident when looking at the number of standards addressed. When mapped onto the *Standards*, the Traver Creek Curriculum appears heavily weighted toward the Life Sciences (19 standards in Standard C) and Environmental Sciences (11 standards in Standard F). This is due in part to how well the content is delineated among the different sub-standards in each section. For example, in the Life Sciences, standard

C.4 addresses the interdependence of organisms and there are five objectives that differentiate the kinds of interactions (Appendix B). All five sub-standards match content addressed in the creek project. At the same time in Standard D on Earth and Space Science, weather phenomena, which were addressed in the context of the impacts a watershed due to flooding, drought, etc, is not part of the 12th grade standards, but of the middle school standards (grades 5-8). Because this study looked at only achievement along the high school standards, understandings about weather and how weather affected the creek were mapped onto less specific standards of D.1.1 – Energy in earth systems, D.2.1 – Geochemical cycles, D.3.3 – Interactions of earth systems, and B.2.5 – States of Matter.

Similarly, there is content in the standards for grades 9-12, i.e. nuclear physics in Standard B.1 that is not part of the standards for the lower grades. If a grade level scale is adopted, for tracking understandings along the standards, there would probably be little difficulty rationalizing that an 8th grader has achieved a 12th grade understanding of radioactive isotopes (Standard B.1.4). However, claiming that a high school senior has only an 8th grade understanding of weather (because standard D.3.2²⁰ for grades 5-8 is the highest level at which weather is addressed in the *Standards*) may be problematic when communicating student achievement to communities of teachers, parents, admissions officers, and policy makers.

Clearly, tool sensitivity, in terms of scalar sensitivity and content specificity, is a

²⁰ D.3.2: Weather changes from day to day and over the seasons. Weather can be described by measurable quantities, such as temperature, wind direction and speed, and precipitation. (NRC, 1996, Chap 6, grades 5-8)

weakness in the criterion for the usefulness of the *Standard* as an assessment tool. To meet these criteria, the methods of using the *Standard* as an assessment tool need further development, including an internal scale by which evaluators can judge the quality of students' understandings.

The fourth major class of criteria that assessment methods need to meet is the impact of the assessment on instruction. This criterion is discussed in the next section.

Impact of the assessment on instruction and classroom practices

The final criterion of tool usefulness is its impact on instruction. The idea that teachers teach toward the test has become part of the conventional wisdom of education, but has its roots in research (see Kulm & Malcom, 1991; Wiggins, 1989b; Wiggins, 1993). The common pattern at the secondary level is for teachers to present the topic, test student to assign grades on the achievement pertaining to the content, and continuing on to the next topic (Treagust, Jacobowitz, Gallagher, & Parker, 2001). School assessments usually ask the learner to identify the products (discourse, things, performances) of others; for example, matching the benthic macro-invertebrates to their pollution tolerance taxa groups (Archbald & Newmann, 1988). In classrooms where the activity of answering recall questions plays a dominant role, this activity often becomes the basis for students' operational definitions of scientific understanding (Anderson & Roth, 1989). Students, who say that they "understand" a concept or topic, often mean that they are prepared to answer recall questions about it; in their experience, this is the sole or primary function of scientific knowledge (Anderson & Roth, 1989). The *Standards* also recognize this when they state:

Assessment and learning are so closely related that if all the outcomes are not assessed, teachers and students likely will redefine their expectations for learning science only to the outcomes that are assessed. (NRC, 1996, p. 82).

The modes of learning called for in the *Standards* imply markedly different roles and tasks for the students in terms of designing, interpreting, explaining, and hypothesizing. More research is needed of what roles students can play in varied science classroom contexts and the types of work they can produce (Anderson & Helms, 2001). There is also an ongoing need for research about the intended and unintended effects of assessments on the ways teachers and students spend their time and think about the goals of education (Linn, et al., 1991). It cannot just be assumed that a more “authentic” assessment will result in classroom activities that are more conducive to learning.

One of the limitations of the present study is that it examined student outcomes as demonstrated in the artifacts they create. This study did not analysis the classroom activities beyond the scope of mapping course content onto the *Standards*. However, the data set upon which this study was based includes extensive field notes and video observations on the classroom instruction. In further studies, this data can be examined in more detail for the impacts of artifact assessments on the instruction in the classroom.

In evaluating the usefulness of the *National Science Education Standards* (NRC, 1996) as an assessment tool, I have applied four criteria: validity, reliability, sensitivity and impact on assessment. In addressing the criterion of assessment validity, the combinations of artifact and the *Standards* meet the requirements of assessment of the valued, content validity and mode validity. For reliability, this method rejects the traditional notion of reliability defined as consistency in the measure over time. However, the method offered here does strive to meet the criteria of internal or task

reliability and scoring reliability through the careful analysis of task structure and the mapping of the assessments onto the *Standards*. Scalar sensitivity and content specificity are weaknesses in the usefulness of the *Standards* as an assessment tool. To meet these criteria, the *Standards* need further development, including an internal scale by which evaluators can judge the quality of students' understandings. The impact on instruction of applying the *Standards* needs further study, some of which can make use of the data gathered during the research described here.

Researchers and classroom teachers are not bound by traditional paper-and-pencil tests as we attempt to identify and evaluate students' scientific understandings. In fact, only the type of learning that we want to measure limits the type of assessment that we decide to use. Moreover, it is often appropriate to combine a variety of assessment methods when evaluating different aspects of student progress (Jones, 1994) in order to achieve a more complete profile of student understandings. Using student artifacts as a method of assessment allows us to evaluate students' structural and functional scientific understandings.

Final Thoughts

In essence, the assessment process of measuring student understandings recapitulates the scientific enterprise (Duschl & Gitomer, 1991). Assessment of student understandings is a sense-making activity that is grounded in student work. Assessors (learners, teachers, researchers, departments of education) make claims about student learning that need to be supported by data and warrants that are recognized as valid within a community. With the agenda of scientific literacy for all Americans and the desire to create a citizenry that has the ability to go beyond the information given and to

exercise knowledge wisely, fluently, and flexibly in interactions with novel experiences, it is important that we develop and use assessment tools that can reveal whether or not we are achieving our objectives.

Traditional science assessments tend to measure discrete isolated bits of knowledge, rather than rich and well-structured understandings called for by the current reform movements. Instead of checking whether students have memorized certain items of information, assessment needs to probe for students' understanding, reasoning, and utilization of knowledge (NRC, 1996). Having students construct representations of their understandings by creating cognitive artifacts is one way of doing this.

Krajcik and colleagues (1998) and Keys & Bryan (2001) have pointed out that much of the research on inquiry has been done in rich demonstration sites or in classes taught by the researchers with a focus at the elementary and middle school levels, thereby limiting our understanding of how inquiry teaching and learning look in an ordinary classroom taught by teachers. This study examined student learning in a public high school in a curriculum designed by the teachers in collaboration with university researchers. Furthermore, this innovative program had been in place for several years. Consequently, this study adds much needed data to our understanding of science learning and assessment in the high school classroom.

In this study, the *National Science Education Standards* (1996) were applied as an assessment standard to characterize high school scientific understandings developed during a creek project. While using this tool shows promise for capturing demonstrations of understanding across multiple representations and time, the tool needs further refinement to increase its sensitivity to changes in understanding. However, using the

Standards as an assessment tool can help increase awareness of the standards and what is valued in terms of scientific literacy in the 21st century.

APPENDICES

APPENDIX A
PARENTAL/STUDENT CONSENT LETTER

August, 1996

Dear Parent/Guardian,

Your child's science teachers, Liz Stern, Kathe Hetter, Madeline Hueble-Drake and Mike Mouradian, at Community High School, are involved in an collaborative effort with faculty members and researchers from the University of Michigan and the National Science Foundation to explore the use of technology to foster sustained inquiry in the science classroom. The aim of this effort is to develop experience using innovative materials and methods to help students and teachers realize the goals of investigative science. Your child will experience state-of-the-art teaching materials and methods, including the use of computer modeling and visualization tools and access to world wide web resources in their science classroom.

All ninth-eleventh grade students will participate in the Foundations of Science (FOS) sequence of science classes. In addition to the instruction we will be taking a closer look at how the students are learning science in this environment. We will be working in your child's science classroom and would like to include your child in our study. Our research methods will include making video and audio recordings of classroom and student activities. We will ask the students in the class who have permission to participate in the study to complete questionnaires. We would also like to video and audio tape specific groups of students as they are engaged in science activities. Some student volunteers will be asked to participate as *focus students* in the study and will be asked to complete short interviews about their activities in the class. In addition, we will be keeping copies or other records of the work the students do in their science classroom.

We ask your permission to allow your child to be involved in this research for the 1996-97 school year. We would like to stress that our research procedures are similar to what normally would happen in these science classrooms. In fact, your child should learn much from the experience and their participation will help us to improve science instruction. You may also withdraw you child from the study at any time by informing one of the researchers or your child's science teacher. Students who do not have permission to participate in the research will receive the same classroom instruction as the other students. We will keep students' names and identities confidential at all times.

Please call us if you have any questions or concerns about this science learning experience or the research project. We can be reached at the University of Michigan at

555-4227 (Valerie) or 555-0597 (Joe), or you may reach us through the science teachers at Community High School, 555-2021. [In the evenings you can reach Valerie at 555-1996 and Nathan at 555-0417.] If you have other concerns about the research, you may contact Dr. Brian Rowan, Associate Dean of Research, UM-School of Education, at 555-9470.

Thank you very much for your cooperation.

Sincerely,

Joe Krajcik	Elliot Soloway	Valerie Talsma	Nathan Bos
Associate Professor	Professor	Research Associate	Research Associate
555-0597	555-1562	555-4227	555-0417

[HSR. Enhancing the Teaching of Science Projects, 11/94] -

Permission Form:

For the Parent/Guardian

I understand that researchers from the University of Michigan are conducting a study of student experiences in technology and sustained inquiry in science at Community High School. My child has been asked to participate in this study as a member of a Foundations of Science class.

I understand that participation in the study may involve video and audio taping in the classroom, student questionnaires, short interviews, and records of student work.

I understand that this participation is entirely voluntary and that I may withdraw my child from the study at any time.

I further understand that no questions or procedures are involved which represent a threat to my child. The child's name will not be used in any subsequent report and their identity will be disguised by the study's authors.

Under the above conditions I give my permission for my child to participate in this study. My child also *has / does not have* (circle one) permission to participate as a "focus student."

Parent's/Guardian's name (please print)

Signature, Parent/Guardian

Date

FOR THE STUDENT

Under the above conditions, I agree to participate in the study of student experiences in sustained inquiry in FOS at Community High School. I *am / am not* (circle one) willing to participate as a "focus student."

Student's name (please print)

Student's signature

Date

***** OR *****

I do NOT give my permission for my child to take part in this research project on sustained inquiry in FOS. My child will participate in the class as a student only.

Parent's/Guardian's name (please print)

Signature, Parent/Guardian

Date

Please return a signed copy of this form to:

Liz Stern, Kathe Hetter, Madeline Hueble-Drake or Mike Mouradian, Science Department, Community High School, 401 N. Division, Ann Arbor, 48104.

APPENDIX B

MATRIX OF UNDERSTANDING STANDARDS

		KEY	scope	0.99	1.00	1.1	2.00	2.1	2.2	3
CONCEPTUAL UNDERSTANDING STANDARDS Content Standard B: Physical Science										
B.1. STRUCTURE OF ATOMS	Travel/Track Curr.	DTL	E/G	Pre test	Essa %	Report	Mode 11	Mode 12	Post test	
B.1.1 Matter is made of small particles called atoms, and atoms are composed of even smaller components. These components have mass, electric charge, and spin.	11/13 structure of atoms 11/13 electron shells 11/13 atomic models Bohr models		E							
B.1.2 The atom nucleus is composed of protons and neutrons, which are much more massive than electrons. While an electron has about 1/1836 the mass of a proton, it has the same charge.	chemical vs. physical changes 11/13 simple replacement reactions. - Spot plane lab		E							
B.1.3 The nuclear forces that hold the nucleus of an atom together, in nuclear fission, are usually stronger than the electric forces that would make it fly apart. Nuclear reactions convert mass into energy.			E							
B.1.4 Radioactive isotopes are unstable and undergo spontaneous nuclear reactions, emitting particles and/or gamma radiations. The decay of a radioactive isotope can be predicted.			E							
B.2. STRUCTURE AND PROPERTIES OF MATTER	Travel/Track Curr.	DTL	E/G	Pre test	Essa %	Report	Mode 11	Mode 12	Post test	
B.2.1 Atoms interact with one another by transferring or sharing electrons that are further from the nucleus. These outer electrons govern the chemical properties of the atom.	chemical vs. physical changes 11/13 simple replacement reactions. - Spot plane lab		E							
B.2.2 An element is composed of a single type of atom. Three elements are found in order according to the number of protons (called the atomic number), properties, and mass of the atoms.	11/13 basic periodic table of elements family properties - naming simple compounds		E							
B.2.3 Bonds between atoms are created when electrons are paired up by being transferred or shared. A substance composed of a single kind of atom is called an element. The atoms may be diatomic.	chemical vs. physical changes 11/13 simple replacement reactions. - Spot plane lab		E							
B.2.4 The physical properties of compounds reflect the nature of the interactions among its molecules. These interactions are determined by the structure of the molecule, including the shape, size, and mass.	11/13 - fact on physical and chemical changes, heterogeneous & homogeneous mixtures, stability and bond angles		E							
B.2.5 Solids, liquids, and gases differ in the distances and angles between molecules or atoms and therefore the energy that holds them together. In solids, the particles are packed closely together.	11/13 matter has space and mass, physical properties & chemical properties chemical (heat, gas, color & reactions) vs. physical changes		E							
B.2.6 Carbon atoms can bond to one another in chains, rings, and branching networks to form a variety of structures, including synthetic polymers, oils, and the large molecules of life.			E							
B.3. CHEMICAL REACTIONS	Travel/Track Curr.	DTL	E/G	Pre test	Essa %	Report	Mode 11	Mode 12	Post test	
B.3.1 Chemical reactions occur all around us, for example in health care, cooking, cosmetics, and agriculture. Complex chemical reactions involving carbon-based molecules take place.			E							
B.3.2 Chemical reactions may release or consume energy. Some reactions such as the burning of fossil fuels release large amounts of energy by losing heat and by releasing gases, light, and sound.	FluKa Chemical WQ Unit - fact sheet 11/13 Pro-grade fu		E							
B.3.3 A large number of important reactions involve the transfer of either electrons (oxidation-reduction reactions) or hydrogen ions (acid-base reactions) between reactants.	Wacker DCI text Canadian Periodicals of 11/13 11/13 ppt - Acids/bases 11/13 double displacement Rxn		E							

CONCEPTUAL UNDERSTANDING STANDARDS Content Standard C: Life Sciences										
C.1. THE CELL		Traver Creek Curr.	QTL	E/G	Pre Test	Essa	Report	Mode	Mode	Post Test
C.1.1	Cells have particular structures that underlie their functions. Every cell is surrounded by a membrane that separates it from the outside world. Inside the cell is a complex network.	Prokaryotes (P) is algae and eukaryotes (E) local eukaryotes		E				1.1	1.2	
C.1.2	Most cell functions involve chemical reactions. Most molecules enter into cells react to provide the chemical constituents needed to synthesize other molecules. Cells break down	1201 Cell scavenger hunt		E						
C.1.3	Cells work and are referenced to guide their functions. The genetic information stored in DNA is used to direct the synthesis of the chemicals of proteins that each cell requires.	1201 Cell scavenger hunt		E						
C.1.4	Cell functions are regulated. Regulation occurs both through changes in the activity of the molecules performed by proteins and through the inherent properties of individual molecules.	1201 Cell scavenger hunt		E						
C.1.5	Most cells contain chloroplasts, the use of photosynthesis use solar energy to combine molecules of carbon dioxide and water into	1201 Cell scavenger hunt		E						
C.1.6	Cells can differentiate, and complex multicellular organisms are formed as a highly organized arrangement of differentiated cells. In the development of these multicellular	1201 Cell scavenger hunt		E						
C.2. THE MOLECULAR BASIS OF HEREDITY		Traver Creek Curr.	QTL	E/G	Pre Test	Essa <td>Report</td> <td>Mode</td> <td>Mode</td> <td>Post Test</td>	Report	Mode	Mode	Post Test
C.2.1	In all organisms, the instructions for specifying the characteristics of the organism are carried in DNA, a large polymer formed from millions of four kinds (A, C, G, and T). The chemical			E				1.1	1.2	
C.2.2	Most of the cells in a human contain two copies of each of 22 different chromosomes. In addition, there is a pair of chromosomes that determine sex, a female carries two X			E						
C.2.3	Changes in DNA (mutations) occur spontaneously at low rates. Some of these changes result in differences in the organism, which can often pass changes to its offspring.			E						
C.3. BIOLOGICAL EVOLUTION		Traver Creek Curr.	QTL	E/G	Pre Test	Essa <td>Report</td> <td>Mode</td> <td>Mode</td> <td>Post Test</td>	Report	Mode	Mode	Post Test
C.3.1	Species evolve over time. Evolution is the consequence of the interaction of (1) the potential for a species to increase in number, (2) the genetic variability of offspring due to			E				1.1	1.2	
C.3.2	The great diversity of organisms is the result of more than 3.5 billion years of evolution that has filled every available niche with life forms.			E						
C.3.3	Natural selection and its evolutionary consequences provide a scientific explanation for the fossil record of ancient life forms, as well as for the striking similarities between			E						
C.3.4	The members of different species of plants, animals, and microorganisms that live on earth today are related by descent from common ancestors.			E						
C.3.5	Biological classifications are based on how organisms are related. Organisms are classified into a hierarchy of groups and subgroups based on similarities which reflect	Identification and classification of local organisms		E						
C.4. THE INTERDEPENDENCE OF ORGANISMS	Traver Creek Curr.	QTL	E/G	Pre Test	Essa <td>Report</td> <td>Mode</td> <td>Mode</td> <td>Post Test</td>	Report	Mode	Mode	Post Test	
C.4.1	The atoms and molecules on the earth cycle among the living and nonliving components of the biosphere.	Water, carbon and nitrogen cycles		E				1.1	1.2	
C.4.2	Energy flows through ecosystems in one direction, from photosynthetic organisms to herbivores to carnivores and decomposers.	1201 Cell scavenger hunt		E						

C.4.3	Organisms both cooperate and compete in ecosystems. The interrelationships and interdependencies of these organisms may vary from ecosystems that are stable to those that are changing.	Ps & Rs- 1211 Cell scavenger hunt Heterotrophs, Autotrophs Food Chain/energy relationships Succession cycles	E																
C.4.4	Living organisms have the capacity to produce populations of almost any size, but environmental and resource limits. The fundamental factors that produce effects on the interactions of organisms are population density, resource availability, and environmental conditions.	Ps & Rs- 1211 Cell scavenger hunt Heterotrophs, Autotrophs Food Chain/energy relationships Succession cycles	E																
C.4.5	Human beings are part of the world's ecosystems. Increasingly, human activity impacts ecosystems as a result of population growth, technology, and consumption. Human activities have profound effects on the interactions of organisms.	Human impact & impacts on water table Human's place in stream food chain TED Streaming is a river video	E																
C.5	MATTER, ENERGY, AND ORGANISMS IN LIVING SYSTEMS	Traver Creek Curr.	OIL	E/G	Pre test	Essa y	Report	Mode 1.1	Mode 1.2	Post test									
C.5.1	All matter tends toward more disorganized states. Living systems require a continuous input of energy to maintain their chemical and physical organization. With death, and the breakdown of the organism, the matter and energy are recycled.	Sun and allochthonous sources of energy, Ps & Rs, Decomposed Food chains & food webs OED energy/energy relationships Sun and allochthonous sources of energy, Ps & Rs 1211 Ps & Rs each project	E																
C.5.2	The energy for life primarily derives from the sun. Plants capture energy by absorbing light and using it to form strong (covalent) chemical bonds between the atoms of carbon-containing molecules.	Ps & Rs 1211 Ps & Rs each project	E																
C.5.3	The chemical bonds of food molecules contain energy. Energy is released when the bonds of food molecules are broken and new compounds with lower energy bonds are formed.	1211 Ps & Rs each project	E																
C.5.4	The complexity and organization of organisms accommodate the need for obtaining, transferring, and storing energy and the ability of the organisms to recycle materials.	Adaptations, Functional feeding groups	E																
C.5.5	The distribution and abundance of organisms and populations in ecosystems are limited by the availability of matter and energy and the ability of the organisms to recycle materials.	Adaptations, Functional feeding groups	E																
C.5.6	Matter and energy flow through different levels of organization of living systems--cells, organs, organisms, communities--and between living systems and the physical environment.	1211 1211 Cell scavenger hunt Ps & Rs 1211 Ps & Rs each project Food Chains	E																
C.6	THE BEHAVIOR OF ORGANISMS	Traver Creek Curr.	OIL	E/G	Pre test	Essa y	Report	Mode 1.1	Mode 1.2	Post test									
C.6.1	Multi-cellular animals have nervous systems that govern behavior. Nervous systems are formed from specialized cells that conduct signals rapidly through the long cell structures.		E																
C.6.2	Organisms have behavioral responses to external changes and to external stimuli. Responses to external stimuli can result from interactions with the organism's own activities.	OED Adaptations of behaviors to environmental types OED types of feeding mechanisms	E																
C.6.3	Like other aspects of an organism's biology, behaviors have evolved through natural selection. Behaviors often have an adaptive function based on survival or efficiency.		E																

CONCEPTUAL UNDERSTANDING STANDARDS Content Standard D: Earth and Space Science																			
D.1	ENERGY IN THE EARTH SYSTEM	Traver Creek Curr.	OIL	E/G	Pre test	Essa y	Report	Mode 1.1	Mode 1.2	Post test									
D.1.1	Earth systems have internal and external sources of energy, both of which create heat. The sun is the major external source of energy. The primary sources of internal energy are:	OED Sun drives photosynthesis, and radiation-temperature effects internal Ps from gas, external Ps from soil, etc.	E																
D.1.2	The constant transfer of earth's internal heat drives convection currents in the mantle that propel the plates comprising earth's surface across the face of the globe. The constant heating of earth's surface and atmosphere by the sun drives convection within the atmosphere and oceans, producing winds and ocean currents.		E																
D.1.3	Global climate is determined by energy transfer between the sun and the earth's		E																

D.3.	GEONEMAL CYCLES	Traver Creek Curr.	OTL	E/O	Pre Test	Essa	Report	Mode 11	Mode 12	Post Test
D.3.1	The earth is a system constantly, essentially a fixed amount of each stable chemical element or element. Each element can exist in several different chemical reservoirs. Each element undergoes movement in nature between reservoirs it drives by the earth's internal and external sources of energy. These movements are closely accompanied by a change in the chemical and physical states of the element.			E						
D.3.2	Movement of matter between reservoirs is driven by the earth's internal and external sources of energy. These movements are closely accompanied by a change in the chemical and physical states of the element.			E						
D.3.3	Interactions among the solid earth, the oceans, the atmosphere, and organisms have resulted in the ongoing evolution of the earth system. The earth system is a dynamic system.			E						
D.3.4	Evidence for sea-level rises of the late Pleistocene - extends back more than 125,000 years. The evidence of the current transgressive cycle is the presence of the earth's sea level.			E						
D.4.	THE ORIGIN AND EVOLUTION OF THE UNIVERSE	Traver Creek Curr.	OTL	E/O	Pre Test	Essa	Report	Mode 11	Mode 12	Post Test
D.4.1	The origin of the universe is the subject of the greatest questions in science. The 'big bang' theory places the origin between 10 and 20 billion years ago, when the universe began as a hot, dense, and rapidly expanding sphere.			E						
D.4.2	Early in the history of the universe, matter, primarily the light atoms hydrogen and helium, clumped together by gravitational attraction to form galaxies, clusters of stars, billions of stars produce energy from nuclear reactions, primarily the fusion of hydrogen to form helium. These and other processes in stars have led to the formation of all the other elements.			E						

CONCEPTUAL UNDERSTANDING STANDARDS CONTENT STANDARD F: Science in Personal and Social Perspectives

F.1.	PERSONAL AND COMMUNITY HEALTH	Traver Creek Curr.	OTL	E/O	Pre Test	Essa	Report	Mode 11	Mode 12	Post Test
F.1.1	Hazards and the potential for accidents exist. Regardless of the environment, the possibility of injury, illness, disability, or death may be present. Humans have a variety of ways to prevent, reduce, or avoid these hazards.			E						
F.1.2	The severity of disease symptoms is dependent on many factors, such as human resistance and the virulence of the disease-producing organism. Many diseases can be prevented.			E						
F.1.3	Personal choice concerning lifestyle and health involves multiple factors. Personal goals, peer and social pressures, ethnic and religious beliefs, and understanding of biological processes are all factors that influence personal choice.	11/13 investigate household sources of PCB and dioxin 12/9 Spawning in a tower water tower - (Middletown, Tenn)		E						
F.1.4	An individual's mood and behavior may be modified by substances. The modifications may be beneficial or detrimental depending on the quantity, type of substance, duration of use, and individual characteristics.			E						
F.1.5	Disorders of tooth and bone growth pattern determine structural balance. Structural balance has a direct effect on growth and development and personal well-being. Personal choices can influence structural balance.			E						
F.1.6	Healthcare serves basic health needs, especially for young children. Regardless of the family structure, individuals have health care needs. A variety of physical, mental, and social factors influence health care needs.			E						
F.1.7	Research is basic to the physical, mental, and social development of humans. Research should understand that human sexuality involves biological, psychological, and social factors.			E						

APPENDIX C
ESSAY ASSIGNMENT

Observation Essay Instructions - This text was included in a handout given to students on September 6, 1996.

Essay #1 Scientific

Written from Traver Creek note-taking observations.

Write an essay that includes 6-7 paragraphs and approximately 300-500 words. This is a descriptive essay about your observations from your trip to Traver Creek. The six to seven paragraphs should be organized as follows:

Introduction (1 paragraph): Where is Traver Creek and what was your purpose in going there. A thesis sentence that makes a judgment about the creek and the surrounding area. The last sentence should be, "We are supposed to observe three major areas of the creek."

Development (3-6 paragraphs): Expand on the three major areas from your note-taking. Write 1-2 paragraphs per area. Make sure that you provide details that support your thesis for each section.

Conclusion (1 paragraph): Summarize your general impressions, give a personal reaction to the area, speculate what's next.

(Reverse side) Criteria for Grading of Essay #1 Scientific:

1. Focus and Organization

- a. Is there a thesis and purpose?
- b. Is there an introduction and conclusion?
- c. Does the essay move energetically with examples and details taken from the note-taking?

2. Content

- a. Does the essay respond directly to the issue posed?
- b. Do each of the sections in the body (panorama, the river, and the banks) of the essay convey a well-developed image to the reader?
- c. Does the writer communicate important scientific observation details and information?

3. Sentence Mechanics/Language

- a. Are there mechanical, punctuation errors?
- b. Are there grammatical or sentence errors
- c. Is language used with fluency and variety
- d. Is meaning clear and concise?
- e. Is writing legible?

© Foundations of Science 1996

APPENDIX D

REPORT ASSIGNMENT

Benthic and Physical Assessment Report Format FOS I 1996-97

TITLE

The title to the report should be concise, but at the same time it should contain enough information so that the reader knows what you investigated. The reader must be able to look at the title and decide if the topic is one that is relevant to them. An example of a self-explanatory title is:

How red-wing black-birds indicate the health of a prairie system

Instead of a title that gives too little information:

Birds on the prairie

INTRODUCTION (About 2 paragraphs)

The introduction should provide a context for the topic under study. The introduction provides the background necessary to understand the rest of the report. In addition the introduction should provide a concise statement of the problem. That is, tell precisely what questions you are trying to answer.

Suggestions for what you should have in the introduction:

- _____ What question were you trying to answer about Traver Creek?
- _____ A description of what benthic organisms are and why they are used as indicators of health for the creek.
- _____ A description of physical forces that influence the benthic community.

MATERIALS AND METHODS (About 2 paragraphs)

This section should provide enough information about how you gathered the data to allow the reader to judge whether you followed proper procedures. This will help the reader decide if your conclusions based on your data are justified.

Suggestions for what to include:

- _____ Describe where (give the section # and what habitats you sampled) and how you collected your benthics
- _____ Describe all the different physical assessments including:
 - physical assessment data sheet, scorecard, transects, and symbol map.
 - *Say that the physical assessment data sheet and scorecard were adapted from the Huron River Watershed Council. List some of the things you were supposed to observe for the physical assessment data sheet.
 - *Describe in detail how you did your transects.
 - *Describe how you calculated discharge.

RESULTS

Results should be brief and objective. This is not the section to draw conclusions. In this section include all your tables and graphs. You will be given a separate sheet that explains tables and graphs.

Suggestions for what to include:

- _____ Benthic spreadsheet.
- _____ Pie chart showing the relative abundance of benthics in each taxa group.
- _____ Symbol map and some way to show where each of them was found.
- _____ Contour map showing creek bottom.
- _____ Pull out the most important details for your section from your physical data sheet and report them in the results.
- _____ Report your score from the physical assessment scorecard.

DISCUSSION

The discussion talks about your data.

Include:

- _____ Complete description of where your section of the creek is and what surrounds it.
- _____ Describe where are the headwaters and mouth of Traver creek and the kind of land it flows through.
- _____ Describe the main physical features of your section itself. Include the substrate, water, banks, dimensions and placement of pools, riffles, and runs. Refer to your maps in the results. Discuss how they can have an impact on the benthics. Emphasize possible habitats or lack of them for benthics.
- _____ Refer to your score on your scorecard and explain what grade you would give the creek.
- _____ For each benthic that you found describe its niche (where it lives and what it eats) and what taxa group it belongs to. Include pictures of the benthics.
- _____ Finish the discussion by relating the benthics you found to you section's physical assessment. Describe why it make sense (or doesn't) for you to have found the benthics that you did. (Refer to your pie chart of taxa group abundance).

CONCLUSION and FUTURE STUDIES (1 paragraph only)

Include:

- _____ Restate your problem statement. The question your were investigating.
- _____ State the major findings and conclusions
- _____ What further investigations could you do to get more information about your problem statement.

APPENDIX E
MODEL I ASSIGNMENT

Text from class handout given to FOS-I students in November 1996. (emphasis in the original handout).

Build a model that shows some of the physical factors of a section of Traver Creek and how they affect the benthic macroinvertebrates.

1. What physical factors of the creek did you measure or observe? (list them)

Choose a few of these factors to focus on. (Circle them.) For each factor that you choose, think about:

a) What affects that factor? (What value did it have for your section? Why did it have that value? What could have made it a different value?)

b.) How does that factor affect the benthics?

c) How does it all fit together? (Are there other relationships between the factors you've chosen?)

Once you have some ideas, you're ready to start building a model. On the server, in the FOS1 folder, there are files with pictures of different sections of the stream, and of different kinds of benthics, for you to use in your model. Make sure you drag the picture files onto the desktop of your powerbook before opening them, so you don't tie up the server.

Your model must convincingly demonstrate your understanding of the physical factors you've chosen to model and how they relate to the benthics. Make sure that you fill out a plan, describe each object and factor, explain each relationship, test your model as you go along, and evaluate it at the end.

Models are due at the end of class on Friday, November 8th.

Coding Sheet for Models

Strategic and Epistemological Understandings in Plans and Evaluation

Plan: Goal:

A.1.1/E.1.1 - Problem Definition: SU

3. defines a reasonable, well focused problem area to be addressed within constraints, identifies essential elements of the problem
makes a prediction, stating possible outcomes,
2. defines , through revision, a reasonable problem area to be addressed within constraints
1. has difficulty defining a reasonable problem area given constraints, problem may be too broad
0. non-existent, no evidence does not define a problem

A.2.2 Describes purpose for model : EU

3. model is constructed in the service of developing and testing ideas
2. specific, explicit purpose for model but focus on reality, not ideas
1. Models are simple copies of reality or identifies no purpose beyond class expectations
0. non-existent, no evidence does not fill in goal statement

Plan: Objects & Factors

A.1.2/E.1.2 Planning model SU

3. designs and implements a method to address problem, including gathering resources, synthesizing information, organizing and presenting findings..
Model plan identifies major objects and factors (66-100% are IDed in plan).
2. designs a method to address problem, may have trouble getting started
(33-66% objects and factors IDed in plan)
1. had difficulty designing a method to address problem, problem may be too broad
(fewer than 1/3 (33%) of objects and factors IDed in plan).
0. non-existent, no evidence, does not attempt to address problem

A.2.5 Explains rationales for plan: EU

3. evaluating which of several designs could be used to serve the model's purpose.
uses empirical evidence to justify or evaluate the inclusion of objects/factors
2. evaluating which of several designs could be used to serve the model's purpose
OR
provides rationale for including specific objects/factors
1. provides non-scientific rational
0. non-existent, no evidence,

Implementation

E.1.3. - Implement a proposed solution/building the model SU

3. constructs a model with explanatory power, elegance and parsimony
2. constructs a model with some detail and explanatory power; OR includes too much detail so that explanatory power is lost.
1. constructs a simple model with little detail and little to no explanations
0. non-existent, no evidence,

Evaluation: How Well Does the Model Work

E.1.4 Evaluates the Model and its behavior: SU

3. Evaluates the model against the stated goals/purpose, provides evidence of how well (or not) the model worked. new criteria not originally considered may be reviewed.
2. Evaluated model against purpose . No evidence, ("it show what I wanted."
1. Naive - "It's great, fine, etc."
0. non-existent, no evidence,

G.2.2 - Nature of knowledge in the models: EU

3. Models can be manipulated and subjected to tests in the service of informing ideas
tests and revises model to better account for evidence or new explanations
 2. Tests of the model are tests of the workability of the model, not tests of underlying ideas
tests and revises model to produce desired outcome
 1. Tests model but does not revise model
 0. non-existent, no evidence of testing model
- ***Need to look at Log Sheets for this information!!!!***

Evaluation: What would you Change:

G.2.3 Tentative nature of knowledge EU

3. science as human construct, used to build explanations of phenomena
theories and models can be revised (tentative nature of science)
"Use better information"
2. describes science as a human process used to build explanations of phenomena
some ideas in science are subject to change.
("add more factors")
1. Science as truth, a static body of knowledge
"change nothing"
0. non-existent, no evidence

Content Understanding in Factors (CU's)

Accuracy of description:

3. Accurate - student description/definition of the factor consistent with prevailing scientific thought,
2. Semi-accurate- student description mixes scientific reasoning with alternative reasoning; or student provides an example instead of a description.

1. nonscientific - student description uses non-scientific description ; or student reiterates the factor, or student used default explanations for population factor
0. no evidence of understanding - student did not provide a coherent description , or student did not fill in dialogue box.

Accuracy of Ranges:

3. domain-accurate/realistic-quantitative--given the factor's scientific domain, are its maximum and minimum values accurate? (e.g., pH has a range from 1-14); or the factor created with some other quantitative range (not default) that seems plausibly realistic (e.g., depth = 0-50 cm) a
2. arbitrary-realistic-qualitative--was the factor created with some qualitative range (not default) that seems plausibly realistic (e.g., substrate rocky to sandy)
 1. arbitrary-default--the factor was created using Model-It's default range of low to high or zero to one hundred?
 0. arbitrary-no criteria--was the factor created with a range that seems unrealistic or totally random (e.g., discharge with a range of -10 to 25)?

Content Understanding in Relationships

Relationship Accuracy Categories

7. in agreement with generally accepted scientific knowledge (Level 3 CU)
or Not in agreement: (expanded subcategories)
6. *incorrect default*: students used the default 1:1 linear relationship.
5. *incorrect* --one factor does affect another, but the relationship the students created is wrong (wrong shaped graph)
4. *wrong orientation* --as one factor changes, another factor changes, but it changes in the opposite direction
codes 4-6 = Level 2 CU
3. *backwards* --the factors are causally related but the sense of the relationship should have been reversed; (dependent/independent reversed)
2. *not causal*--factors in the relationship are not strictly causally related, but they may be correlated.
 1. *nonexistent* --there is no scientific relationship between the two factors; - or relationship is redundant of another relationship
codes 1-3 = Level 1 CU

Content in relationship Explanations

Explanation Accuracy

The "because" "How" or "why" of the relationship.

3. Accurate - student explanation consistent with prevailing scientific thought or reasoning, no naive conceptions .
2. Semi-accurate- student explanation mixes scientific explanations with alternative explanations.
 1. nonscientific - student explanation uses non-scientific reasoning, or reiterates relationship or uses defaults on population factors

0 no evidence of understanding - student did not provide a coherent explanation, or did not fill out response box

Explanation Elaboration

3 Rich/elaborate: explanation integrates and applies knowledge presented in class, contains detail and elaborates on the relationship, may includes examples (original or from curriculum)

2. Moderate or Partially complete: explanation lacks detail and elaboration, may be missing minor constructs/relationships. may includes examples (from curriculum)

1. Poor/shallow, incomplete: explanation expressed in generalities, may be missing several major constructs/relationships and minor constructs/relationships.

0. Nonsensical, or no evidence (e.g. student only reiterates relationship.) or no response

Overall Calculations Take Median Value

Overall Accuracy in Factors: definitions

3. Accurate - At least 50% of the factors were coded as "3" in the individual coding, and none were coded as 0 or 1.

2. Semi-accurate - At least 50% of the factors were coded as "3" or "2" in the individual coding.

1. nonscientific understanding - at least 50% of the factors coded as 3 or 2 (or 4)

0. lack of understanding - 50% or more of the individual descriptions received a rating of "1" or "0".

ranges:

4. Scientifically accurate - 50% or more of the factors received category 4.

3. arbitrary quantitative - 50% or more of the factors received category 3 or categories 4 and 3.

2. arbitrary qualitative - 50% or more of the factors were coded as category 2.

1. Default - 50% or more of the factors had the default range of values.

0. arbitrary- no criteria - 50% or more of the factors were coded as category 0.

Model's overall accuracy in relationships: _____

4. Scientifically accurate - 50% or more of the relationships received category 7. None were coded as 1, 2, 3, 4,.

3. Semi-accurate - 50% or more of the relationships classified as categories 5 and/or 6, 7

2. Inaccurate-wrong - 50% or more of the relationships classified as categories 4 and/or 3,

1. Arbitrary - 50% or more of the relationships classified as categories 3, 2 and/or 1.

Overall explanations:

Explanation Accuracy:

4. Accurate - At least 50% of the explanations were coded as "4" in the individual coding.
3. Semi-accurate - At least 50% of the explanations were coded as "4" or "3" in the individual coding.
2. nonscientific understanding - at least 50% of the explanations coded as 3 or 2 (or 4)
1. lack of demonstrated understanding - 50% or more of the individual explanations received a rating of "1" or "0"

Explanation Overall Elaboration:

4. Rich/Elaborate - at least 50% of the explanations were coded as "4", none as 1 or 0
3. Moderate - at least 50% of the explanations were coded as 3 or 4.
2. Shallow - at least 50% of the explanations were coded 2 or higher.
1. Nonsensical 50% of the explanations were coded 1 and/or 0

APPENDIX F

MODEL 2 ASSIGNMENT

FOS1 January Model-It Assignment

For your final Traver Creek project, you will build a model that demonstrates in-depth understanding of a stream ecosystem.

Design your own project: Build a model that represents some aspect of a stream ecosystem. You may use data you have collected about the stream, and other resources such as manuals, textbooks, teachers, and mentors to help you set up the relationships.

Your model should include physical, chemical, and biological factors of the stream.

Choose one area to focus on, and build the best representation you can that demonstrates how it works in depth.

Talk it over with your partner, and brainstorm ideas. Use the back of this sheet to sketch your ideas if you want. We will also brainstorm ideas in class.

Final models are due at the beginning of class on Friday, January 17th.

You must turn in a printout of the model and its factor map. On the 17th, we will have a round-robin where you will look at and evaluate other student's models.

You will be graded on:

Overall Model

1. Well chosen project that meets the criteria: includes relationships between physical, chemical, and biological factors
2. Scientifically accurate / correct

3. Well-structured/ designed, complex, shows in-depth understanding.
4. Works when you run it.
5. At least one independent factor affects other factors so model can be manipulated.

Model Components:

1. Objects / factors make sense.
2. Object, factor descriptions are stand-alone and inform about what they are and measure.
3. Relationships are explained (not just restated).
4. Plan filled in specifically.
5. Evaluation filled in thoughtfully.

Text from classroom handout given to students on January 8, 1997

APPENDIX G
PRE/POST TEST FORM A

COMMUNITY HIGH SCHOOL

FOUNDATIONS OF SCIENCE - I

Fall Semester Post-test

Student's Name: _____

Period: _____ Block 1 - 8:00 - 9:30 MW (Kathe)
 _____ Block 3 - 9:40 - 11:10 MW (Liz)
 _____ Block 5 - 12:00 - 1:30 MW (Kathe)
 _____ Block 7 - 1:40 - 3:10 MW (Liz)

Instructions: Carefully follow the directions for each set of questions. Your instructors will use this post-test to help assess what you have learned about the stream project, water quality, chemistry, ecology, and photosynthesis this semester. Your responses will also be used in the study about what students learn when engaged in extended projects like the study of Traver Creek. Do your very best and answer all of the questions with the attitudes and knowledge you have today. Thanks!

1. List three rivers/creeks within 50 miles of your school.
 - a.
 - b.
 - c.
2. In what watershed is your school located?
3. Have you ever visited Traver Creek?
4. What are some kinds of activities you like to do in or around a river or creek?
5. How much do you agree with each of the following statements. (Circle one)
 1 = I don't agree at all... 5 = I agree very much... X = not sure/don't know

Water pollution is not a problem because people don't drink river/creek water directly.	1	2	3	4	5	X
From an environmental standpoint, how I live my life doesn't matter much.	1	2	3	4	5	X
There's too much unnecessary worry about pollution	1	2	3	4	5	X
The laundry detergent that cleans my shirt will help clean the water in the river.	1	2	3	4	5	X
As you go down the river/creek, the water is likely to become cleaner	1	2	3	4	5	X
How we care for our lawns has an impact on the river/creek.	1	2	3	4	5	X
Trying to clean up the river/creek just seems too much of a pain.	1	2	3	4	5	X
I don't think I should have to change the way I live if no one else is willing to.	1	2	3	4	5	X
It's hard to find examples where citizen action has really made a difference in the environment.	1	2	3	4	5	X
There are a number of things I can do to help clean up the river/creek.	1	2	3	4	5	X
Each time we find a solution to an environmental problem, it seems to cause new ones.	1	2	3	4	5	X
Individual efforts to save water may seem small, but they add up.	1	2	3	4	5	X
It would be embarrassing to be the only one in my school who picked up litter.	1	2	3	4	5	X

Golf courses can contribute to water pollution.	1	2	3	4	5	X
I have trouble finding information that would inform me about environmental issues.	1	2	3	4	5	X
Trying to live in balance with nature isn't possible -- we've altered it too much.	1	2	3	4	5	X

6. Rank the following environmental problems from one to five. Put a one next to the item that you think is the most pressing problem facing Americans today, and a two next to the second most pressing environmental issue, and so forth, until you get to five.

- ___ water pollution
- ___ air pollution
- ___ indoor radon
- ___ pesticides and chemicals in our food
- ___ toxic chemicals in our environment (PCB's, dioxin, lead in the soil, etc.)

7. List five (5) things you could personally do to reduce your contribution to water pollution.

- a.
- b.
- c.
- d.
- e.

8. Into what body of water does Traver Creek empty?

9. You see someone polluting the river/creek, to whom would you report that person?

Multiple choice: Circle the BEST answer.

10. How would you rate the quality of Traver Creek Water?
 a. awful
 b. pretty dirty
 c. all right
 d. pretty clean
 e. great
11. Which Dissolved Oxygen (DO) saturation level would be the best for the creek?
 a. 30%
 b. 60%
 c. 90%
 d. 120%
 e. Don't know
12. Which type of substrate would you expect to find in a riffle area in a stream?
 a. fine silts
 b. sand
 c. gravel
 d. cobbles and boulders

13. Fecal coliform is: (Check [v] all that apply)
- _____ a chemical from industry that pollutes waterways.
 - _____ another name for untreated sewage.
 - _____ usually a non-point source pollutant in the creek.
 - _____ a bacteria that lives in the guts of mammals.
 - _____ a benthic macro-invertebrate that decomposes organic pollution.

14. Which of these would indicate an improvement in Traver Creek's water quality? (Check [v] as many as appropriate.) Then choose ONE of the items that you checked and explain why you think that would indicate an improvement in water quality.

- _____ An increase in the number of fecal coliforms.
- _____ A rise in the amount of total solids
- _____ A shift in the Water Quality Index from 55 to 71.
- _____ Finding a lot of different kinds of insect larvae in the water.

(Why?)

15. What would happen if extra phosphates enter the creek? Describe this process in terms of photosynthesis, cellular respiration, Dissolved oxygen (DO) and Biochemical Oxygen Demand (BOD). Include any relevant chemical equations in your answer.

16. Use the Periodic Table to complete the following chart.

Atomic number	Elemental name	Symbol	Atomic mass	number of electrons
19				
		Cl		
	Cadmium			

17. How do each of the following affect the water quality of Traver Creek/Huron River? (makes it worse, makes it better, doesn't affect it.) Choose **TWO** that you think might change and describe what related factors or parameters might also be affected or involved in the change?

Fertilizing your lawn.

Water runoff during a heavy storm.

Installing a combined storm and sanitary sewer system.

Charing the trees and bushes from the banks of the stream.

Expanding a parking lot at a shopping mall.

Multiple Choice: Circle the BEST answer.

18. A solution in a dish contains 3 grams of salt dissolved in 100 grams of water. If 50 grams evaporates from the solution, how much salt remains in the dish?
- a. 0.0 grams
 - b. 1.0 grams
 - c. 1.5 grams
 - d. 3.0 grams

19. Juan thinks that water will evaporate faster in a warm place than in a cool one. He has two identical bowls and a bucket of water. He wants to do an experiment to find out if he is correct. Which of the following should he do?
- a. Place two bowls with the same amount of water in a warm place.
 - b. Place a bowl of water in a cool place and a bowl with twice the amount of water in a warm place.
 - c. Place a bowl of water in a cool place and a bowl with half of the amount of water in a warm place.
 - d. Place a bowl of water in a cool place and a bowl with the same amount of water in a warm place.

20. The droplets of water on the outside of the glass containers pictured below most likely result from which of the following?



- a. Water is leaking through the container wall.
 - b. Water in the air outside the container is cooling and changing from vapor to liquid.
 - c. Air above the ice inside the container is warming and changing from vapor to liquid.
 - d. Cold air is carrying water from the inside to the outside of the container.
21. Currently, scientists make many predictions about the way solids behave by assuming that the atoms in the solid behave as if they are connected by tiny springs. This means that which of the following is true?

- a. There must be tiny springs connecting atoms together.
- b. Scientists will always find it useful to think of atoms as connected by springs.
- c. An intelligent life form on another planet would almost certainly think about solids in the same fashion.
- d. Thinking of atoms as being connected by springs constitutes a useful model of solids.
- e. Because it is a successful model, the spring model of the atom can be classified as a fact.

22. **Fill in the Blank:** Aquatic insects can be grouped according to how and what they eat. These different "functional feeding" groups reveal the many roles aquatic insects play in stream ecosystems. From the list below, fill in the name of the functional feeding group described on the blanks provided.

- | | |
|----------------------|----------------------|
| | Grazers/Scrapers |
| Predators | Filtering Collectors |
| Gathering Collectors | Shredders |

- a. The black fly larva (*Simuliids*) is found in fast current areas where it catches small organic particles with its extended fan-shaped brushes. It is an example of _____ which strain minute or tiny particles out of the water.
- b. Some Caddisfly (*Trichoptera*) larvae are _____. This functional feeding group eats coarse dead plant material: leaves, grasses and aquatic plants. They break large pieces of organic matter into finer material that is released in their feces.
- c. _____, such as mayflies, feed on algae that's growing on stones and other substrates in the water. This functional feeding group are usually flattened dorso-ventrally and very streamlined.

Multiple Choice: Circle the Best Answer.

- 23. Group 1 elements are similar in all of the following ways **except**:
 - a. they all lose one electron
 - b. they react in similar ways
 - c. they all have the same number of protons
 - d. they are all metals
- 24. Which of the following is an example of an homogenous mixture
 - a. bean soup
 - b. chicken broth
 - c. pond water
 - d. salt water
- 25. Which of the following is a chemical change?
 - a. tearing a piece of paper
 - b. a candle melting in the hot sun
 - c. settling of pond water
 - d. rusting steel wool

26. **Metals and Non-metals:**

a. Name the compound that is represented by the chemical formula:

LiF _____

AlCl₃ _____

Mg₃(PO₄)₂ _____

b. Write the chemical formula for these compounds.

Potassium iodide _____

Magnesium fluoride _____

Calcium nitrate _____

c. Write the complete equation for this double displacement reaction using chemical symbols. Balance each compound. [You do not need to balance the equation!]

potassium chloride and calcium sulfate

27. **Fill in the Blank:** Write the term from the list below that best fits the phrase in the answer blank provided.

- | | |
|----------------|-----------------------|
| aquatic plants | embeddedness |
| eroding banks | fine sediment |
| flow rate | meters |
| pools | predominant substrate |
| riffles | stable habitat |
| tenths of feet | |

a. _____ measures how deeply covered the rocks and gravel are by sand and silt.

b. _____ is another term for hiding places for small animals and teeny

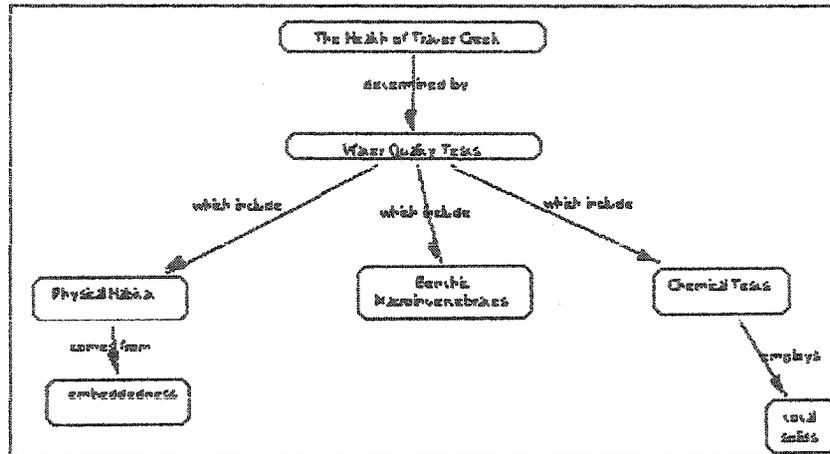
creatures. It includes rubble, gravel, submerged logs, and undercut banks.

c. One of the things you will look for along the side of the stream are areas of bare soil, scoured banks, and sloughing banks, exposed roots, and leaning trees. This best fits the description for-_____.

d. The areas of the stream water that bubbles over rocks (mini-rapids) are called _____.

28. What was the most interesting thing that you discovered, learned or did during the Traver Creek project?

29. Finish the concept map by adding (at least) 4 concepts under benthics, 2 more concepts under physical habitat and 3 more concepts under Water quality tests. Use the following words (or others of your choice) to show the relationships between the concepts: employs, comes from, determines, depends on, is a kind of, for example, needed by, covers, leads to, changes, results in, etc. Also draw lines and arrows to indicate how the concepts under each of the three main assessments are related to each other. Finally write a paragraph explaining the relationships you drew in your map.



30 Scenario 1: Read carefully, then answer the questions below.

Fishing along the Huron River is a popular activity. Barton Pond is locally known for its bass, bluegill, perch and carp recreational fishery. Occasionally anglers will catch walleye and pike in this impoundment. This summer, however, reported catches are way down. Michigan Department of Natural Resources fishery specialists were sent to investigate. They found that the populations of the fish caught in the survey nets to be way down. They also discovered a species of fish never before reported in the Huron River.

The fish, the ruffe (*Gymnalespethus asraukus*) is a small European member of the perch family. It was first introduced into the Great Lakes area around 1935. Ruffe grow rapidly and can reproduce in their first year. Females in the Great Lakes area have been known to lay between 45,000 and 90,000 eggs a year. Ruffe are primarily bottom feeders. They rarely grow bigger than 5 inches, although the sharp spines on their gill covers, dorsal and anal fins make them difficult for larger fish to eat.

a. What do you think is happening to the native fish community?

b. How can you find out if your hypothesis is correct? What will you investigate?

c. How will you use the investigation results to see if your hypothesis is correct?

31. Scenario 2: Read carefully, then answer the questions below.

During the summer of 1986, wind-surfing on the Hwon River in Gallup Park was a popular activity. Boards could be rented from the concession stand. Then several Ann Arbor students came down with hepatitis, a disease that attacks the liver. Public health officials determined that the only activity the students had in common was the wind-surfing, yet when they tested the water in Gallup Park, they could not find the source of the disease. The students at Hwon High School decided to carry out their own investigations and they were able to confirm that the source of the disease was the river water, and that the highest risk for contracting the disease came after heavy rains.

a. What do you think is happening to the river that makes it dangerous to windsurfers?

b. How can you find out if your hypothesis is correct? What will you investigate?

c. How will you use the investigation results to see if your hypothesis is correct?

d. If the test supports your hypothesis... what needs to happen to make the river safe for wind-surfing?

Answer Key for PRE/POST ASSESSMENT

Form A

1. List three rivers/creeks within 50 miles of your school.
2. In what watershed is your school located?
3. Have you ever visited Traver Creek?
4. What are some kinds of activities you would like to do in or around a river/creek?
5. How much do you agree with each of the following statements. (Circle one)
1 = I don't agree at all...5 = I agree very much...
X=not sure/don't know
 - a Water pollution is not a problem because people don't drink river/creek water directly.
 - p Trying to live in balance with nature isn't possible -- we've altered it too much
6. Rank the following environmental problems from one to five. Put a number one next to the item that you think is the most pressing problem facing Americans today, and a two next to the second most pressing environmental issue, and so forth, until you get to five.
 - ___ water pollution
 - ___ air pollution
 - ___ indoor radon
 - ___ pesticides and chemicals in our food
 - ___ toxic chemicals in our environment (PCB's, dioxin, lead in the soil, etc.)
7. List five (5) things you could personally do to reduce your contribution to water pollution.
8. What body of water does Traver Creek empty into?
9. You see someone polluting the river/creek, to whom would you report that person?
of agencies

Form B

- # Correct: e.g.
Huron: Traver, Honey, Fleming, Allen, Mallets, Mill, Clinton, Detroit, Rouge, Saline, Raison, Kalamazoo, Grand River etc.
- 3 = L3, 2=L2, 1=L1
- 0=no response, 1= Huron, 2=don't know, 3= other
- 0=no, 1=yes, 2 = unknown/NR
- list. 1-4,
- Code 1-5, X=9 examine frequency distribution
- As long as my drinking water is clean, I don't have to worry about water pollution.
- Trying to live in balance with nature isn't possible -- we've altered it too much
- Code:
- 6a. 1-5 water pollution
 - 6b. 1-5 air pollution
 - 6c. 1-5 indoor radon
 - 6d. 1-5 pesticides
 - 6e. 1-5 toxic chemicals
- count # of items
- 0 = wrong or no response, 1= Huron River,
9. List the governmental agencies or public officials who are responsible for Traver Creek?
of agencies.

10. How would you rate the quality of Traver Creek Water?

- a. awful
- b. pretty dirty
- c. all right
- d. pretty clean
- e. great

- 1 = a. awful
- 2 = b. pretty dirty
- 3 = c. all right
- 4 = d. pretty clean
- 5 = e. great

11. Which Dissolved Oxygen (DO) saturation level would be the best for the creek?

- 1= a. 30%
- 2= b. 60%
- 3= c. 90%
- 4= d. 120%
- 5= e. Don't know

11. Which pH level (degree of acidity) would be best for the creek?

- 1= a. 3.5
- 2= b. 5.5
- 3= c. 7.5
- 4= d. 9.5
- 5= e. Don't Know

12. Which type of substrate would you expect to find in a riffle area in a stream?

- 1= a. fine silts
- 2= b. sand
- 3= c. gravel
- 4= d. cobbles and boulders

12. Which type of substrate would you expect to find in a pool area in a stream?

- 1= a. fine silts
- 2= b. sand
- 3= c. gravel
- 4= d. cobbles and boulders

13. Fecal coliform is: (Check [✓] all that apply)

- a. a chemical from industry that pollutes waterways.
- b. another name for untreated sewage.
- c. usually a non-point source pollutant in the creek.
- d. a bacteria that lives in the guts of mammals.
- e. a benthic macro-invertebrate that decomposes organic pollution.

13a 0 = unmarked, 1=marked

13b. 0 = unmarked, 1=marked

13c. 0 = unmarked, 1=marked

13d 0 = unmarked, 1=marked

13e 0 = unmarked, 1=marked

14. Which of these would indicate an improvement in Traver Creek's water quality?

total 4 = L3, 3=L2, 2 or 1=L1

14 a. An increase in the number of fecal coliforms. 0 = unmarked, 1=marked

14a. A decrease in the number of fecal coliforms. 0 = unmarked, 1=marked *Fecal coliform numbers have a big impact on WQ and indicate problems such as untreated sewage entering a system. A decrease would indicate that that source of pollution is decreasing*

14b. A rise in the amount of total solids 0 = unmarked, 1=marked

14 b. A rise in the turbidity 0 = unmarked, 1=marked

14c. A shift in the Water Quality Index from 55 to 71. 0 = unmarked, 1=marked *In the WQI, higher values indicate a better water quality. This would be the result of an improvement in one or more of the 9 parameters, such as DO, pH, FC, BOD. etc.*

14c. A shift in the Water Quality Index from 76 to 68. 0 = unmarked, 1=marked

14d. Finding a lot of different kinds of insect larvae in the water. 0 = unmarked, 1=marked *High diversity is usually an indication of better environmental quality.*

14d. Finding a lot of one kind of insect larvae in the water. 0 = unmarked, 1=marked *It depends... if you find a lot of "Taxa 1" insects, (mayflies, stoneflies, etc.) It may indicate good water quality. Lots of a Taxa 3 would indicate poor water quality. Usually its better to find a diversity of species rather than dominance by a single species.*

15. What would happen if extra phosphates enter the

Extra phosphates might come from human wastes, animal wastes, fertilizers, human disturbance of land & vegetation, draining wetlands (F.5.2)

Phosphates are a plant nutrient (fertilizer). Extra phosphates in the creek will cause algae to grow (population growth). rapid growth = algal bloom (C.5.5, F.2.1)

Algae carry out photosynthesis (fix energy in the form of carbon in the presence of sunlight ($6\text{CO}_2 + 6\text{H}_2\text{O} \Rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$)). More algae carrying out Ps means more oxygen will be produced during the day (increased DO) (C.5.2) Algae (living things) carry out cellular respiration ($\text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2 \Rightarrow 6\text{CO}_2 + 6\text{H}_2\text{O}$). At night, use of O_2 exceeds production and DO drops. (C.5.2)

After the phosphates are used up, the algae dies and begins to decompose. Decomposition by bacteria requires DO (= BOD) and so DO levels drop further. (C.5.1)

3 = Accurate- student explanation consistent with prevailing scientific thought. Explanation contains detail, and elaborates, may include examples

2 = Semi-accurate - student mixes scientific reasoning with alternative reasoning; explanation lacks detail and elaboration, may be missing constructs/relationships

1 = non scientific - student explanation uses non scientific reasoning, may be expressed in generalities - missing several major constructs/relationships

0 = no understanding - students did not provide a coherent description although they did attempt to write some sort of response.

9 = no response, nothing written in space (=blank).

Other living things (e.g. benthics) may be affected by the changing DO levels. E.g. Taxa 1 organisms (e.g. mayflies) cannot survive in low DO waters (C.5.5)

16. Use the Periodic Table to complete the following chart.

Atomic number	Elemental name	Symbol	Atomic mass	number of electrons
20	a Calcium	b. Ca	c. 40	d. 20
e. 11	f. Sodium	Na	g. 23	h. 11
i. 15	Phosphorus	j. P	k. 31	l. 15

number of electrons (d,h,l) (3,2,1),

Atomic number	Elemental name	Symbol	Atomic mass	number of electrons
19	Potassium	K	39	19
17	Chloride	Cl	35.5	17
48	Cadmium	Cd	112.4	48

Naming/symbols (a,b,f,j,) 4 - L3, 3= L2, 2 or 1 = L1

atomic # and Mass (c,e,j,i, k) 5=L3, 3-4 = L2, 2or 1 = Level 1.

17. How do each of the following affect the water quality of Traver Creek/Huron River? (makes it worst, makes it better, doesn't affect it.) Choose **TWO** that you think might change and describe what related factors or parameters might also be affected or involved in the change?

Raw coding:

0 = no response

1 = makes it worse,

2 = makes it better,

3 = doesn't affect it

correctness coding. 3 = correct, 1= wrong response, 0=no response

Explanations:

3 = Accurate- student explanation consistent with prevailing scientific thought. Explanation contains detail, and elaborates, may include examples

2 = Semi-accurate - student mixes scientific reasoning with alternative reasoning; explanation lacks detail and elaboration, may be missing constructs/relationships

1 = non scientific - student explanation uses non scientific reasoning, may be expressed in generalities - missing several major constructs/relationships

0 = no understanding - students did not provide a coherent description although they did attempt to write some sort of response.

9 = no response. nothing written in space. (=blank)

total = average of response and explanation, if given.

17 a. Fertilizing your lawn.

3 Worse- NOx and PO4 from the fertilizer may wash into the stream, causing algal blooms, increasing BOD, etc.

17 b. Water runoff during a heavy storm.

3 Worse- runoff will carry pollutants into the stream or

2 Better - the flooding will scour the stream bottom, improving habitat

17 c. Installing a combined storm and sanitary sewer system.

3 Better - storm runoff (with NPS pollutants) will be treated before reaching the stream. unless

2 Worse - if system capacity is exceeded, untreated sewage may be released into the stream to prevent flooding of the treatment plant.

17. d Clearing the trees and bushes from the banks of the stream.

3 Worse- clearing decreases shading which allows temperature to increase. -> DO2 decreases Less vegetation increases bank erosion, leading to silting of habitats. Also less leaves for food chain.

1= better

17 e. Expanding a parking lot at a shopping mall.

3 Worse - increases rapid runoff. less water soaks into ground. Stream flow regime becomes flashier. runoff carries pollutants into stream

17 a. Using a high phosphate detergent to wash the family car in the drive way.

3 Worse- PO4 from the detergent may wash into the stream, via storm drains, causing algal blooms, increasing BOD, etc.

17 b. Six weeks of hot summer days without rain.

3 Worse - less water + heat increases water °T, --> decreased DO and heat stress

17 c. Separating storm sewers from sanitary sewers.

2 Better - Less chance of untreated sewage being released into the water way ...

but

3 Worse - since NPS pollutants will no longer be treated in storm runoff

17. d Restoring and protecting natural wetlands along the stream.

3 Better - increases water retention, decreasing "flashiness" of stream. wetlands act a filters - retaining NPS pollutants and silt. and absorbing nutrients

1 = worse

17 e. Expanding the Leslie golf course

3 Worse - construction leads to erosion and silting of stream habitats. more course --> more fertilizer--> more nutrients entering stream.. (and more geese?)

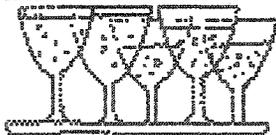
18. A solution in a dish contains 3 grams of salt dissolved in 100 grams of water. If 50 grams evaporates from the solution, how much salt remains in the dish?

- 1= a. 0.0 grams
- 2= b. 1.0 grams
- 3= c. 1.5 grams
- 4= d. 3.0 grams

19. Juan thinks that water will evaporate faster in a warm place than in a cool one. He has two identical bowls and a bucket of water. He wants to do an experiment to find out if he is correct. Which of the following should he do?

- 1= A. Place two bowls with the same amount of water in a warm place.
- 2= B. Place a bowl of water in a cool place and a bowl with twice the amount of water in a warm place.
- 3= C. Place a bowl of water in a cool place and a bowl with half of the amount of water in a warm place.
- 4= d. Place a bowl of water in a cool place and a bowl with the same amount of water in a warm place.

20. The droplets of water on the outside of the glass containers pictured below most likely result from which of the following?



- 1= A. Water is leaking through the container wall.
- 2= B. Water in the air outside the container is cooling and changing from vapor to liquid.
- 3= C. Air above the ice inside the container is warming and changing from vapor to liquid.
- 4= D. Cold air is carrying water from the inside to the outside of the container.

18. A solution in a dish contains 5 grams of salt dissolved in 100 grams of water. If 50 grams evaporates from the solution, how much salt remains in the dish?

- 1= a. 0.0 grams
- 2= b. 1.0 grams
- 3= c. 2.5 grams
- 4= d. 5.0 grams

19. Yvette thinks that water will evaporate faster in a windy place than in a still one. She has two identical bowls and a bucket of water. She wants to do an experiment to find out if she is correct. Which of the following should she do?

- 1= A. Place two bowls with the same amount of water in a warm, windy place.
- 2= B. Place a bowl of water in a protected place and a bowl with twice the amount of water in a windy place.
- 3= C. Place a bowl of water in a protected place and a bowl with half of the amount of water in a windy place.
- 4= d. Place a bowl of water in a protected place and a bowl with the same amount of water in a windy place.

20. The droplets of water on the outside of the glass containers pictured below most likely result from which of the following?



- 1= A. Water is leaking through the container wall.
- 2= B. Water in the air outside the container is cooling and changing from vapor to liquid.
- 3= C. Air above the ice inside the container is warming and changing from vapor to liquid.
- 4= D. Cold air is carrying water from the inside to the outside of the container.

21. Currently, scientists make many predictions about the way solids behave by assuming that the atoms in the solid behave as if they are connected by tiny springs. This means that which of the following is true?

1= A. There must be tiny springs connecting atoms together.

2= B. Scientists will always find it useful to think of atoms as connected by springs.

3= C. An intelligent life form on another planet would almost certainly think about solids in the same fashion.

4= **D. Thinking of atoms as being connected by springs constitutes a useful model of solids.**

5= E. Because it is a successful model, the spring model of the atom can be classified as a fact.

22 0 = wrong/NR, 1 = correct.

a. Filtering Collectors

b. Shredders.

c. Grazers/Scrapers ,

23 Group 1 elements are similar in all of the following ways **except**:

1=a. they all lose one electron

2=b. they react in similar ways

3=c. **they all have the same number of protons**

4=d. they are all metals

24 Which of the following is an example of an homogeneous mixture

1= a. bean soup

2= b. chicken broth

3= c. pond water

4= **d. salt water**

25 Which of the following is a chemical change?

1= a. tearing a piece of paper

2= b. a candle melting in the hot sun

3= c. settling of pond water

4= **d. rusting steel wool**

26. **Metals and Non-metals:**

a. Name the compound that is represented by the chemical formula:

a1 LiF Lithium fluoride

a2 AlCl₃ Aluminum chloride

a3 Mg₃(PO₄)₂ Magnesium phosphate

b. Write the chemical formula for these compounds.

b1 Potassium iodide KI

b2 Magnesium fluoride MgF₂

b3 Calcium nitrate Ca(NO₃)₂

c. Write the complete equation for this double displacement reaction using chemical symbols. Balance each compound. [You do not need to balance the equation!]

potassium chloride and calcium sulfate $2KCl + CaSO_4 \rightleftharpoons CaCl_2 + K_2SO_4$ (post test only)

21. Currently, scientists make many predictions about the way solids behave by assuming that the atoms in the solid behave as if they are connected by tiny springs. This means that which of the following is true?

d. Thinking of atoms as being connected by springs constitutes a useful model of solids.

22. 0 = wrong/NR, 1 = correct.

a. Filtering Collectors.

b. Shredders.

c. Predators,

functional group understandings = sum (3,2,1)

23. Group 7 elements are similar in all of the following ways **except**:

1=a. they all gain one electron.

2=b. they react in similar ways

3=c. they are all non metals

4= **d. they all have the same atomic number**

24. Which of the following is an example of an heterogeneous mixture

1= a. **chicken broth**

2= b. salt water

3= c. clean air

4= d. rubbing alcohol

25. Which of the following is a physical change?

1= a. rusting steel wool

2= **b. a candle melting in the hot sun**

3= c. the color change when testing for nitrates

4= d. mixing two liquids, which form a solid precipitate

26. **1= correct response, 0 = wrong or NR**

KI Potassium iodide

MgF₂ Magnesium fluoride

Ca(NO₃)₂ Calcium nitrate

understanding = average of responses

b. Write the chemical formula for these compounds.

Lithium fluoride LiF

Aluminum chloride AlCl₃

Magnesium phosphate Mg₃(PO₄)₂

understanding = average of responses

27 Fill in the Blank: Write the term from the list below that best fits the phrase in the answer blank provided.

- a. embeddedness
- b. stable habitat
- c. -eroding banks
- d. riffles

S1a(28/30). Problem definition/hypothesis
CU's

C.4.3 = identified problems as biotic interaction with invasive species

3 = competition

2 = predation

1 = other, vague, general

C.5.5 = Other probable cause, e.g. changing environment w/ differential environmental tolerances

3 - env. changed, natives no longer can survive, invasive take over.

2

1

S1b(28/30). Investigation design

SU A.1.2

3= scientifically accurate - suggests an experimental design that directly addresses identified problem suggests data. Explanation integrates and applies knowledge, controls variables.

2 =developing - employs some methodology, perhaps survey, - suggests data to collected. may mix scientific and non-scientific approaches. may not control variables.

1 = nonscientific - suggests other approach, e.g. reading, talking to people

0= nonsensical - lacks a coherent design or no response.

S1c(28/30). checking results against hypothesis ((interpretation)

SU: A15 - construction of argument

3 = Supported - predicts results based upon design that would support or refute hypothesis

2 = semi-accurate - uses some supports for hypothesis

1 = unsupported, nonscientific -

0= nonsensical - lacks a coherent argument or no response.

S2a(29/31). Problem definition/hypothesis
CU - F52

3= identified problem as possible human environmental impact and suggest probable cause (e.g. A - sewage; B- lawn chemicals).

2 = Other probable cause, e.g. toxins

1 = non-scientific cause.

27 1= correct response, 0 = wrong or NR

- a. fine sediment
- b. aquatic plants
- c. predominant substrate
- d. pools

4 = L3, 3-2 = L2, 1= L1,

S1a(28/30). Problem definition/hypothesis
SU's - A.1.1

3= states reasonable thesis - well defined, makes a prediction stating possible outcomes

2 = states reasonable thesis, no prediction, may be general

1= vague, undefined or non-scientific thesis,

0= nonsensical - lacks a coherent construct or - no response

S1b(28/30).. Investigation design

EU - A.2.2 - purpose of investigation

"because... so that... , to find out if..."

3 - inquiry for verification or exploration

2 - inquiry to test an idea and see if it is right

1. no purpose beyond task/ class expectations

0. nonsensical or no response

S1c(28/30). checking results against hypothesis ((interpretation)

EU A.2.5 - explains rationales

3. uses empirical evidence to justify or evaluate an argument or state position

2. Uses some evidence to support argument

1. does not use evidence or does not justify.

S2a(29/31). Problem definition/hypothesis
SU's - A.1.1

3= states reasonable thesis - well defined, makes a prediction stating possible outcomes

2 = states reasonable thesis, no prediction, may be general

1= vague, general, undefined or non-scientific

thesis, 0= nonsensical - lacks a coherent construct or - no response

S2b(29/31). Investigation design

SU A.1.2

3= scientifically accurate - suggests an experimental design that directly addresses identified problem. Explanation integrates and applies knowledge, controls variables.

2 = semi accurate - employs some methodology, perhaps survey, may mix scientific and non-scientific approaches. may not control variables.

1 = nonscientific - suggests other approach, e.g. reading, talking to people

0= nonsensical - lacks a coherent design or no response.

S2c(29/31) checking results against hypothesis ((interpretation)

SU: A15 - construction of argument

3 = Supported - predicts results based upon design that would support or refute hypothesis

2 = semi-accurate - uses some supports for hypothesis

1 = nonscientific -

0= nonsensical - lacks a coherent argument or no response.

S2d(29/31) Solution

3=Accurate/Elaborate - solution directly addresses identified problem, feasible

2 = semi-accurate, realistic/ moderately complete solution,

1 = vague, general, not feasible or nonscientific - solution address other problems/issues

0= nonsensical - lacks a coherent explanation or no response.

30. (pretest)What issue or question are you most interested in studying during the stream project?

(28(30)). (post-test) What did you find most interesting during the stream project.

31. Concept map

(29(31)) on post test

S2b(29/31) Investigation design

EU - A.2.2 - purpose of investigation

3 - inquiry for verification or exploration

2 - inquiry to test an idea and see if it is right

1. no purpose beyond class expectations

0. nonsensical or no response

S2c(29/31) checking results against hypothesis

((interpretation)

EU A.2.5 - explains rationales

3. uses empirical evidence to justify or evaluate an argument or state position

2. Uses some evidence to support argument

1. does not use evidence or does not justify.

S2d(29/31) Solution

CU - F-5 or F6 - solution addresses human contributions to the problem

3 - scientifically accurate

2. semi accurate

1 - non scientific.

keep list - look for themes

31a - Number of concepts added

31b - Number of links added

explanations (levels 3-1)

if maps warrant - use more sophisticated scoring guide.

APPENDIX H
PRE/POST TEST FORM B

COMMUNITY HIGH SCHOOL

FOUNDATIONS OF SCIENCE - I
Fall Semester Post-test

Student's Name: _____

Period: _____ Block 1 - 8:00 - 9:30 MW (Kathe)
_____ Block 3 - 9:40 - 11:10 MW (Liz)
_____ Block 5 - 12:00 - 1:30 MW (Kathe)
_____ Block 7 - 1:40 - 3:10 MW (Liz)

Instructions: Carefully follow the directions for each set of questions. Your instructors will use this post-test to help assess what you have learned about the stream project, water quality and other environmental issues this semester. Your responses will also be used in the study about what students learn when engaged in extended projects like the study of Traver Creek. Do your very best and answer all of the questions with the attitudes and knowledge you have today.
Thanks!

1. List three rivers/creeks within 50 miles of your school.
 - a.
 - b.
 - c.
2. In what watershed is your school located?
3. Have you ever visited Traver Creek?
4. What are some kinds of activities you like to do in or around a river or creek?
5. How much do you agree with each of the following statements. (Circle one)
 1 = I don't agree at all... 5 = I agree very much... X = not sure/don't know

As long as my drinking water is clean, I don't have to worry about water pollution.	1	2	3	4	5	X
I see no reason for me to do anything about environmental problems since most of these situations will fix themselves.	1	2	3	4	5	X
People these days worry too much about pollution.	1	2	3	4	5	X
The detergent we use in the dishwasher will help clean the water in the river.	1	2	3	4	5	X
As you go down the river/creek, the water is likely to become cleaner.	1	2	3	4	5	X
How we wash our cars has an impact on the river/creek.	1	2	3	4	5	X
Trying to clean up the river/creek just seems too much of a pain.	1	2	3	4	5	X
I didn't create environmental problems, so it's not my responsibility to solve them.	1	2	3	4	5	X
It's hard to find examples where citizen action has really made a difference in the environment.	1	2	3	4	5	X
There are a number of things I can do to help clean up the river/creek.	1	2	3	4	5	X
Once we solve an environmental problem, it seems that we create a new problem.	1	2	3	4	5	X
The government will eventually clean up the river.	1	2	3	4	5	X

It would be embarrassing to be the only one in my school who rode their bike to school.	1	2	3	4	5	X
Shopping malls can contribute to water pollution.	1	2	3	4	5	X
I can easily find the information I need to inform me about environmental issues.	1	2	3	4	5	X
Trying to live in balance with nature isn't possible -- we've altered it too much.	1	2	3	4	5	X

6. Rank the following environmental problems from one to five. Put a one next to the item that you think is the most pressing problem facing Americans today, and a two next to the second most pressing environmental issue, and so forth, until you get to five.

___ water pollution

___ air pollution

___ indoor radon

___ pesticides and chemicals in our food

___ toxic chemicals in our environment (PCB's, dioxin, lead in the soil, etc.)

7. List five (5) things you could personally do to reduce your contribution to water pollution.

a.

b.

c.

d.

e.

8. Into what body of water does Traver Creek empty?

9. List the governmental agencies or public officials who are responsible for Traver Creek?

Multiple choice: Circle the BEST answer.

10. How would you rate the quality of Traver Creek Water?
- awful
 - pretty dirty
 - all right
 - pretty clean
 - great
11. Which pH level (degree of acidity) would be best for the creek?
- 3.5
 - 5.5
 - 7.5
 - 9.5
 - Don't Know
12. Which type of substrate would you expect to find in a pool area in a stream?
- fine silts
 - sand
 - gravel
 - cobbles and boulders
13. Fecal coliform is: (Check [✓] all that apply)
- a chemical from industry that pollutes waterways.
 - another name for untreated sewage.
 - usually a non-point source pollutant in the creek.
 - a bacteria that lives in the guts of mammals.
 - a benthic macro-invertebrate that decomposes organic pollution.
14. Which of these would indicate an improvement in Traver Creek's water quality? (Check [✓] as many as appropriate.) Then choose ONE of the items that you checked and explain why you think that would indicate an improvement in water quality.
- A decrease in the number of fecal coliforms.
 - A rise in the turbidity
 - A shift in the Water Quality Index from 76 to 68.
 - Finding a lot of one kind of insect larvae in the water.
- (Why?)

15. What would happen if extra phosphates enter the creek? Describe this process in terms of photosynthesis, cellular respiration, Dissolved oxygen (DO) and Biochemical Oxygen Demand (BOD). Include any relevant chemical equations in your answer.

16. Use the Periodic Table to complete the following chart.

Atomic number	Elemental name	Symbol	Atomic mass	Number of electrons
19				
		Cl		
	Cadmium			

17. How do each of the following affect the water quality of Traver Creek/Huron River? (makes it worse, makes it better, doesn't affect it.) Choose **TWO** that you think might change and describe what related factors or parameters might also be affected or involved in the change?

Using a high phosphate detergent to wash the family car in the driveway.

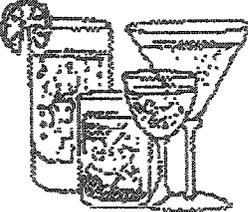
Six weeks of hot summer days without rain.

Separating storm sewers from sanitary sewers.

Restoring and protecting natural wetlands along the stream.

Expanding the Leslie golf course

Multiple Choice: Circle the **BEST** answer.

18. A solution in a dish contains 5 grams of salt dissolved in 100 grams of water. If 50 grams evaporates from the solution, how much salt remains in the dish?
- 0.0 grams
 - 1.0 grams
 - 2.5 grams
 - 5.0 grams
19. Yvette thinks that water will evaporate faster in a windy place than in a still one. She has two identical bowls and a bucket of water. She wants to do an experiment to find out if she is correct. Which of the following should she do?
- Place two bowls with the same amount of water in a warm, windy place.
 - Place a bowl of water in a protected place and a bowl with twice the amount of water in a windy place.
 - Place a bowl of water in a protected place and a bowl with half of the amount of water in a windy place.
 - Place a bowl of water in a protected place and a bowl with the same amount of water in a windy place.
20. The droplets of water on the outside of the glass containers pictured to the right most likely result from which of the following?
- 
- Water is leaking through the container wall.
 - Water in the air outside the container is cooling and changing from vapor to liquid.
 - Air above the ice inside the container is warming and changing from vapor to liquid.
 - Cold air is carrying water from the inside to the outside of the container.
21. Currently, scientists make many predictions about the way solids behave by assuming that the atoms in the solid behave as if they are connected by tiny springs. This means that which of the following is true?
- There must be tiny springs connecting atoms together.
 - Scientists will always find it useful to think of atoms as connected by springs.
 - An intelligent life form on another planet would almost certainly think about solids in the same fashion.
 - Thinking of atoms as being connected by springs constitutes a useful model of solids.
 - Because it is a successful model, the spring model of the atom can be classified as a fact.

22. **Fill in the Blank:** Aquatic insects can be grouped according to how and what they eat. These different "functional feeding" groups reveal the many roles aquatic insects play in stream ecosystems. From the list below, fill in the name of the functional feeding group described on the blanks provided.

- | | |
|----------------------|----------------------|
| Predators | Grazers/Scrapers |
| Gathering Collectors | Filtering Collectors |
| | Shredders |

- a. The net spinning caddisflies (*Hydropsychids*) are found in fast currents where they build their nets to catch tiny minute particles out of the water column. They are an example of _____.

- b. Crane fly larvae (*Tipulids*) like to live in the leaf packs that catch on rocks and other obstructions in the stream. They chomp up the leaves and the bacteria decomposers that cover the leaf surfaces. Crane fly larvae are an example of _____.

- c. Since they feed on other aquatic insects, _____ often have large pincer-like jaws (heiligammles), or spear-like mouth parts (water striders).

Multiple Choice: Choose the Best Answer.

23. Group 7 elements are similar in all of the following ways except:
 - a. they all gain one electron
 - b. they react in similar ways
 - c. they are all non metals
 - d. they all have the same atomic number

24. Which of the following is an example of a heterogeneous mixture?
 - a. chickenbroth
 - b. salt water
 - c. clean air
 - d. rubbing alcohol

25. Which of the following is a physical change?
 - a. rusting steel wool
 - b. a candle melting in the hot sun
 - c. the color change when testing for nitrates
 - d. mixing two liquids, which form a solid precipitate

26. **Metals and Non-metals:**

a. Name the compound that is represented by the chemical formula:

KI _____

MgF₂ _____

Ca(NO₃)₂ _____

b. Write the chemical formula for these compounds.

Lithium chloride _____

Aluminum chloride _____

Magnesium phosphate _____

c. Write the complete equation for this double displacement reaction using chemical symbols. Balance each compound. [You do not need to balance the equation!]

potassium chloride and calcium sulfate

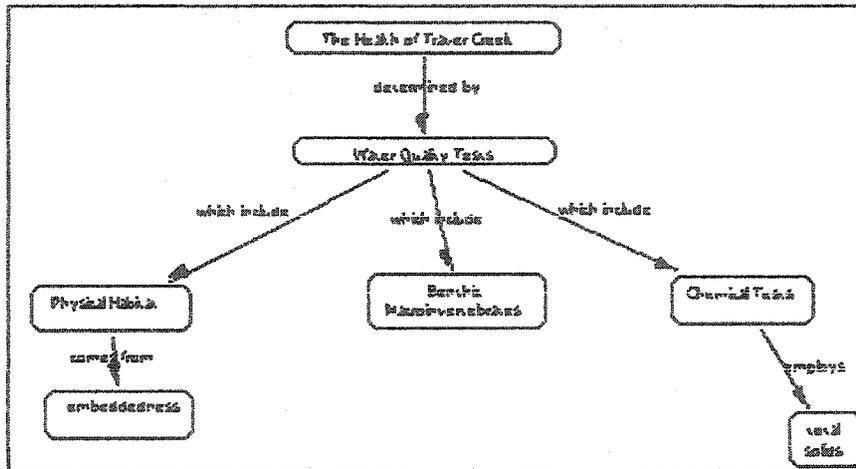
27. **Fill in the Blank:** Write the term from the list below that best fits the phrase in the answer blank provided.

- | | |
|----------------|-----------------------|
| aquatic plants | embeddedness |
| eroding banks | fine sediment |
| flow rate | meters |
| poles | predominant substrate |
| riffles | stable habitat |
| tenths of feet | |

- a. _____ is silt, clay, and sand that covers the little creature's habitat.
- b. Plants found in the water and that provide food and homes for animals are called _____.
- c. The major type of stream bottom cover (like rocks or sand) is called _____.
- d. The areas of the stream that are slow moving, wide and tend to have a lot of sediment deposited in them are called _____.

28. What was the most interesting thing that you discovered, learned or did during the Traver Creek project?

29. Finish the concept map by adding (at least) 4 concepts under benthics, 2 more concepts under physical habitat and 3 more concepts under Water quality tests. Use the following words (or others of your choice) to show the relationships between the concepts: employs, comes from, determines, depends on, is a kind of, for example, needed by, covers, leads to, changes, results in, etc. Also draw lines and arrows to indicate how the concepts under each of the three main assessments are related to each other. Finally write a paragraph explaining the relationships you drew in your map.



30 Scenario 1: Read carefully, then answer the questions below.

A graduate student from the University of Michigan was studying the native mussels and clams in the Huron River. When comparing the 1996 data to past years, the scientist noticed that the numbers and diversity of native clams was greatly reduced. This is a real concern as several of the Huron River's native clams are on the Federal Endangered Species list. In addition, the scientist discovered that the zebra mussel (*Dreissena polymorpha*) had invaded Argo Pond and was found in high numbers.

Zebra mussels are small, fingernail-sized mussels native to the Caspian Sea region of Asia. They were first discovered in the Great Lakes in 1988. Female zebra mussels can produce as many as 1 million eggs per year. These develop into microscopic, free-swimming larvae. At about three weeks, the sand grain-size larvae start to settle and attach to any firm surface. They will cover rock, gravel, metal, rubber, wood, crayfish, native mussels and each other. Zebra mussels filter plankton from the surrounding water, each mussel can filter about one liter of water each day.

a. What do you think is happening to the native mussels and clams?

b. How can you find out if your hypothesis is correct? What will you investigate?

c. How will you use the investigation results to see if your hypothesis is correct?

APPENDIX I

DATA SUMMARY

Key	Expect	0-99	1-1.49	1.5-1.99	2-2.49	2.5-2.99	Σ	
		0-23	1-1.49	1.5-1.99	2-2.49	2.5-2.99		
Conceptual Understandings		No	95	42	21	48	45	95
		Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest
B.1	STRUCTURE OF ATOMS							
B.1.1	Matter is made of minute particles called atoms, and atoms are composed of even smaller components. These components have measurable properties, such as mass and electrical charge. Each atom has a positively charged nucleus surrounded by negatively charged electrons. The electric force between the nucleus and electrons holds the atom together.	E						
		O						
		ns	25					51
		ns	95					95
		L3	21.58					56.25
		L2	3.13					4.17
		L1	11.46					3.13
		L0	53.34					35.46
		nc (1-3)						56(46+10-)
		0						0.000
B.2	STRUCTURE AND PROPERTIES	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest
B.2.2	An element is composed of a single type of atom. When elements are listed in order according to the number of protons (called the atomic number), repeating patterns of physical and chemical properties clearly link up elements with similar properties. This 'Periodic Table' is a consequence of the repeating pattern of outermost electrons and their permitted energies.	E						
		O						
		ns	41					81
		ns	95					95
		L3	19.73					51.21
		L2	2.08					21.35
		L1	20.24					13.93
		L0	57.29					7.39
		nc (1-3)						56(71+8-)
		0						0.000
B.2.5	Solids, liquids, and gases differ in the distances and angles between molecules or atoms and therefore the energy that binds them together. In solids the structure is nearly rigid; in liquids molecules or atoms move around each other but do not move apart; and in gases molecules or atoms move almost independently of each other and are mostly far apart.	E						
		O						
		ns	79	5	9	23		87
		ns	95	23	14	41		95
		L3	47.92			2.08	4.15	54.38
		L2		13.15	14.98	25.25		
		L1	20.21	9.08	2.05	19.22		29.17
		L0	31.38	72.73	81.25	52.17		6.25
		nc (1-3)		12(4-8-)	4(2-0-)	5(1-3-)	41(39-8-)	56(40+16-)
		0		0.388	0.500	1.000		0.005
B.3	CHEMICAL REACTIONS	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest
B.3.1	Chemical reactions occur all around us, for example in health care, cooking, cosmetics, and automobiles. Complex chemical reactions involving carbon-based molecules take place constantly in every cell in our bodies. [See Central Standard C (grades 9-12)]	E						
		O						
		ns	25					30
		ns	95					95
		L3	28.04					52.08
		L2						
		L1						
		L0	73.95					47.92
		nc (1-3)						56(49+17-)
		0						0.000
B.3.3	A large number of important reactions involve the transfer of either electrons (oxidation-reduction reactions) or protons (acid-base reactions) between reacting ions, molecules, or atoms. In other reactions, chemical bonds are broken by heat or light as found in many reactions in our bodies. Radical reactions control many processes such as the prevention of cancer and greenhouse gases in the atmosphere.	E						
		O						
		ns					1.17	
		ns					5.00	
		ns					5.00	
		L3						
		L2						
		L1					2.17	
		L0					10.87	
		L0					85.98	
		nc (1-3)						
		0						

C. CONTENT STANDARD C.14: Science								
C.3. BIOLOGICAL EVOLUTION								
	Reference	Pre-Test	Essay	Report	Model 1	Model 2	Post-Test	Pre/Post
C.3.5. Biological classifications are based on how organisms are related. Organisms are classified into a hierarchy of groups and subgroups based on similarities which reflect their evolutionary relationships. Species is the most fundamental unit of classification.	E							
	O	11.06	1.42	1.75	3.00		1.15	
	99	16	12	20	1		27	
	95	26	12	36	1		26	
	L3			12.82			1.04	
	L2	1.04	11.11	77.17	2.08		2.08	
	L1	15.63	19.36				29.03	
	L0	83.53	73.33	9.09	57.52		71.88	
	ac (a-)		12(11:0-)	1(0:0-)	0		75(0:11-)	55(21:3-)
	S		0.001	1.000			0.000	0.045
C.4. THE INTERDEPENDENCE OF ORGANISMS								
C.4.1. The atoms and molecules on the earth cycle among the living and nonliving components of the biosphere.	E							
	O					1.00		
	99					1.00		
	95					1.00		
	L3							
	L2						2.17	
	L1						97.83	
	L0							
	ac (a-)							
	S							
C.4.2. Energy flows through ecosystems in one direction, from photosynthetic organisms to herbivores to carnivores and decomposers.	E							
	O			2.00	1.87	3.00		
	99			7	9	10		
	95			37	17	20		
	L3							
	L2				31.82	12.90	21.74	
	L1					5.25		
	L0				68.18	21.25	73.26	
	ac (a-)				0	0	7(5: 0-)	
	S						0.063	
C.4.3. Organisms both cooperate and compete in ecosystems. The interrelationships and interdependencies of these organisms may generate ecosystems that are stable for hundreds or thousands of years.	E							
	O			1.5	1.0	1.0		
	99	27		5	1	2	75	
	95	70		24	2	19	84	
	L3	15.83					48.98	
	L2	15.67		22.73		13.21	19.79	
	L1	5.25		4.35	2.05	5.45	10.42	
	L0	51.46		55.18	87.82	73.26	30.83	
	ac (a-)				0(0)	1(0:0-)	15(10:2-)	55(18: 8-)
	S						1.000	0.000
C.4.4. Living organisms have the capacity to produce populations of infinite size, but environments and resources are finite. This fundamental tension has profound effects on the interactions between organisms.	E							
	O			1.75	1.0	1.15		
	99			4	4	8		
	95			15	5	11		
	L3							
	L2				13.64	8.33	13.04	
	L1				4.55		2.17	
	L0				31.82	21.67	64.79	
	ac (a-)				0(0)	0(0)		
	S							
C.4.5. Human beings live within the world's ecosystems. Increasingly, humans modify ecosystems as a result of population growth, technology, and consumption. Human disturbance of habitats through direct harvesting, pollution, atmospheric changes, and other factors is threatening critical global resources, and if not addressed, ecosystems will be irreversibly affected.	E							
	O	1.18				3.00	3.00	
	99	55				4	88	
	95	95				8	95	
	L3	12.50					12.50	
	L2	21.59				5.70	54.17	
	L1	22.82					25.05	
	L0	42.71				51.30	7.29	
	ac (a-)					8(5:2-)	8(5:2-)	55(52: 24-)
	S					0.289	0.289	0.000

C.5.	MASTER, ENERGY, AND ORGANIZATION IN LIVING SYSTEMS	Reference	Pgs/Text	Essay	Report	Model 1	Model 2	Posttest	Pgs/Post
C.5.1	All matter tends toward more disorganized states. Living systems require a continuous input of energy to maintain their chemical and physical organizations. With death, and the cessation of energy input, living systems rapidly disintegrate. [See Underlying Concepts and Processes]	E							
		O	1.00			1.00	1.73	1.71	
		ns	3			1	19	83	
		nv	95			1	35	98	
		L3	3.13			2.02	2.17	3.28	
		L2					28.26	27.06	
		L1					10.87	29.17	
		O	35.65			37.32	39.70	34.38	
		nc (4-)				113 (0-)	110 (0-)	3417 (17-)	96183 (3-)
		P				1.000	1.000	0.064	0.001
C.5.2	The energy for life primarily derives from the sun. Plants capture energy by absorbing light and using it to form strong (covalent) chemical bonds between the atoms of carbon-containing (organic) molecules. These molecules can be used to assemble larger molecules such biological energy including protein, DNA, sugar, and starch. In addition, the energy stored is used because the atoms (chemical energy) can be used as sources of energy.	E							
		O	1.30	2.01	1.75	2.50	1.38	1.55	
		ns	95	2	8	4	14	73	
		nv	95	2	32	8	29	95	
		L3					4.35	3.13	
		L2	3.13	4.44	31.82	8.33	17.39	37.50	
		L1	7.29		9.08		8.70	35.42	
		O	39.38	33.18	39.09	31.57	38.57	23.82	
		nc (4-)		212 (0-)		205	610 (0-)	2817 (11-)	95157 (4-)
		P		0.500		1.000	1.000	0.210	0.001
C.5.4	The complexity and organization of organisms accommodates the need for obtaining, transforming, transporting, releasing, and eliminating the matter and energy used to sustain the organism.	E							
		O	1.35		1.35		1.47	1.42	
		ns	83		18		3	35	
		nv	95		45		6	95	
		L3	14.33					52.08	
		L2	29.17		77.73		4.75	34.38	
		L1	21.35		9.03		2.17	12.50	
		O	32.29				32.48	1.04	
		nc (4-)			45118 (14-)		312 (1-)	519 (0-)	95155 (7-)
		P			0.455		0.500	0.062	0.000
C.5.5	The distribution and abundance of organisms and populations in ecosystems are limited by the availability of matter and energy and the ability of the ecosystem to recycle materials.	E							
		O	1.25	1.57	2.11	1.94	1.35	2.21	
		ns	91	7	21	32	40	93	
		nv	95	7	78	59	77	95	
		L3	2.02		11.82	8.35	4.35	37.90	
		L2	17.71	8.88	53.84	40.00	76.09	47.32	
		L1	33.31	6.57		10.42	6.52	11.48	
		O	46.80	34.44	72.73	33.33	13.04	3.13	
		nc (4-)		715 (1-)	13 (0-)	4817 (7-)	5119 (8-)	74131 (15-)	95173 (4-)
		P		0.215	0.250	1.000	1.000	0.041	0.000
C.5.6	As matter and energy flow through different levels of organization of living systems—cells, organs, organisms, communities—and between living systems and the physical environment, chemical elements are recombined in different ways. Each recombination results in storage and dissipation of energy into the environment as heat. Matter and energy are conserved in each change.	E							
		O	1.33		1.67	1.00	1.33	1.94	
		ns	9		6	4	3	51	
		nv	95		18	9	6	95	
		L3						18.67	
		L2	3.13		18.18		2.17	37.50	
		L1	8.25		9.03	8.33	4.35	30.21	
		O	30.53		77.73	31.57	33.48	15.63	
		nc (4-)			1817 (0-)	0	110 (0-)	512 (3-)	95175 (0-)
		P			0.000		1.000	1.000	0.000

C. 6. THE BEHAVIOR OF ORGANISMS		Reference	Pie Test	Essay	Report	Model 1	Model 2	Posttest	Pie/Post	
C. 6. 2.	Organisms have behavioral responses to external changes and to external stimuli. Responses to external stimuli can result from interactions with the organism's own species and others, as well as environmental changes. These responses can be innate or learned. The broad patterns of behavior exhibited by animals have evolved to ensure reproductive success. Animals often live in unpredictable environments, and so their behavior must be flexible enough to	E								
		O			100	100				
		ns			1	1				
		nr					1			
		L3					100			
		L2			2.72					
		L1								
		O					97.92			
nc a-r					0					
e										
D. CONTENT STANDARD D: Earth and Space Sciences										
D. 1. ENERGY IN THE EARTH SYSTEM		Reference	Pie Test	Essay	Report	Model 1	Model 2	Posttest	Pie/Post	
D. 1. 1.	Earth systems have internal and external sources of energy, both of which create heat. The sun is the major external source of energy. Two primary sources of internal energy are the decay of radioactive isotopes and the gravitational energy from the earth's original formation.	E								
		O		100		100	100	100		
		ns		2		3	15	13		
		nr		2		19	25	24		
		L3					10.43	4.99		
		L2		4.44		22.73	12.50	21.74		
		L1					3.33	2.17		
		O		99.56		77.27	88.78	71.74		
		nc a-r				0	910.00	1214.40		
		e					1.000	1.000		
D. 2. GEOCHEMICAL CYCLES		Reference	Pie Test	Essay	Report	Model 1	Model 2	Posttest	Pie/Post	
D. 2. 1.	The earth is a system containing essentially a fixed amount of each stable chemical atom or element. Each element can exist in several different chemical reservoirs. Each element on earth moves among reservoirs in the solid earth, oceans, atmosphere, and organisms as part of geochemical cycles.	E								
		O		100				100	100	
		ns		34				9	59	
		nr		34				16	55	
		L3		19.42					57.29	
		L2						19.22		
		L1						4.35		
		O		84.36				80.43	42.71	
		nc a-r						15110.00	1517.00	26135.14
		e						0.102	1.000	0.004
D. 2. 2.	Movement of matter between reservoirs is driven by the earth's internal and external sources of energy. These movements are often accompanied by a change in the physical and chemical properties of the matter. Carbon, for example, occurs as carbonate rocks such as limestone, as the atmosphere as carbon dioxide gas, as matter as dissolved carbon dioxide, and as all organisms as complex molecules that control the chemistry of life.	E								
		O				100	100			
		ns					1			
		nr						100		
		L3								
		L2					2.08	2.17		
		L1								
		O					97.92	97.92		
		nc a-r							0	
		e								
D. 3. THE ORIGIN AND EVOLUTION OF THE EARTH SYSTEM		Reference	Pie Test	Essay	Report	Model 1	Model 2	Posttest	Pie/Post	
D. 3. 3.	Interactions among the solid earth, the oceans, the atmosphere, and organisms have resulted in the ongoing evolution of the earth system. We can observe some changes with an earthquake and volcanic eruptions on a human time scale, but many processes such as mountain building and plate tectonics take place over hundreds of millions of years.	E								
		O		1.45	1.14	2.05	2.02	1.83	2.30	
		ns		55	11	12	34	34	55	
		nr		55	11	51	54	54	55	
		L3		11.46			12.54	10.43	2.70	57.29
		L2		21.25	15.56	63.64	52.08	47.83	34.28	
		L1		28.04	3.89	9.09	8.33	17.39	7.29	
		O		31.25	75.86	13.03	29.17	35.80	1.04	
		nc a-r			1113.10	910.00	2013.40	4515.13	62142.13	26172.80
		e			0.039	1.000	0.049	0.098	0.000	0.001

CONTENT STANDARD 1: Science in Personal and Social								
F.1. PERSONAL AND COMMUNITY HEALTH	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest	Pre/Post
F.1.3. Personal choice concerning fitness and health involves multiple factors. Personal goals, peer and social pressures, ethnic and religious beliefs, and understanding of biological consequences can all influence decisions about health practices.	E							
	O		1.30					
	NS		2					
	SP		2					
	LS							
	LI		2.22					
	OS		2.22					
Σ		99.96						
	O							
F.2. POPULATION GROWTH	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest	Pre/Post
F.2.1. Populations grow or decline through the combined effects of births and deaths, and through migration and immigration. Populations can increase through linear or exponential growth, with effects on resource use and environmental pollution.	E							
	O	1.00						
	NS	2			2	8	54	
	SP	95			4	17	95	
	LS						1.04	
	LI	2.08			2.08	15.22	35.54	
	OS	97.92			5.83	82.61	43.73	
Σ (1-7)					114.0-2	212.0-2	171.2-2	951.3-2
Σ					0.125	0.500	0.105	0.000
F.2.3. Populations can reach limits to growth. Carrying capacity is the maximum number of individuals that can be supported in a given environment. The limitation is due to the availability of space, but the number of people is related to resources and the capacity of earth systems to support human beings. Changes in technology can cause significant changes, either positive or negative, in carrying capacity.	E							
	O							
	NS					1		
	SP							
	LS						2.17	
	LI							
	OS						97.81	
Σ (1-7)								
Σ								
F.3. NATURAL RESOURCES	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest	Pre/Post
F.3.1. Human populations use resources in the environment in order to maintain and improve their existence. Natural resources have been and will continue to be used to maintain human populations.	E							
	O		2.00					
	NS		1		5		4	
	SP		1		19		5	
	LS							
	LI		2.22		18.18		9.70	
	OS		97.78		77.27		91.30	
Σ (1-7)				110.2		0		
Σ				1.000				
F.3.3. Humans use many natural systems as resources. Natural systems have the capacity to reuse waste, but that capacity is limited. Natural systems can change to an extent that exceeds the limits of organisms to adapt naturally or humans to adapt technologically.	E							
	O				1.00		1.00	
	NS				15	1	3	
	SP							
	LS						2.17	
	LI				88.18	2.08	2.17	
	OS				31.82	97.92	92.48	
Σ (1-7)				110.1-1	0			
Σ				1.000				

F.5.	NATURAL AND HUMAN-INDUCED HAZARDS	Reference	Pre Test	Essay	Report	Model 1	Model 2	Posttest	Pre/Post
F.5.2	Human activities can enhance potential for hazards. Acquisition of resources, urban growth, and waste disposal can accelerate rates of natural change.	E							
		O	1.20	1.20	2.11		2.13	2.15	
		ns	34	8	9		9	25	
		ns	34	8	30		15	26	
		L3	1.25		2.21		2.17	25.42	
		L2	12.94	6.67	16.36		19.22	14.38	
		L1	41.67	11.11				21.38	
		LO	43.75	32.22	59.09		32.61	3.33	
		nc (4-)		3(4+ 0-)	0		2(2+ 0-)	15(16+ 4-)	98(99+ 7-)
		g		0.125			0.909	0.734	0.000
F.5.3	Climate hazards, such as earthquakes, volcanic eruptions, and severe weather, are rapid and spectacular. But slow and progressive changes that also result in problems for individuals and societies. For example, change in stream channel position, erosion of bridge foundations, subsidence of lands and harbors, coastal erosion, and changing erosion and washing of soil and landscapes can all separately affect society.	E							
		O	1.00		2.00	2.00	2.00	3.00	
		ns	25		2	2	2	35	
		ns	25		4	5	5	59	
		L3	22.56					61.38	
		L2			2.00	6.31			
		L1							
		LO	78.04		30.91	93.45	38.34		
		nc (4-)			4(241-)	(06)	19(2+ 3-)	98(94+ 3-)	
		g			1.000	1.000	1.000	0.000	
F.5.4	Natural and human-induced hazards present the need for humans to assess potential danger and risk. Many changes in the environment designed by humans bring benefits to society, as well as cause risks. Students should understand the costs and benefits of various hazards—ranging from those with minor risk to a few people to major catastrophes with major risk to many people. The scale of events and the accuracy with which humans act	E							
		O		1.50	2.00				
		ns		2	1				
		ns							
		L3							
		L2		2.22	4.55				
		L1		2.22					
		LO		93.36	95.45				
		nc (4-)			0				
		g							
F.5.5	Humans have a major effect on other species. For example, the influence of humans on other organisms occur through land use—which decreases space available to other species—and pollution—which changes the chemical composition of air, soil, and water.	E							
		O		1.57			2.21		
		ns		3			14		
		ns		1			28		
		L3					6.62		
		L2		4.44			23.91		
		L1		1.22					
		LO		93.33			59.57		
		nc (4-)					3(06)		
		g					1.000		
		KEY	Expect	0-35	1-1.45	1.5-1.95	2-2.45	2.5-2.95	3+
		obs		0-35	1-1.45	1.5-1.95	2-2.45		3+
		E	expected						
		O	observed						
		ns	number of students demonstrating that understanding						
		ns	Number of students demonstrating that understanding						
		L3	Level of Understanding 3, scientifically accurate at the standard						
		L2	Level of Understanding 2, partially scientific						
		L1	Level of Understanding 1, non scientific						
		LO	no evidence of understanding						
		nc (4-)	Number of Cases in which significant (number increasing, -number decreasing)						
		g	probability from signed signed test						

BIBLIOGRAPHY

BIBLIOGRAPHY

- AAAS, American Association for the Advancement of Science (1993). Benchmarks for Science Literacy. New York: Oxford University Press.
- Aikenhead, G. S. (1982). Science: A Way of Knowing. In V. N. Wanchoo (Ed.), World Views on Science Education (pp. 206-215). New Delhi, India: Oxford & IBH Publishing Co.
- Aikenhead, G. S., & Ryan, A. G. (1992). The development of a new instrument: "Views on Science – Technology – Society "(VOSTS). Science Education, 76, 477 –491.
- Anderson, C. W., & Roth, K. J. (1989). Teaching for meaningful and self-regulated learning of science. Advances in Research on Teaching: A research annual. J. Brophy, (ed.), 1(1), 265-309.
- Anderson, R. D., & Helms, J. V. (2001). The ideal of standards and the reality of schools: Needed research. Journal of Research in Science Teaching, 38(1), 3-16.
- Archbald, D. A., & Newmann, F. M. (1988). Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School (1 ed.). Reston: National Association of Secondary School Principals.
- Barden, L. M. (1997). Strategies exhibited by high school biology students during laboratories. In A paper presented at the Annual Meeting of the National Association for Research in Science Teaching, . Oakbrook, IL:
- Bhattacharyya, G. K., & Johnson, R. A. (1977). Statistical Concepts and Methods. New York, NY: John Wiley & Sons.
- Blumenfeld, P. C., Marx, R. W., Patrick, H., Krajcik, J., & Soloway, E. (1995). Teaching for Understanding. In B. J. Biddle, T. L. Good, & I. F. Goodson (Eds.), The International Handbook of Teachers and Teaching Dordrecht, The Netherlands: Kluwer.
- Blumenfeld, P. C., Marx, R. W., Patrick, H., Krajcik, J., & Soloway, E. (1997). Teaching for Understanding. In B. J. B. e. al. (Ed.), International Handbook of Teachers and Teaching (pp. 819-878). Netherlands: Kluwer.

- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating Project-Based Learning: Sustaining the Doing, Supporting the Learning. Educational Psychologist, 26(3/4), 369-398.
- Bos, N., Krajcik, J., & Soloway, E. (1997). Student publishing in a WWW digital library: Goals and instructional support. In P. Bell (Chair) Artifact-building in computer learning environments: supporting students' scientific inquiry. (Ed.), A paper presented at the annual meeting of the American Educational Research Association (AERA), . Chicago, IL:
- Brody, M. J. (1990/91). Understanding of Pollution among 4th, 8th, and 11th grade students. The Journal of Environmental Education, 22(3), 24-34.
- Brody, M., Chipman, E., & Marion, S. (1988/89). Student knowledge of scientific and natural resource concepts concerning acidic deposition. The Journal of Environmental Education, 20, 32-42.
- Brown, A. L., & Palincsar, A. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), Knowing, learning, and instruction : essays in honor of Robert Glaser (pp. 393-451). Hillsdale, NJ: L. Erlbaum Associates.
- Brown, A. L., & Campione, J. C. (1990). Communities of learning and thinking, or a context by any other name. Contributions to Human Development, 21, 108-121.
- Brown, A. L., & Campione, J. C. (1991). Restructuring grade school learning environments to promote scientific literacy (Annual Paper No. HBJ 1991). Summer Institute: Council of Chief State School Officers.
- Brown, A. L., & Campione, J. C. (1994). Guided Discovery in a Community of Learners, Chapter 9. In K. McGilly (Ed.), Classroom Lessons: Integrating cognitive theory and classroom practice (pp. 229-270). Cambridge, MA: MIT Press/Bradford Books.
- Brown, A. L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J. C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), Distributed cognitions: Psychological and educational considerations (pp. 188-228). New York, NY: Cambridge University Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of learning. Educational Researcher, 18(1 (Jan-Feb)), 32-42.
- Burger, S. E., & Burger, D. L. (1994). Determining the Validity of Performance-Based Assessment. Educational Measurement: Issues and Practice, 13(1), 9-15.
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. Educational Psychologist, 28(3), 235-251.

- Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). 'An experiment is when you try it and see if it works': A study of grade 7 students' understanding of the construction of scientific knowledge. International Journal of Science Education, 11(Special Issue), 514-529.
- Champagne, A. B., & Newell, S. T. (1992). Directions for Research and Development: Alternative methods of assessing scientific literacy. Journal of Research in Science Teaching, 29(8), 841-860.
- Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.
- Confrey, J. (1990). A review of the research on student conceptions in mathematics, science and programming. In C. Cazden (Ed.), Review of Research in Education (pp. 3-56). Washington, D.C.: American Educational Research Association.
- Connolly, P. (1989). Writing and the ecology of learning. In P. Connolly & T. Vilardi (Eds.), Writing to learn mathematics and science (pp. 1-14). New York, NY: Teachers College, Columbia University,.
- Cooley, W. W., & Klopfer, L. E. (1961). Test on understanding science. Princeton, NJ: Educational Testing Service.
- CoVIS (1996). Project Pedagogy in CoVIS (<http://www.covis.nwu.edu/Geosciences/philosophy/projects.html>) No. The Learning Sciences program of the School of Education and Social Policy at Northwestern University.
- Crawford, B. A. (2000). Embracing the essence of Inquiry: New roles for science teachers. Journal of Research in Science Teaching, 37(9), 916-937.
- CTGV, Cognition and Technology Group at Vanderbilt (1990). Anchored Instruction and its Relationship to Situated Cognition. Educational Researcher, 19(6 (Aug-Sept)), 2-10.
- CTGV, Cognition and Technology Group at Vanderbilt (1992a). The Jasper Series as an example of anchored instruction: Theory, program description, and assessment data. Educational Psychologist, 27(3), 291-315.
- CTGV, Cognition and Technology Group at Vanderbilt (1992b). Anchored Instruction in Science and Mathematics: Theoretical basis, developmental projects, and initial research findings. In R. A. Duschl & R. J. Hamilton (Eds.), Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice (pp. 116-146). Albany, NY: State University of New York (SUNY) Press.
- CTGV, Cognition and Technology Group at Vanderbilt (1996). Scientists in Action. In Nashville, TN: Learning Technology Center At Peabody College of Vanderbilt University.

- Daiute, C. (1989a). Play as thought: thinking strategies of young writers. Harvard Educational Review, 59(Feb.), 1-23.
- Daiute, C. (1989b). Research currents: play and learning to write. Language Arts, 66(6 Oct.), 656-664.
- Daiute, C. (1992). Multimedia composing: extending the resources of kindergarten to writers across the grades. Language Arts, 69(Apr.), 250-260.
- Demastes, S. S., Good, R. G., & Peebles, P. (1995). Students' conceptual ecologies and the process of conceptual change in evolution. Science Education, 79(6), 637-666.
- Dewey, J. (1956). The school and society ; and, The child and the curriculum. Chicago, IL: University of Chicago Press, "A centennial publication of the University of Chicago Press c1990".
- Driver, R., & Easley, J. (1978). Pupils and Paradigms: A review of literature related to concept development in adolescent science students. Studies in Science Education, 5, 61-84.
- Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Successful enculturation: Constructing scientific knowledge in the classroom. Educational Researcher, 23(7), 5-12.
- Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). Children's ideas in science. UK: Milton Keynes, UK: Open University Press.
- Duschl, R. A., & Gitomer, D. H. (1991). Epistemological Perspectives on Conceptual Change: Implications for Educational Practice. Journal of Research in Science Teaching, 28(9), 839-858.
- Edelson, D. C. (1997). Realizing Authentic Science Learning through the Adaptation of Scientific Practice. In K. Tobin & B. Fraser (Eds.), The International Handbook of Science Education The Netherlands: Kulwer.
- Ernest, P. (1995). The One and the Many. In L. Steffe & J. Gale (Eds.), Constructivism in Education (pp. 459-486). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fellows, N. J. (1994). A Window into Thinking: Using student writing to understand conceptual change in science learning. Journal of Research in Science Teaching, 31(9), 981-1001.
- Fitzpatrick, R., & Morrison, E. J. (1971). Performance and Product Evaluation. In R. L. Thorndike (Ed.), Educational Measurement (pp. 237-270). Washington, D.C.: American Council on Education.
- Freisinger, R. A. (1980). Cross-disciplinary writing workshops: Theory and practice. College English, 42, 154-166.

- Gallagher, J. J. (2000). Teaching for understanding and application of science knowledge. School Science and Mathematics, 100(6), 310-318.
- Gallagher, J., & Tobin, K. (1987). Teacher management and student engagement in high school- science. Science Education, 71 (4), 535-555.
- Gallas, K. (1991). Arts as Epistemology: Enabling children to know what they know. Harvard Educational Review, 61(1), 40-50.
- Gardner, H. (1983). Frames of Mind: the theory of multiple intelligences. New York, NY: Basic Books.
- Gardner, H. (1991). The Unschooled Mind: How children think and how schools should teach. New York, NY: Basic Books.
- Gilbert, J. K., & Watts, D. M. (1983). Concepts, Misconceptions and Alternative Conceptions: Changing perspectives in science education. Studies in Science Education, 10, 61-98.
- Gilbert, S. W. (1991). Model building and a definition of science. Journal of Research in Science Teaching, 28(1), 73-79.
- Glaser, R. (1984). Education and thinking: The role of knowledge. American Psychologist, 39(2), 93-104.
- Goldberg, M. R. (1992). Expressing and Assessing Understanding Through the Arts. Phi Delta Kappan, 73(Apr), 619-623.
- Gomez, L. M., & Gordin, D. N. (1995). A Case Study of Open-Ended Scientific Inquiry in a Technology-Supported Classroom. In J. Greer (Ed.), Proceedings of AI-Ed '95, Seventh World Conference on Artificial Intelligence in Education, (pp. 17-24, <http://www.covis.nwu.edu/Papers/Tech-class.html>). Charlottesville, VA: Association for the Advancement of Computing in Education.
- Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. Journal of Research in Science Teaching, 28(9), 799-822.
- Haertel, E. H. (1991). Form and function in assessing science education. In G. Kulm & S. M. Malcolm (Eds.), Science Assessment in the Service of Reform (pp. 133-245). Washington, D.C.: American Association for the Advancement of Science.
- Haney, W., & Madaus, G. (1989). Searching for Alternatives to Standardize Tests: Whys, Whats and Whithers. Phi Delta Kappan, 70(9), 683-687.
- Hestenes, D. (1992). Modeling games in the newtonian world. American Journal of Physics, 60(8), 732-748.

- Hogan, K., & Fisherkeller, J. (1996). Representing students' thinking about nutrient cycling in ecosystems: Bidimensional coding of a complex topic. Journal of Research in Science Teaching, 33(9), 941-970.
- Huebel-Drake, M., Finkel, L., Stern, E., & Mouradian, M. (1995). Planning a Course for Success: Using an integrated curriculum to prepare students for the twenty-first century. Science Teacher, 62(7), 18-21.
- Hulland, C., & Mumby, H. (1994). Science, Stories, and Sense-making: a comparison of qualitative data from a wetlands unit. Science Education, 78(2), 117-136.
- Jackson, S. (1995). Model-It. In Ann Arbor, MI: HI-CE Research Group, Elliot Soloway, Director, The Regents of the University of Michigan.
- Jackson, S. L., Krajcik, J., & Soloway, E. (in press (1998)). The design of guided learner-adaptable scaffolding in interactive learning environments. Submitted to CHI '98.
- Jackson, S. L., Stratford, S. J., Krajcik, J. S., & Soloway, E. (1996a). A Learner-Centered Tool for Students Building Models. Communications of the ACM, 39(4), 48-49.
- Jackson, S. L., Stratford, S. J., Krajcik, J. S., & Soloway, E. (1996b). Making System Dynamic Modeling Accessible to Pre-College Science Students. Interactive Learning Environments, 4(3), 233-257.
- Jackson, S. L., Stratford, S. J., Krajcik, J. S., & Soloway, E. (in press (1996c)). Model-It: A Case Study of Learner-Centered Design Software for Supporting Model Building. .
- Jackson, S., Krajcik, J., Soloway, E. (2001). Model-It: A Design Retrospective. In Jacobson, M. and Kozma, R (Eds.), Advanced Designs For The Technologies Of Learning: Innovations in Science and Mathematics Education. Hillsdale, NJ: Erlbaum.
- Jones, M. G. (1994). Assessment Potpourri. Science & Children, 32(2), 14-17.
- Jones, R. A. (1985). Research Methods in the Social and Behavioral Sciences. Sunderland, MA: Sinauer Assoc., Inc.
- Kafai, Y., & Resnick, M. (Eds.). (1996). Constructionism in Practice: Designing, thinking and learning in a digital world. Mahwah, NJ: Lawrence Erlbaum Assoc.
- Kendall, J. S., & Marzano, R. J. (1996). CONTENT KNOWLEDGE: A Compendium of Standards and Benchmarks for K-12 Education. Mid-continent Regional Educational Laboratory (McREL) (<http://www.mcrel.org/standards-benchmarks/>) and Association of Supervision and Curriculum Development (ASCD).

- Keys, C. W. (1994). The development of scientific reasoning skills in conjunction with collaborative writing assignments: An interpretive study of six ninth-grade students. Journal of Research in Science Teaching, 31(9), 1003-1022.
- Keys, C. W., & Bryan, L. A. (2001). Co-Constructing Inquiry-based science with teachers: Essential research for lasting reform. Journal of Research in Science Teaching, 38(6), 631-645.
- Kilpatrick, W. H. (1918). The Project Method. Teachers College Record, pp 319-335
- Kimball, M. E. (1967 –1968). Understanding the nature of science: A comparison of scientists and science teachers. Journal of Research in Science Teaching, 5, 110 –120.
- Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentations: a developmental study. Cognitive Psychology, 25:111-146.
- Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., & Soloway, E. (1994). A Collaborative Model for Helping Science Teachers Learn Project-Based Instruction. Elementary School Journal, 94(5 (May)), 483-97.
- Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., & Soloway, E. (2000). Instructional, Curricular and Technological Supports for Inquiry in Science Classrooms. In J. Minstrell & E. H. VanZee (Eds.), Inquiring into Inquiry Learning and Teaching in Science. Washington, DC: American Association for the Advancement of Science. ., pgs. 283 –315
- Krajcik, J. S., Simmons, P. E., & Lunetta, V. N. (1988). A research strategy for the dynamic study of students' concepts and problem solving strategies using science software. Journal of Research in Science Teaching, 25(2), 147-155.
- Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J., & Soloway, E. (1998). Middle school students' initial attempts at inquiry in project-based science classrooms. Journal of the Learning Sciences. McCutchan Publishers. 7, 313-350.
- Krajcik, J., Czerniak, C., & Berger, C. (1999). Teaching Children Science: A project-based approach (1st ed.). Boston, MA: McGraw-Hill College.
- Kulm, G., & Malcom, S. (Eds.). (1991). Science Assessment in the Service of Reform. Washington, D.C.: American Association for the Advancement of Science.
- Lamon, M., Secules, T., Pelrosino, A. J., Hackett, R., Bransford, J. D., & Goldman, S. R. (1996). Schools for Thought: Overview of the project and lessons learned from one of the sites. In L. Schauble & R. Glaser (Eds.), Contributions of Instructional Innovation to Understanding Learning Hillsdale, NJ: Lawrence Erlbaum.

- Langer, J. A., & Applebee, A. N. (1987). How writing shapes thinking : a study of teaching and learning (NCTE research report No. 22). Urbana, IL: National Council of Teachers of English.
- Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1995). Children's ideas about ecology 1: Theoretical background, design and methodology. International Journal of Science Education, *17*(6), 721-732.
- Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996a). Children's ideas about ecology 2: Ideas found in children aged 5-16 about the cycling of matter. International Journal of Science Education, *18*(1), 19-34.
- Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996b). Children's ideas about ecology 3: Ideas found in children aged 5-16 about the interdependency of organisms. International Journal of Science Education, *18*(2), 129-141.
- Lee, O., & Brophy, J. (1996). Motivational patterns observed in sixth-grade science classrooms. Journal of Research in Science Teaching, *33*(3), 303-318.
- Lehrer, R. (1993). Authors of Knowledge: Patterns of hypermedia design. Chapter 7. In S. P. Lajole & S. J. Derry (Eds.), Computers as Cognitive Tools (pp. 197-227). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Linn, M. C., & Songer, N. B. (1991a). Cognitive and Conceptual Change in Adolescence. American Journal of Education, *99*(4), 379-417.
- Linn, M. C., & Songer, N. B. (1991b). Teaching Thermodynamics to Middle School Students: What Are Appropriate Cognitive Demands? Journal of Research in Science Teaching, *28*(10), 885-918.
- Linn, M. C., Songer, N. B., Lewis, E. L., & Stern, J. (1990). Using technology to teach thermodynamics: Achieving integrated understanding. In D. L. Ferguson (Ed.), Advanced Technologies in the Teaching of Mathematics and Science Berlin: Springer-Verlag.
- Linn, M.C. and Hsi, S. (2000) Computers, Teachers, Peers: Science Learning Partners. Lawrence Erlbaum Associates. Mahwah, NJ.
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, Performance-Based Assessment: Expectations and validation criteria. Educational Researcher, *20*(8), 15-21.
- Madaus, G. F. (1994). A technological and historical consideration of equity issues associated with proposal to change the nation's testing policy. Harvard Educational Review, *64*(1), 76-95.
- Madaus, G. F., & Tan, A. G. A. (1993). The growth of assessment. Yearbook (Association for Supervision and Curriculum Development), 53-79.

- Maeroff, G. I. (1991). Assessing Alternative Assessment. Phi Delta Kappan, 73(4), 272-281.
- Magnusson, S. J., Templin, M., & Boyle, R. A. (1997). Dynamic science assessment: A new approach for investigating conceptual change. The Journal of the Learning Sciences, 6(1), 91-142.
- Maimon, E. P., Belcher, G. L., Hearn, G. W., Nodine, B. F., & O'Connor, F. W. (1981). Writing in the Arts and Sciences. Boston, MA: Little, Brown and Company, Inc.
- Marx, R. W., Blumenfeld, P. C., Krajcik, J. S., & Soloway, E. (1997). Enacting Project-Based Science. The Elementary School Journal, 97(4), 341-358.
- McGilly, K. (1994). Cognitive Science and Educational Practice: An introduction. In K. McGilly (Ed.), Classroom Lessons: Integrating cognitive theory and classroom practice. Cambridge, MA: The MIT Press.
- McGinley, W. J., & Madigan, D. (1990). The research "story": A forum for intergrating reading, writing, and learning. Language Arts, 67(Sept.), 474-483
- MDOE, Michigan Department of Education (1991). Michigan Essential Goals and Objectives for Science Education K-12: New Directions for Science Education in Michigan. (Educational Policy No. Michigan State Board of Education, Lansing, Michigan.
- Mehrens, W. A. (1992). Using Performance Assessment for Accountability Purposes. Educational Measurement: Issues and Practice, 11(1), 3-9, 30.
- Meichtry, Y. J. (1992). Influencing student understanding of the nature of science: Data from a case of curriculum development. Journal of Research in Science Teaching, 29, 389-407.
- Messick, S. (1989). Meaning and Values in Test Validation: The science and ethics of assessment. Educational Researcher, 18(2), 5-11.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. Educational Researcher, 23(2), 13-23.
- Metcalf, S. J. Krajcik, J., & Soloway, E. (2000) MODEL-IT : a design retrospective in Innovations in science and mathematics education : advanced designs for technologies of learning, Jacobson, M. J. and Kozma, R.B. Eds. L. Erlbaum, : Mahwah, N.J
- Metcalf, Shari Jackson (1999) The design of guided learner-adaptable scaffolding in interactive learning environments. . Doctor of Philosophy (Computer Science), University of Michigan.
- Miller, M. D., & Legg, S. M. (1993). Alternative Assessment in a High-Stakes Environment. Educational Measurement: Issues and Practice, 12(2), 9-15.

- Mistler-Jackson, M., & Songer, N. B. (2000). Student motivation and Internet technology: are students empowered to learn science? Journal of Research in Science Teaching, 37(5), 459-479.
- Mitchell, M. K., & Stapp, W. B. (1994). Field Manual for Water Quality Monitoring (6 ed.). Dexter, MI: Thomson-Shore, Inc.
- Mitman, A. L., Mergendoller, J. R., Marchman, V. A., & Packer, M. J. (1987). Instruction addressing the components of scientific literacy and its relation to student outcomes. American Educational Research Journal, 24(4), 611-633.
- Moss, P. A. (1994). Can there be validity without reliability? Educational Researcher, 23(2), 5-12.
- NESIC, National Education Standards and Improvement Council (1993). Promises to keep: Creating high standards for American students. Report on the review of educational standards from the Goals 3 and 4 Technical Planning Group to the National Educational Goals Panel (Malcom Report No. Washington, DC: National Goals Panel.
- Nickerson, R. S. (1995). Can technology help teach for understanding. In D. N. Perkins, J. L. Schwartz, M. M. West, & M. S. Wiske (Eds.), Software Goes to School: Teaching for understanding with new technologies (pp. 7-22). New York, NY: Oxford University Press.
- Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. New York, NY: Cambridge University Press.
- NRC, National Research Council (1996). National Science Education Standards No. National Academy of Sciences.
- NSTA, National Science Teachers Association, Aldridge, B. G., & Strassenburg, A. A. (Eds.). (1995). Scope, sequence, and coordination of national science education content standards: An addendum to the content core based on the 1994 draft national science education standards. Arlington, VA: National Science Teachers Association (NSTA).
- NSTA, National Science Teachers Association, Pearsall, M. K. (Ed.). (1993). Scope, sequence, and coordination of secondary school science: The content core. A guide for curriculum designers. Washington, DC: National Science Teachers Association (NSTA).
- O'Loughlin, M. (1992). Rethinking Science Education: Beyond Piagetian Constructivism Toward a Sociocultural Model of Teaching and Learning. Journal of Research in Science Teaching, 29(8), 791-820.

- O'Neill, D. K., & Gomez, L. (1994). The Collaboratory Notebook: A distributed knowledge-building environment for project-enhanced learning. In T. Ottmann & I. Tomek (Eds.), Educational Multimedia and Hypermedia, 1994: Proceedings of Ed-Media '94, (pp. <http://www.covis.nwu.edu/Papers/edmedia94.html>). Charlottesville, VA: AACE.
- Osborne, R., & Freyberg, P. (1985). Learning in Science: The Implications of Children's Science. Auckland, NZ: Heinemann Education.
- Palincsar, A. S., Anderson, C. W., & David, Y. M. (1993). Pursuing scientific literacy in the middle grades through collaborative problem solving. The Elementary School Journal, 93(5), 643-658.
- Papert, S. (1991). Situating constructionism. In I. Harel & S. Papert (Eds.), Constructionism: Research reports and essays, 1985-1990 (pp. 1-11). Norwood, NJ: Ablex Pub. Corp.
- Papert, S. (1993). Literacy and Letteracy in the Media Ages. Wired magazine.
- Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1997). Building functional models: designing an elbow. Journal of Research in Science Teaching, 34(2), 125-143.
- Perkins, D. N. (1986). Knowledge as Design. Hillsdale, NJ: Lawrence Erlbaum Assoc. Publ.
- Perkins, D. N. (1991). Educating for Insight. Educational Leadership, 49(Oct), 4-8.
- Perkins, D. N. (1992). Smart Schools: from training memories to educating minds. New York, NY: Free Press.
- Perkins, D. N., & Simmons, R. (1988). Patterns of misunderstanding: an integrative model for science, math, and programming. Review of Educational Research, 58(Fall), 303-326.
- Perkins, D. N., Crismond, D., Simmons, R., & Unger, C. (1995). Inside Understanding. In D. N. Perkins, J. L. Schwartz, M. M. West, & M. S. Wiske (Eds.), Software Goes to School: Teaching for understanding with new technologies New York, NY: Oxford University Press.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211-227.
- Price, S., & Hein, G. E. (1994). Scoring Active Assessment. Science & Children, 32(2), 26-29.

- Raizen, S. A., Baron, J. B., Champagne, A. B., Haertel, E., Mullis, I. V. S., & Oaks, J. (1989). Assessment in Elementary School Science Education (Report No. The National Center for Improving Science Education).
- Resnick, L. B. (1987). Education and learning to think. Washington, D.C.: National Academy Press.
- Roth, W.-M. (1994). Experimenting in a constructivist high school physics laboratory. Journal of Research in Science Teaching, *31*(2), 192-223.
- Roth, W.-M. (1995). Authentic School Science: Knowing and learning in open-inquiry science laboratories. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Roth, W.-M., & Roychoudhury, A. (1992). The social construction of scientific concepts or the concept map as conscription device and tool for social thinking in high school science. Science Education, *76*(5), 531-557.
- Rubba, P. A., & Anderson, O. (1978). Development of an instrument to assess secondary school students' understanding of the nature of scientific knowledge. Science Education, *2*, 449-458.
- Rutherford, F. J., Ahlgren, A., & the American Association for the Advancement of Science S (1990). Science for All Americans. New York, NY: Oxford University Press.
- Rye, J. A., & Rubba, P. A. (1996). An Exploratory Study of the Concept Map as a Tool to Facilitate the Externalization of Students' Understandings about Global Atmospheric Change in the Interview Setting. In J. (. Thomas R. Koballa (Ed.), A paper presented at the Annual Meeting of the National Association for Research in Science Teaching, . St. Louis, MO: University of Georgia.
- Salomon, G., & Perkins, D. N. (1989). Rocky Roads to Transfer: Rethinking mechanisms of a neglected phenomenon. Educational Psychologist, *24*(2), 113-142.
- Salomon, G., Perkins, D., & Globerson, T. (1991). Partners in cognition: Extending human intelligences with intelligent technologies. Educational Researcher, *20*(3), 2-9.
- Scardamalia, M., & Bereiter, C. (1996). Student Communities for the Advancement of Knowledge. Communications of the ACM, *39*(4), 36-37.
- Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' Understanding of the Objectives and Procedures of Experimentation in the Science Classroom. The Journal of the Learning Sciences, *4*(2), 131-166.

- Schauble, L., Klopfer, L. E., & Raghaven, K. (1991). Students' transition from an engineering model to a science model of experimentation. Journal of Research in Science Teaching, 28(9), 859-882.
- Schneider, R.M., Krajcik, J., Marx, R., & Soloway, E. (2002). Performance of students in project-based science classrooms on a national measure of science achievement. Journal of Research in Science Teaching. 39(5) 410-422.
- Schoenfeld, A. H. (1985). Mathematics, technology, and higher order thinking. In R. S. Nickerson & P. P. Zoghbiates (Eds.), Technology and education: Looking toward 2020 (pp. 67-96). Hillsdale, NJ: Erlbaum.
- Schwab, J. J. (1964). Structure of the Disciplines: Meanings and significances. In G. W. Ford & L. Pugno (Eds.), The Structure of Knowledge and the Curriculum (pp. 1-30). Chicago, IL: Rand McNally & Co.
- Shavelson, R. J., & Baxter, G. P. (1992a). Linking Assessment with Instruction. In F. K. Oser Andreas Dick and Jean-Luc Patry (Ed.), Effective and Responsible Teaching: The New Synthesis San Francisco: Jossey-Bass Publishers.
- Shavelson, R. J., & Baxter, G. P. (1992b). What We've Learned About Assessing Hands-On Science. Educational Leadership, 49(8), 20-25.
- Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance Assessments: Political rhetoric and measurement reality. Educational Researcher, 21(4), 22-27.
- Smith, M. L., & Glass, G. V. (1987). Research and Evaluation in Education and the Social Sciences. Boston, MA: Allyn and Bacon.
- Solomon, J., Scott, L., & Duveen, J. (1996). Large-scale exploration of pupils' understanding of the nature of science in Great Britain. Science Education, 80, 493 -508.
- Songer, N. B. (1996). Exploring Learning Opportunities in Coordinated Network-Enhanced Classrooms: A case of Kids as Global Scientists. The Journal of the Learning Sciences, 5(4), 297-327.
- Songer, N. B. (1997). Can Technology Bring Students Closer to Science? In K. Tobin & B. Fraser (Eds.), The International Handbook of Science Education (pp. 18). The Netherlands: Kluwer.
- Songer, N. B., & Linn, M. C. (1991). How Do Students' Views of Science Influence Knowledge Integration? Journal of Research in Science Teaching, 28(9), 761-784.
- Spitulnik, M. W. (1995). Students Modeling Concepts and Conceptions: What connections do they make? In Paper presented at the National Association for Research in Science Teaching Annual Meeting. San Francisco. April 21-25, 1995. .

- Spitulnik, M. W., Stratford, S., Krajcik, J., & Soloway, E. (1996). Using Technology to Support Students' Artifact Construction in Science. In International Handbook of Science Education Netherlands: Kluwer Publishers.
- Spoehr, K. T. (1994). Enhancing the Acquisition of Conceptual Structures through Hypermedia. Chapter 4. In K. McGilly (Ed.), Classroom Lessons: Integrating cognitive theory and classroom practice (pp. 75-101). Cambridge, MA: The MIT Press.
- Stapp, W. B., & Mitchell, M. K. (Eds.). (1995). Field Manual for Global Low-Coast Water Quality Monitoring (1 ed.). Dexter, MI: Thomson-Shore, Inc.
- Stratford, S. J. (1996). Investigating processes and products of secondary science students using dynamic modeling software. Doctor of Philosophy (Education), University of Michigan.
- Stratford, S. J., Krajcik, J., & Soloway, E. (1997a). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing and testing models of stream ecosystems. In A paper presented at the annual meeting of the American Educational Research Association (AERA), . Chicago, IL:
- Stratford, S. J., Krajcik, J., & Soloway, E. (1997b). Technological Artifacts Created by Secondary Science Students: Examining Structure, Content, and Behavior of Dynamic Models. In A paper presented at the Annual Meeting of the National Association for Research in Science Teaching (NARST), . Oakbrook, IL:
- Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice (pp. 147-176). Albany, NY: State University of New York Press.
- Talsma, V. L. (1992). The Interactive Rouge River Water Quality Monitoring Program: An Evaluation of an Environmental Education Program. Master of Science - Natural Resources, The University of Michigan.
- Tamir, P. (1994). Israeli students' conceptions of science and views about the scientific enterprise. Research in Science & Technological Education, 12, 99 -116.
- Tobin, K., & Gallagher, J. (1987). What happens in high school science classrooms? Journal of Curriculum Studies, 19(6), 549-560.
- Treagust, D. F., Jacobowitz, R., Gallagher, J. L., & Parker, J. (2001). Using assessment as a guide in teaching for understanding: A case study of a middle school science class learning about sound. Science Education, 85, 137-157.
- Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. Canadian Journal of Fisheries and Aquatic Science, 37, 130-137.

- Wallace, J. D., & Mintzes, J. J. (1990). The Concept Map as a Research Tool: Exploring Conceptual Change in Biology. Journal of Research in Science Teaching, 27(10), 1033-1052.
- Wandersee, J. H., & Fisher, K. M. (1995). A semantic network-based analysis of the AAAS Benchmarks for biology. In Annual Meeting of the American Educational Research Association, . San Francisco, CA:
- Webster's (1985). Webster's Ninth New Collegiate Dictionary. Springfield, MA: Merriam-Webster Inc.
- Welch, W. W., & Pella, M. D. (1967 –1968). The development of an instrument for inventorying knowledge of the process of science. Journal of Research in Science Teaching, 5, 64 –68.
- White, R., & Gunstone, R. (1992). Probing Understanding. London • New York • Philadelphia: The Falmer Press.
- Whitehead, A. N. (1929). The Aims of Education. Cambridge, UK: Cambridge University Press.
- Wiggins, G. (1989a). A True Test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70(9), 703-713.
- Wiggins, G. (1989b). Teaching to the (authentic) test. Educational Leadership, 48(5), 18-25.
- Wiggins, G. (1991). Standards, not standardization: Evoking quality student work. Educational Leadership, 48(5), 18-25.
- Wiggins, G. (1993). Assessment: Authenticity, Context, and Validity. Phi Delta Kappan, 75(3), 200-214.
- Wiley, D. E., & Haertel, E. H. (1996). Extended assessment tasks: Purposes, definitions, scoring and accuracy. In R. Mitchell & M. B. Kane (Eds.), Implementing performance assessment : promises, problems, and challenges Mahwah, NJ: L. Erlbaum Associates.
- Wilkinson, L., & Bjercknes, M. (1989). MYSTAT 1.1. In Evanston, IL: SYSTAT, Inc.
- Winn, W. (1993). Instructional Design and Situated learning: Paradox or Partnership? Educational Technology, 33(3), 16-21.
- Wisnudel, M. (1994). Constructing hypermedia artifacts in math and science classrooms. Journal of Computers in Mathematics and Science Teaching, 13(2), 127-145.
- Wisnudel-Spitulnik, M., Stratford, S., Krajcik, J., & Soloway, E. (1996). Using Technology to Support Students' Artifact Construction in Science. In International Handbook of Science Education Netherlands: Kluwer Publishers