INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

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 bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI 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.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
THE PROBLEM-SOLVING APPROACH TO PROGRAM EVALUATION:
DEVELOPMENT AND APPLICATION IN A MATHEMATICS CONTEXT

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Kelly M. Costner, M.A., M.A.

* * * * *

The Ohio State University
2002

Dissertation Committee:
Professor Sigrid Wagner, Adviser
Professor Douglas T. Owens
Professor Diane W. Birckbichler

Approved by:

Sigrid Wagner
Adviser
College of Education
ABSTRACT

This study developed and piloted the Problem-Solving Approach to program evaluation, which involves the direct application of the problem-solving process as a metaphor for program evaluation. A rationale for a mathematics-specific approach is presented, and relevant literature in both program evaluation and mathematics education is reviewed.

The Problem-Solving Approach was piloted with a high-school level integrated course in mathematics and science that used graphing calculators and data collection devices with the goal of helping students to gain better understanding of relationships between mathematics and science. Twelve students participated in the course, which was co-taught by a mathematics teacher and a science teacher. Data collection for the evaluation included observations, a pre- and posttest, student questionnaires, student interviews, teacher interviews, principal interviews, and a focus group that involved both students and their teachers.

Results of the evaluation of the course are presented as an evaluation report. Students showed improvement in their understandings of mathematics-science relationships, but also showed growth in terms of self-confidence, independence, and various social factors that were not expected outcomes. The
teachers experienced a unique form of professional development by learning and relearning concepts in each other's respective fields and by gaining insights into each other's teaching strengths.

Both the results of the evaluation and the evaluation process itself are discussed in light of the proposed problem-solving approach. The use of problem solving and of specific problem-solving strategies was found to be prevalent among the students and the teachers, as well as in the activities of the evaluator. Specific problem-solving strategies are highlighted for their potential value in program evaluation situations. The resulting Problem-Solving Approach, revised through the pilot application, employs problem solving as a recursive process at three interconnected levels: the evaluation level (where the evaluator is problem solver), the program level (where teachers or other program administrators are problem solvers), and the mathematical task level (where students are problem solvers).
To my wife, Beth, who defended her dissertation on the same day, and whose support for me was unfailing. To my parents, Flay and Doris, who have seen me off to school practically every year of my life. To the memory of my grandfather, John L. Sain (1901-2001), who always wanted me to be "caught up" on my work—as of the day and moment that I submit this document, I am.
ACKNOWLEDGMENTS

You know how it is. You undertake a project such as this thinking it's all yours—your hard work, your thought processes, your writing, and your product. But in reality, all those around you help shape your thoughts and opinions, and all your past experiences and contacts influence where you are, what you think, and what you've become.

I would like to thank several of my high school teachers: Judy Tysinger and Barbara Murray, for making mathematics fun for me, and for being models of the caring teacher I hoped to become; and Dixie Dellinger, who taught me how to do research using qualitative methods before I even knew what "qualitative" (or research, for that matter) was. I would also like to thank my undergraduate mentor, Dr. Jane T. Mitchell, who sent me to Ohio State—she was right for doing so, even though I didn't go into foreign language education as she had planned.

My experiences in program evaluation were instrumental in setting my mind to pursue this project, and for those experiences I am grateful to all the people I worked with—especially Diane W. Birckbichler, who also served on my dissertation committee, Jean-Louis P. Dassier, and Hiroaki Kawamura.
Finally, I would like to thank those who were directly involved with this project: my adviser and committee members for ideas, input, and multiple readings of an unconventional project; the teachers and students I worked with at Valley Circle High School (a pseudonym, of course); Michelle K. Reed for her proofreading efforts and unbridled use of The Red Pen; Marlena F. Herman for shuttling documents, books, and messages back and forth from one person or place to another; Mike Kositzke for his assistance with computer access and difficulties; Patsy and Waller Cooper for welcoming me and my wife into their home on our many trips back to Columbus; and most of all my wife, Beth, who knew just what to do when it came to the "technicalities" of such a project: tables, figures, page numbers, sections, formatting, graphics—and me.
VITA

July 13, 1968................Born – Shelby, North Carolina

1990........................B.A. in Mathematics and French
The University of North Carolina at Greensboro

1990-1993 ................Mathematics and French Teacher
Central Davidson High School
Lexington, North Carolina

1993-1995 ...............Graduate Teaching Associate
Department of French and Italian
The Ohio State University

1995........................M.A. in French
M.A. in Mathematics Education
The Ohio State University

1995-2000 ..............Graduate Research and Administrative Associate
The Ohio State University

2000 – present ..........Instructor of Mathematics Education
Thayer School of Education, Wingate University
Wingate, North Carolina

FIELDS OF STUDY

Major Field: Education
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td>CHAPTER RATIONALE</td>
<td>1</td>
</tr>
<tr>
<td>1. What is Program Evaluation?</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Assessment and Program Evaluation</td>
<td>4</td>
</tr>
<tr>
<td>1.2 The Problem: A Lack of Program Evaluation Models Specific to Mathematics Education</td>
<td>7</td>
</tr>
<tr>
<td>1.2.1 Existing Mathematics Program Evaluation Guides</td>
<td>7</td>
</tr>
<tr>
<td>1.2.2 Need for a Mathematics-Specific Approach</td>
<td>10</td>
</tr>
<tr>
<td>1.3 Research Questions</td>
<td>12</td>
</tr>
</tbody>
</table>

viii
2. REVIEW OF RELATED LITERATURE ............................................................13

Generic Approaches to Program Evaluation .............................................13
    The Positivist View ...........................................................................14
Beyond the Positivists........................................................................15
    Goal-Oriented Approach ..................................................................15
    Decision-Focused Approach ...........................................................15
    User-Oriented Approach ..................................................................15
    Responsive Approach ......................................................................16
An Alternative Continuum ......................................................................17
    Expertise-Oriented Approach ...........................................................17
    Adversary-Oriented Approach ............................................................18
Stufflebeam's Presentation .....................................................................19
The Accompanying Paradigm Shift ..........................................................21
    The Responsive Approach ................................................................22
    Focus on Changing Issues .................................................................22
    Fluidity of Evaluation Events ..........................................................23
    Extension to Participatory Approaches .............................................24
    Influence on the Problem-Solving Approach .....................................24
The Connoisseurship Approach ..............................................................25
    Essentials of Connoisseurship .........................................................25
    An Analogy ....................................................................................26
    Application to Educational Evaluation ..............................................27
    Influence on the Problem-Solving Approach .....................................27
The Fourth Generation Approach .............................................................27
    The First Generation: Measurement ..................................................28
    The Second Generation: Description ................................................28
    The Third Generation: Judgment .......................................................29
    Problems of the First Three Generations ..........................................29
    The Fourth Generation: Responsive Constructivist Evaluation ...........30
    Influence on the Problem-Solving Approach .....................................30
Summary ...............................................................................................31
Language Education Approaches to Evaluation ........................................32
Some Examples of Mathematics Education Program Evaluations .............37
    Curriculum Evaluations ...................................................................38
    College Level Examples ...................................................................38
    K-12 Comprehensive Examples .......................................................38
    High School Level Examples ............................................................40
    Middle School Level Examples ........................................................42
    Elementary Level Examples .............................................................43
    Professional Development and Systemic Initiative Projects ...............46
    Nontraditional Approaches in Mathematics and Science ..................49
Development of the Problem-Solving Approach ......................................51
3. METHODOLOGY..............................................................................................58

The Integrated Laboratory Mathematics-Science Course ........................................59
  Site ...........................................................................................................59
  Participants .............................................................................................59
  Course Development ...........................................................................59
  Time Span ................................................................................................61
  Data Collection .......................................................................................61
    Classroom Observations ........................................................................62
    Discussions with the Teachers .............................................................62
    Pre- and Posttest .................................................................................64
    Student Questionnaire ..........................................................................64
    Student Interviews ................................................................................64
    Focus Group ..........................................................................................65
    Teacher Interviews ...............................................................................65
    Principal Interviews ............................................................................66

Summary .............................................................................................................66

4. EVALUATION REPORT ....................................................................................68

Student Participants ...........................................................................................68
Chronology .........................................................................................................71
  Pretest ......................................................................................................71
  Early Questionnaire ...............................................................................72
  Student Early Interviews .......................................................................76
  Teachers' Early Interview .....................................................................77
    Goals for the Students .........................................................................78
    Professional Development .....................................................................80
    Early Student Progress .........................................................................80
    Pacing ...........................................................................................81
    Order of Topics ...................................................................................82
    Student Grouping Procedures .............................................................82
    Effects of Evaluation Process ...............................................................83
    Teachers' Relationship with Each Other .............................................83
  Focus Group ............................................................................................84
    General Impressions of the Course .....................................................85
    Repetition ........................................................................................86
    Balance of Mathematics and Science ..................................................86

x
APPENDICES ................................................................................................................162

Appendix A ..................................................................................................................162
  Pretest and Posttest .................................................................................................164
Appendix B ...................................................................................................................169
  Student Early Questionnaire ....................................................................................170
Appendix C ...................................................................................................................171
  Student Late Questionnaire ......................................................................................172
Appendix D ...................................................................................................................173
  Student Early Interview ............................................................................................174
Appendix E ...................................................................................................................175
  Student Late Interview .............................................................................................176
Appendix F ...................................................................................................................177
  Focus Group Questions ............................................................................................178
Appendix G ...................................................................................................................179
  Teacher Early Interview ............................................................................................180
Appendix H ...................................................................................................................181
  Teacher Late Interview .............................................................................................182
Appendix I ...................................................................................................................183
  Principal Early Interview ..........................................................................................184
Appendix J ...................................................................................................................185
  Principal Late Interview ............................................................................................186

LIST OF REFERENCES .................................................................................................187

xiii
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Categorizations of Approaches to Program Evaluation</td>
<td>21</td>
</tr>
<tr>
<td>2.2 The Steps in Lynch’s <em>Context Adaptive Model</em> (1990a)</td>
<td>35</td>
</tr>
<tr>
<td>2.3 Problem-Solving Strategies with Applications to Program Evaluation</td>
<td>55</td>
</tr>
<tr>
<td>3.1 Data Collection Activities for Laboratory Mathematics-Science Pilot Course</td>
<td>67</td>
</tr>
<tr>
<td>4.1 Student Descriptions</td>
<td>69</td>
</tr>
<tr>
<td>4.2 Pretest Scores</td>
<td>71</td>
</tr>
<tr>
<td>4.3 Early Questionnaire: Question 3. How are mathematics and science related?</td>
<td>74</td>
</tr>
<tr>
<td>4.4 Late Questionnaire: Question 2. How are mathematics and science related? Give an example of this relationship.</td>
<td>91</td>
</tr>
<tr>
<td>4.5 Pretest and Posttest Scores</td>
<td>99</td>
</tr>
<tr>
<td>4.6 Pretest to Posttest Gain by Student</td>
<td>99</td>
</tr>
<tr>
<td>5.1 Problem-Solving Strategies Employed in Pilot Course Evaluation</td>
<td>139</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>148</td>
</tr>
</tbody>
</table>

- Figure 6.1: The Problem-Solving Approach to Program Evaluation
CHAPTER 1

RATIONALE

In an age of accountability it becomes increasingly important to document what we do and how well we do it. One way to document successes, failures, progress, and change in mathematics education is through program evaluation. Although program evaluation is often seen as a fairly technical process carried out at the directive of administrators primarily for financial accountability reasons, it holds promise for showing how all aspects of an educational program work together toward any degree of success. This project involved the development and piloting of a new approach to program evaluation that uses the problem-solving process as a metaphor for program evaluation.

What is Program Evaluation?

Worthen, Sanders, and Fitzpatrick (1997) define evaluation as the process of "determining the worth or merit of an evaluation object (whatever is evaluated)" (p. 5), and argue that the purpose of evaluation is "to render judgments about the value of what is being evaluated" (p. 8). Because the purpose of educational research is to determine the effectiveness of some new approach or of new teaching materials, program evaluation should seem an
essential task for educators. Indeed, Worthen et al. argue that, although program evaluation had its first real boost from evaluation mandates for federally funded education programs in the 1960s, that boost in evaluation activity led many agencies to appreciate evaluation for the information it provided. Thus, evaluation was institutionalized, and is now an essential component of nearly all funded projects.

Shadish, Cook, and Leviton (1991) argue that program evaluation is an important endeavor because of the distinct lack of a single, agreed-upon list of criteria for determining the worth of something evaluated. Unlike the business sector, education and other social programs do not use a profit margin as the criterion for evaluation. In education, while it is often the case that a program is instituted in order to raise achievement test scores, some programs are designed, for example, to address only affective components of the learning process. This lack of a single, well-defined criterion for judgment makes the evaluation task that much more important, because even the criteria by which a program is judged must be justified. Further, the wide variation in judgment criteria implies that every evaluation is different, specialized to some degree, and very dependent on context.

Kemmis and Stake (1988) state that the evaluation process is a "search for understanding" (p. 8), and involves communicating that understanding to others. Further, evaluation is an attempt to establish interpretations—judgments of both the object of the evaluation and of the evaluation process itself. A third aspect of evaluation is an orientation to action, often toward decision making for
program change. Finally, evaluation is political; it is always linked through the questions asked and the judgments made to those who are in power to ask the questions and make the judgments.

The purpose of program evaluation varies from case to case, but Tyler (1991) points out that there are at least six essential aims. His list includes (a) monitoring existing programs; (b) choosing a new program to replace one that has been determined ineffective; (c) aiding in the development of new programs; (d) identifying various effects of the program on different subgroups of those participating; (e) providing estimates of effects and/or costs in order to inform consumers or potential consumers of services; and (f) testing the validity of the guiding principles of a program. With such varied purposes, one can see why program evaluation takes so many different forms.

Because program evaluation is defined as a process of determining value or worth, and seeks to gain understanding of the object being evaluated, it seems a natural process to apply to questions that are currently at the forefront of education. In particular, the field of mathematics education is under constant scrutiny, and with national, state, and local standards to which teachers must aspire, approaches are needed to demonstrate the degree to which the reform efforts being advocated and implemented are effective. Program evaluation, because it looks at more than just a single element or issue, provides a comprehensive view of the effectiveness of a newly implemented (or well-established) curriculum or other program, and is a means of providing the
accountability that is expected by administrators and the public. Program evaluation can also provide data useful in the successful dissemination of reform efforts, either at the school or higher education level.

Assessment and Program Evaluation

Recent calls for reform in mathematics education have focused the public's attention on assessment. Broad policy documents emphasize the need for assessment that goes beyond standardized tests that focus on computational skills and instead better reflects the current notion of what should be in a mathematics curriculum. In essence, the argument is that assessment guides the curriculum—what is tested gets taught, and if we want to change what is taught we must change what gets tested (MSEB, 1991; MSEB & NRC, 1993).

Herman (1997) states that assessment serves a communication purpose, providing feedback on an individual student's progress as well as focusing attention on what is important. Logically, assessment must reflect the standards we have set for students. All too often, Herman argues, assessments are created before standards—an illogical order that ends up making assessments a less valid (if not completely invalid) match with the intended curriculum.

The logical order—standards first, then assessments—makes assessment a key factor in reform efforts. The MSEB (1993) summarizes reform literature that argues that assessment can in fact serve as a catalyst for reform, thus making the role of assessment ever more important. Change in assessment must come in terms of content as well as form: not only must the content be important and nontrivial, but the form must match the content in a logical way. If, for example, problem-solving skills are being assessed, the assessment must reflect the
meaningful use of those skills—the right answer, in this case, would not be enough, and some form of exposition would likely be in order. Danielson (1997) argues similarly on the classroom level, saying that not only do assessments guide what teachers choose to teach, they also determine students’ focus: we must assess the skills we want our students to learn for the students to consider them important enough to study.

The National Council of Teachers of Mathematics (NCTM) Standards documents (1989, 1991, 1995, 2000) generally focus on assessment at the student level (as a means of determining individual student progress and achievement) and at the classroom level (as a tool for the teacher to use in making instructional decisions). Although many standardized tests are intended to determine achievement in a summative way, the arguments presented thus far, along with the NCTM focus on classroom and student-level uses of assessment, demonstrate a formative view of assessment—one that establishes assessment as a means of obtaining information that assists in making decisions to improve teaching and thereby further student progress. In this sense, assessment becomes that form of communication, providing feedback on student progress and focusing attention on what is important, to which Herman (1997) referred. Assessment becomes a way of determining what progress has been made, either on the national level in terms of reform efforts or on the student level in terms of skills and conceptual understanding, and determining how to make further progress.
With this emphasis on the assessment of student learning, it would seem logical that there would be interest in assessing ourselves as teachers and the courses we teach, gathering data to use in a formative way—a way that would inform the thinking and decision-making processes at work in the classroom and in the policy-making board rooms. To this end, program evaluation, whether at the course, school, district, state, or national level, would be an appropriate task because it takes a comprehensive view of the situation, including the perspectives of administrators, teachers, parents, students, and even the general public. The NCTM *Assessment Standards* (1995) define one of the purposes of the assessment standards to be for evaluating programs. They argue that there is a “close connection between assessment for the purpose of making educational decisions and assessment for the purpose of evaluating programs” (p. 66).

Three recommendations are made in the *Assessment Standards* (1995) to make program evaluation more consistent with the NCTM ideas of reform in mathematics education. The first recommendation is to make use of data from a variety of high-quality sources rather than that of a single test or even a single testing format. Second, there should be a move away from the reporting of a single group mean toward detailed analysis of data, including the examination of variations in responses and the disaggregation of data by subgroups. Finally, the NCTM suggests more attention be paid to the professional judgments of classroom teachers about student progress and achievement rather than relying solely on outside tests and data sources.
In light of these suggestions, a specific model for evaluation of mathematics programs is needed. Such a model should accommodate the national focus on assessment, consider program evaluation as an essential component of mathematics education reform, and incorporate suggestions from the NCTM and mathematics education specialists. The purpose of the present study was to develop such a model.

The Problem: A Lack of Program Evaluation Models

Specific to Mathematics Education

Mathematics educators—content specialists dedicated to the subject of mathematics—would naturally look for models of program evaluation within the field of mathematics education. Unfortunately, only a few such models exist, and indeed, a search of related literature revealed only three “how-to” guides on program evaluation in mathematics. None of these guides assist in identifying and analyzing the interrelatedness of various program factors and their contributions to a program, which is a hallmark feature of true program evaluation.

Existing Mathematics Program Evaluation Guides

The first such guide is a publication of the NCTM titled *How to Evaluate Your Mathematics Program* (1981). This 19-page booklet consists of a set of 21 standards for school mathematics programs established by a collaborative in Maryland, and grouped into the following categories: instruction, curriculum and instructional materials, the teacher, and physical facilities and equipment. The authors state that “all twenty-one standards are attainable” (p. 3, emphasis
in original) and that attainment of all the standards indicates a successful program. They continue by explaining that the definition of "attained" or "not attained" is up to the individual school, and present a detailed instrument (11 of the 19 pages) to assist in that decision. Since this document was published several years earlier than NCTM's *Curriculum and Evaluation Standards* (1989) and the ensuing reform efforts, a current program would have somewhat different goals and objectives from those indicated in these 1981 standards. Further, although the authors state that all the goals are "attainable," this evaluation instrument seems rather narrowly defined and prescriptive in its expectations, and the booklet does not adequately address the utility of results from such a self-study.

The second mathematics-specific guide is a joint project of NCTM and the Association for Supervision and Curriculum Development, titled *A Guide for Reviewing School Mathematics Programs* (1991). This document is more extensive in scope, including curricular goals by grade levels and aspects of school administration, but remains little more than a survey-type instrument. Some guidance is given as to how the guide should be used—that individual schools should tailor its use to the school context, and suggesting that the instrument be part of a textbook adoption process. In general, however, this 65-page document is of limited use in investigating the interrelatedness of factors contributing to the success (or failure) of a total program.

The final mathematics-specific guide is also from the Association for Supervision and Curriculum Development, the *Mathematics Assessment Process: A Curriculum Alignment Strategy* (Pechman, 1992). The result of an extensive
project with national consultants, this two-volume publication takes the idea of program evaluation much further. Although designed for K-8 programs, the author states that the process is easily adaptable to high school programs, as well. The first volume is a 98-page user's manual that outlines the process of curriculum change and provides guidance on how the instruments are to be used. The second volume contains the actual instruments used for evaluating a mathematics program. The review process outlined in this guide takes a somewhat more comprehensive programmatic view than the first two examples, and includes criteria in nine categories: curriculum, instruction, thinking processes, developmental diversity, attitudes, relevance, collegiality, community, and continuing assessment and redesign. The instruments in volume 2 include surveys as well as interview protocols for several different constituency groups. Overall, this assessment process differs from the other two in that its stated purpose is to assist in making changes in a mathematics program to help the program become more aligned with the NCTM Standards. Thus, this mathematics program evaluation is intended to do more than assess the current status of a program, and is in fact designed specifically to help guide a program toward very specific goals.

In summary, however, none of these examples provides a comprehensive guide for overall evaluation of mathematics programs. They are either limited in scope, in evaluation activities, or in grade level applicability, and essentially provide instruments tied to specific standards rather than a flexible approach or
model. Chapter 2 provides an overview of examples of mathematics program evaluations at various levels and a review of the more general program evaluation literature to provide further background for the current study.

Need for a Mathematics-Specific Approach

The importance of content knowledge is always of concern in a program evaluation, but the way in which content knowledge is integrated into the evaluation process varies. Worthen and Sanders (1984) explain that content can be addressed in one of three ways: (a) by engaging a professional evaluator who consults with content specialists as needed; (b) by working with an evaluator who has substantial content knowledge; or (c) by employing a content specialist who has some expertise in program evaluation. Worthen and Sanders focus on the person in the role of evaluator, and they conclude that the best choice is to engage a professional evaluator who consults content specialists as needed. This conclusion is based on their argument that on the one hand, an evaluator with substantial content knowledge might still not have the expertise to deal with all content issues (and therefore would still need to call on content specialists), and on the other, a content specialist would be so close to the subject matter as to influence the program through interfering or to bias the evaluation toward his own views of content issues. Overall, they argue that the best way to formulate an evaluation team is to have the professional evaluator consult with content specialists.

However, program evaluations are often conducted by an individual (Worthen & Sanders, 1984), not a team of experts in various fields. If indeed a professional evaluator is chosen, several problems arise. One problem is that of
respect and credibility: The program evaluator who is not a mathematician lacks knowledge of mathematics itself, and may not share the implicit values and culture of mathematics education. Further, Worthen and Sanders acknowledge that any conclusions or recommendations that are specific to the content area may be viewed as suspect, since the recommendations do not come from one who knows and understands the content. Wolf (1990) also notes that subject-matter specialists have significant and worthwhile concerns about evaluations—in terms of both focus and processes—and that those concerns merit attention. Stufflebeam (2001) expresses concerns about the subjectivity of a content expert conducting an evaluation, but adds that an evaluation conducted by an expert in a certain field will provide for more insightful analysis.

Together these factors call for an approach to evaluation that mathematics specialists can undertake themselves. The development of an evaluation approach that is feasible for mathematics specialists to implement would result in evaluators who would better understand the implicit agenda of a mathematics program, who would be more familiar with the culture of mathematics, and who would be more aware of what is expected within that mathematical culture. With this context and point of view, a mathematics-specific program evaluation will result in potentially richer data with conclusions that will be better understood by those implementing a mathematics program. Furthermore, a mathematics-specific approach would potentially bring to light more specific content-based recommendations for program improvement. Overall, such
advantages increase the likelihood that an evaluation report will be used, because of both its source (a fellow mathematician as evaluator) and its process (a familiar mathematical approach).

Research Questions

The present project consisted of the systematic development of a mathematically-responsive approach to program evaluation. As part of the development process, the Problem-Solving Approach, which used the mathematical process of problem solving as a metaphor, was piloted with a course that integrated mathematics and science at the high school level. Research questions for the overall project were as follows:

1. How can evaluation of mathematics programs and courses be approached in a way that reflects the values of the mathematics education community?

2. How does the Problem-Solving Approach to program evaluation affect the role of students, teachers, and evaluators in the evaluation process?

3. How does the Problem-Solving Approach affect the conclusions and recommendations that result from a program evaluation?
CHAPTER 2

REVIEW OF RELATED LITERATURE

This chapter will present numerous ways of viewing, categorizing, and conceptualizing approaches to program evaluations. Examples of evaluations, primarily in areas of mathematics education, will be provided. Finally, the Problem-Solving Approach used in this study will be developed.

Generic Approaches to Program Evaluation

Approaches to program evaluation abound, and are categorized differently by various researchers. One of the most commonly known distinctions is that of summative and formative evaluation as first delineated by Scriven (1967). As Scriven's work has been interpreted, summative evaluation occurs after a program is completed and is used to determine the program's effectiveness, while formative evaluation occurs while a program is in progress and is used to change or further develop the program in some way. Other evaluators have argued that this dichotomy is too limited, and have posited that evaluations can be (and often are) both summative and formative (Worthen, Sanders, & Fitzpatrick, 1997). That is, evaluations can and often do provide "results" to an administrative body or funding source while simultaneously
serving as feedback for improvement to those involved with day-to-day operations of the program. Scriven (1991) later argued that his work had been largely misinterpreted and reduced to the dichotomy we now know, and that his actual intent was to represent program evaluation as having different roles. In other words, a single evaluation could be used for summative purposes and for formative purposes, depending on presentation of data. Since this is more in line with his later critics, the argument is less important. The point remains that evaluations are often viewed and construed as having only one or the other purpose.

The Positivist View

Fitz-Gibbon and Morris (1987) present a very positivistic approach to program evaluation. They offer six ways to design an evaluation: (1) true control group, pretest-posttest; (2) true control group, posttest only; (3) nonequivalent control group pretest-posttest; (4) single group time series; (5) time series with nonequivalent control group; and (6) before and after. These designs assume an experimental orientation to program evaluation, and can only be effectively implemented in nearly ideal conditions. The authors acknowledge and others concur (Guba & Lincoln, 1989; Stufflebeam, 2001) that program evaluators rarely have ideal conditions, and state that their suggestions are merely guidelines for evaluation projects.
Beyond the Positivists

Stecher and Davis (1987) go beyond the experimental, presenting four additional approaches to program evaluation in a progressive continuum.

**Goal-Oriented Approach**

The first, the *goal-oriented* approach, emphasizes the goals and objectives of the program and attempts to measure the success of the program in terms of its goals. In this type of evaluation, the evaluator takes on the role of measurement specialist. An advantage of this orientation is the clarity gained by relating goals and outcomes, but they point out that that focus on original goals may leave some unintended but important outcomes unnoticed.

**Decision-Focused Approach**

Second, the *decision-focused* approach is guided by the decision needs of the program—data are gathered in relation to issues on which changes are already being contemplated or at crucial steps in implementation or progress. Effective program management is a goal, and the evaluator must work backwards from clear decision points to determine what information is needed to inform the decision-making process. The strength of this type of evaluation is its focus on the direct needs of those in charge of a program. Its weakness lies in the fact that many decisions are not made at definite points in time, and instead occur gradually, thus making the timing of data collection unclear.

**User-Oriented Approach**

Third, the *user-oriented* approach concentrates on providing information that will be useful to those who are in a position to take action based on the data. The emphasis is on people, and the evaluator involves user groups throughout
the evaluation, which is a strength of this approach. The main weaknesses lie in
the influence of strong personalities and changing interests as user groups
change in composition. A special consideration in this approach is that the
evaluator becomes heavily involved in the operations of the program, and
ultimately acts as an internal collaborator rather than an external evaluator.

Responsive Approach

Fourth, the responsive approach puts the evaluator in the position of
seeking understanding of all evaluation questions from the point of view of all
participants, program staff, administrators, and others. This approach uses more
qualitative data collection methods, and therefore is functional in less focused
evaluations. Drawbacks are potential subjectivity, and the fact that priorities in
the evaluation process are not established because of the attempt to eventually
address all perspectives of every question or problem. The evaluator in this case
acts more as a counselor or facilitator. This conception of the evaluator’s role
increases the amount of communication between the evaluator and those
involved in the program. That focus on communication is in keeping with
several evaluation specialists’ views, including Stake (1975), Cousins and Earl
(1995), and Sanders (2000), who states that “three aspects of good program
evaluation [are] communication, communication, and communication” (p. 1).

This continuum of program evaluation approaches, as presented by
Stecher and Davis (1987), shows the changing points of view between the
approaches in an orderly fashion. The experimental approach uses a distinctly
positivistic orientation and includes mostly quantitative measures. The goal-oriented, decision-focused, user-oriented, and responsive approaches employ increasingly naturalistic orientations and qualitative measures.

**An Alternative Continuum**

Worthen, Sanders, and Fitzpatrick (1997) discuss the previous four approaches to evaluation similarly. They also present the approaches in a continuum from more positivistic to more naturalistic orientations, but refer to the goal-oriented approach as the *objective-oriented approach*, the decision-focused approach as the *management-oriented approach*, the user-oriented approach as the *consumer-oriented approach*, and the responsive approach as the *participant-oriented approach*. However, they add two additional ways to view evaluation, which fall on the continuum between the consumer-oriented and the participant-oriented approaches.

**Expertise-Oriented Approach**

The *expertise-oriented approach* is the term Worthen et al. (1997) use to denote the oldest and perhaps best known form of evaluation, in which a professional in the particular field is called upon to evaluate the specific program. In this approach, an expert or a team of experts in the relevant field is called upon to provide professional, subjective judgments about the program being evaluated. A review for accreditation is one example that would fall under this category of evaluation.
Adversary-Oriented Approach

Adversary-oriented approaches make up the other category offered by Worthen et al. (1997). This approach borrows ideas from the field of law, in that the evaluation plan attempts not only to generate but also balance opposing viewpoints on a program. This process often involves conducting open "hearings" in which all viewpoints are freely expressed. The authors point out that this approach helps eliminate criticism that an evaluation is one-sided, and in fact incorporates a form of metaevaluation (evaluation of the evaluation process itself) since the processes employed by each side are sure to be under close scrutiny from those holding opposing views.

Worthen et al. refer to the approaches as the "theoretical and conceptual underpinnings of most of today's program evaluations" (1997, p. 172). They point out, however, that program evaluation is a relatively new field, and that it is not considered yet to be a "full-grown discipline" (p. 172). They continue by stating that the approaches presented do not meet the criteria to be considered scientific models, and indeed that most so-called models for program evaluation are difficult to emulate. Further, these categories of approaches are not theories since they are not nor can they be written in axiomatic style, and they do not allow us to make predictions or conjectures of any sort. They conclude that these approaches are in fact simply "individuals' conceptions about the field of evaluation,...sets of categories, lists of things to think about, descriptions of different kinds of evaluation" (p. 173)—useful information, but not theories or models, a point on which Stufflebeam (2001) agrees. Indeed, Worthen et al. point out that the field continues to expand, and that it will be much longer
before we can attempt to put together all we know about program evaluation into a model. They argue against too much consolidation, however, because "evaluation contexts are so different that it is difficult to conceive of any one or two models that would be relevant to all" (p. 175).

Stufflebeam’s Presentation

In keeping with the argument against consolidation made by Worthen et al., Stufflebeam (2001) presents an extensive listing of 22 approaches to evaluation. He groups the approaches into four categories. The first, pseudoevaluations, are those he considers to render incomplete or invalid results, and include the politically-controlled approach and the public relations-inspired approach. He dismisses these as essentially invalid pursuits, and recommends that they not be accepted as valid means for evaluation. The second category is questions/methods-oriented approaches. Stufflebeam lists 13 approaches in this category, all of which are essentially variations on the objective-oriented or goal-oriented approach described in the previous categorizations, although experimental studies are also included. The third category is improvement/accountability-oriented approaches, which includes only three separate approaches that are variations on the management-oriented and consumer-oriented approaches described earlier. Finally, Stufflebeam’s fourth category is social agenda/advocacy approaches. He presents four approaches in this category, all of which build on the ideas of the responsive or participant-oriented approach described in the categorizations presented earlier.
Stufflebeam's (2001) presentation of approaches shows that there are many ways of looking at the task of evaluation, and while some simplify and group the approaches more broadly, others emphasize fine distinctions in methodology or participants or orientation. Stufflebeam goes further in his monograph and evaluates the 22 approaches he presents, suggesting that only nine of them deserve further pursuit. Here, for simplicity and because direct relationships to Stufflebeam's proliferation of approaches can be drawn, the discussion will be limited to the simpler categorizations of Stecher and Davis (1987) and Worthen et al. (1997).

As a point of reference, the categories of approaches as presented by Stecher and Davis (1987), Worthen et al. (1997), and Stufflebeam (2001) are compared in Table 2.1.
Table 2.1: Categorizations of Approaches to Program Evaluation

<table>
<thead>
<tr>
<th>Expertise-Oriented</th>
<th>Objective-Oriented</th>
<th>Questions/Methods-Oriented Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal-Oriented</td>
<td>Management-Oriented</td>
<td>Improvement/Accountability-Oriented Group</td>
</tr>
<tr>
<td>User-Oriented</td>
<td>Consumer-Oriented</td>
<td>Social Agenda/Advocacy Group</td>
</tr>
<tr>
<td>Responsive</td>
<td>Participant-Oriented</td>
<td></td>
</tr>
<tr>
<td>Adversary-Oriented</td>
<td></td>
<td>(Pseudoevaluations)</td>
</tr>
</tbody>
</table>

The Accompanying Paradigm Shift

Although still considered a relatively young field, the history of program evaluation reflects a move from a reliance on experiments and quasi-experiments in the 1960s to more case-study orientations and interpretive designs (Walberg & Reynolds, 1998). These varying approaches to evaluation actually reflect a philosophical shift that has been paralleled in most areas of educational research. Greene (1994) describes the evolution of program
evaluation in terms of an underlying move through several philosophical paradigms. The categorizations by Stecher and Davis (1987) and Worthen, Sanders, and Fitzpatrick (1997) illustrate the history of this shift of paradigms. Greene essentially explains that the goal-oriented approach reflects a postpositivist philosophical framework; the decision-oriented approach shows a pragmatic viewpoint; and the user-oriented and responsive approaches demonstrate an interpretivist framework.

The Responsive Approach

Greene (1994) mentions several program evaluation researchers who have made significant contributions toward the paradigmatic shift. Stake's responsive evaluation (1975, as cited in Greene 1994) was one model that helped move program evaluation toward the interpretivist paradigm. The responsive approach is focused on identifying and addressing the concerns of all the stakeholders in a program, and examines those concerns in terms of program improvement. Moreover, Stake (1975, 1976a) offers the responsive approach as an alternative to what he deems preordinate studies, those that are focused on objectives, hypotheses, and questions determined before the inception of the actual evaluation. Rather, the responsive approach allows the evaluator to organize the project around "phenomena encountered—often unexpectedly—as the programme goes along" (Stake, 1976b, p. 20).

Focus on Changing Issues

Stake (1975) contends that responsive evaluations orient themselves around program activities more than program objectives, respond to stakeholder requests for information both during and after the evaluation is
completed, and take into account the differing perspectives of the various stakeholders. He warns that the evaluator should not allow a prescribed list of questions or a set of data instruments chosen early in the evaluation to get in the way of directing attention to those issues that truly concern the stakeholders as the program is progressing. Rather, data instruments should be chosen based on observations that have revealed the major issues of the participants. Indeed, *issues* is Stake's preferred term for the organizing factors in an evaluation, since it reflects "complexity, immediacy, and valuing" (p. 15).

*Fluidity of Evaluation Events*

The ability to respond to concerns is the hallmark of this evaluation design, and is represented in Stake's recommended series of evaluation tasks. He emphasizes that any event can follow any event—for example, data collection processes may be redesigned while data collection or even analysis is taking place—depending on the needs of those involved (including the evaluator) at the time. The tasks include talking with clients, identifying scope, overviewing program activities, discovering purposes and concerns, conceptualizing issues, identifying data needs, selecting data collection vehicles, observing, preparing and presenting portrayals, validating preliminary conclusions, sorting and eliminating information, and assembling reports (adapted from p. 19). The interchangeability of these actions frees the evaluator to work with the stakeholders to stay focused on the issues at hand at any given moment in the evaluation process.
Extension to Participatory Approaches

Stake's responsive approach has been expanded over the years, and as was seen in the categorizations presented earlier, is sometimes considered a part of a broader group of approaches referred to as participatory because of the focus on participants in the program. The case for participatory approaches is made by Cousins and Earl (1995) within the framework of the professionalization of the teaching profession. Specifically, they argue that the move to develop teachers who focus on improving their practice through inquiry encourages their involvement in the evaluation process. Paralleling the professionalism of teaching is a movement to engage all those in the school community in organizational change to improve the overall school program, which also encourages the involvement of educators in evaluation processes. Cousins and Earl emphasize that the idea of participatory evaluation involves both professional evaluators and practitioners, and that some aspect of training of the practitioners in evaluation processes is involved. Overall, their version of participatory evaluation is an extension of the responsive approach because it elevates the role of stakeholders to co-evaluators. Such use of collaboration in evaluations is a trend in the field noted by Walberg and Reynolds (1998).

Influence on the Problem-Solving Approach

Although Stake (1975) does not completely discount what he calls preordinate evaluation designs (those focused on predetermined lists of objectives or questions), his conceptualization of the evaluation process opened the way for evaluators such as Cousins and Earl to follow less traditional and less positivistic approaches. Stake emphasizes that preordinate design
approaches remain valid in some evaluation situations, and that combinations of those approaches with the responsive approach may sometimes be in order. Aspects of the responsive and participatory approaches were incorporated in the approach created for the present study due to the focus on program improvement, the involvement of and attention to all stakeholders, and the provision for fluid evaluation design.

The Connoisseurship Approach

Another milestone in the evolution of program evaluation was, according to Greene (1994), Eisner’s acknowledgement of the evaluator in the evaluation (1991, as cited in Greene, 1994). Eisner’s approach focuses on interpretation and judgment by a connoisseur—an expert in the field of the program being evaluated. It should be noted that Stake (1975) issued gentle warnings against such specialized approaches, referring to the “temptation” (p. 16) of some researchers to apply their particular areas of expertise and styles of inquiry to evaluations. He contends that although it might seem sensible to apply the methodologies of one’s own field to an evaluation, those methodologies might not always be a good match to the issues and questions important in the evaluation.

Essentials of Connoisseurship

Eisner (1976), in his first iteration of connoisseurship, argued that the specialist is the one who knows the best methods to apply to a set of issues within a field. Although originally intended to provide an alternative for those involved in evaluating programs in the arts (he is essentially applying the metaphor of art and literary criticism to evaluation), he extended his idea to
apply to education and educators in general. Eisner (1985) emphasized a need for sensitivity to context, and the ability to identify and interpret slight nuances of meaning. His beliefs developed over his concerns about the fact that evaluation had come to be defined by testing, which he felt placed too much attention on outcomes and not enough on the conditions that bring about those outcomes. Of particular interest for the current project is the fact that another of Eisner’s concerns (first explained in his 1976 article but unchanged even in his 1991 publication) about traditional evaluation was that it has tended to put educational problems into research paradigms rather than find appropriate procedures to fit the problems. In the case of the current project, the aim was to find mathematical procedures to fit mathematical problems.

An Analogy

Eisner (1985) compares the connoisseur to a wine taster. The wine taster develops expertise based on experience across time, with attention to the slightest details and nuances of flavor. His memories of wines tasted in the past forms a backdrop against which to judge others. Indeed, it is the experience and the expertise that enable him to make judgments about the quality of a wine. Eisner argues that these are indeed judgments—not just preferences—because they are “grounded in reasons” (p. 104). Similarly, the educational connoisseur is capable of judgments based on reasoning developed from past experiences in classrooms, in teaching, and in learning.
Application to Educational Evaluation

The application of connoisseurship to educational evaluation requires both connoisseurship, the appreciation aspect, and criticism, which is the disclosure of the judgments (Eisner, 1985). While connoisseurship is private, criticism is public and essentially parallels the reporting aspect of program evaluation. To further explain the connoisseurship metaphor, Eisner points out that it is essentially practiced by teachers every day, when, for example, a teacher realizes the difference between “the noise of children working and just plain noise” (p. 109).

Influence on the Problem-Solving Approach

Eisner’s (1976) application of his own field of expertise to the practice of educational evaluation is, he explains, an acknowledgement of the fact that evaluators will use their own background as a frame of reference, and will essentially create a template for viewing a situation based on the practices of their field. Further, he notes that whatever template is chosen “both constrains and makes possible what one is able to represent” (1991, p. 182). By extension of Eisner’s idea, the current project seeks to create a mathematical template to make evaluation both accessible and feasible for mathematicians. After all, as Eisner states, “People do what they know how to do” (p. 251).

The Fourth Generation Approach

Greene (1994) further explains that in terms of the paradigm shift, we have moved beyond interpretivism to an era when critical theory is being applied to program evaluation, specifically through the work of Guba and Lincoln (1989, as cited in Greene, 1994) in what they call fourth generation evaluation. Greene explains that this approach uses a constructivist framework,
which "requires that evaluation catalyze social action" (p. 540). Guba and Lincoln (1989) call their approach, which is outlined in the next section, *fourth generation* because of their view of the development of the field of educational evaluation.

*The First Generation: Measurement*

Guba and Lincoln (1989) extend beyond the beginning of program evaluation as such, classifying the first period (or generation) as that of *measurement*, when evaluation was focused exclusively on tests and what truths they could measure. This period encompassed approximately the early 1900s to 1930s (Cowin, 1996), and was aided by the rise to prominence of the application of science to social realms and the increasing use of science to inform management decisions and procedures in business and industry.

*The Second Generation: Description*

The second generation is that of *description*, which arose from the realization that the focus of the first generation was on students, when it was becoming clear that it was school curricula that needed revision. This period spanned approximately the 1930s to the 1950s, and moved the focus from strictly human performance to the effectiveness of curricular objectives, strategies for implementation, and organizations (such as schools) themselves. Description became important to clarify patterns of both strengths and weaknesses of school programs, while measurement became "one of several tools" (Guba & Lincoln, 1989, p. 28) in the evaluation process. Cowin (1996) notes that this period marked the earliest beginnings of program evaluation.
The Third Generation: Judgment

The third generation is *judgment*, roughly the 1960s and 1970s (Cowin, 1996). In this post-Sputnik era, it became important not only to measure and describe, but also to make judgments about the quality and effectiveness of a program. These judgments were based not only on what the goals were but also on whether they were appropriate. Thus, evaluators began to problematize even the objectives they had been using (in previous generations) as their evaluation guides. The questions turned from a focus on whether programs were being successfully implemented to whether programs were worth implementing.

Problems of the First Three Generations

Guba and Lincoln (1989) proposed their fourth generation version of evaluation because of several deficiencies they identified with the evaluation process to date. First, they saw prior forms of evaluation as *managerialistic*—that is, the evaluator was elevated to a role that made it possible for the evaluator to decide directions and changes in a program, in much the same way that a manager would do. Second, prior forms did not allow for multiple value systems. Whereas the objectivity of the evaluator had been assumed in prior eras, Guba and Lincoln contend that the supremacy of the evaluator meant that the evaluator’s value system (not objectivity) was imposed on the program. Finally, Guba and Lincoln argued that prior forms of evaluation place too much emphasis on the scientific mode of inquiry. This led to an overreliance on testing and other quantitative measures, and a removal of context that ignored the fact that no program actually existed in an ideal, controlled environment.
The Fourth Generation: Responsive Constructivist Evaluation

Guba and Lincoln's (1989) proposed new generation of evaluation, which they call responsive constructivist evaluation, is an attempt to make up for the shortcomings they identified in previous modes of evaluation. Their approach is responsive in the way that Stake (1975) proposed—it is formed by continuous interaction with all stakeholders negotiating the progress and direction of the evaluation. Further, as an extension of Stake’s approach, Guba and Lincoln contend that stakeholders must confront each other’s issues and problems rather than simply allow the evaluator to mediate them. Fourth generation evaluation is constructivist in that it operates in a way that is significantly different from the traditional scientific mode. This represents complete adherence to a paradigm (which Guba and Lincoln call the constructivist or interpretivist paradigm) in which the existence of a single reality is denied, and it is assumed that there are multiple realities, all socially or even individually constructed. Interaction between subject and evaluator is acknowledged and the idea of true objectivity (that is, of complete valueless detachment) of the observer is denied. Indeed, the evaluation becomes a constructed reality in and of itself, formed from the interaction between the evaluator and those involved in the program—an evaluation conducted by a different evaluator might result in a different constructed reality.

Influence on the Problem-Solving Approach

Although Guba and Lincoln's Fourth Generation Evaluation (1989) is generally considered to be a treatise on the essentials of a new paradigm in overall educational research, it is also, as Greene (1994) pointed out, another of
the major influences in the development of the field of program evaluation. Greene points out that the role of the evaluator in fourth generation evaluation is a particular divergence in this approach, and describes that role as one who facilitates social change. The view of the evaluator as facilitator of change is the influence of fourth generation evaluation on the Problem-Solving Approach: Problem-Solving evaluation is intended to focus on some aspect of mathematics program improvement, and does so with the expectation that the evaluator will enable those engaged in the program to improve it.

Summary

These three approaches—Stake’s responsive approach, Eisner’s connoisseurship approach, and Guba and Lincoln’s fourth generation approach—were major influences in the development of the Problem-Solving Approach. Stake’s contribution encouraged substantial involvement of the stakeholders at various points through the evaluation process, and left it open to changes in procedures and data collection instruments based on reaction to prominent issues that arose during the course of the evaluation. Eisner’s approach encouraged the direct application of the problem-solving metaphor to make evaluation by a connoisseur of mathematics a reality. Finally, Lincoln and Guba’s approach guided the orientation of the evaluator in Problem-Solving evaluation as a mediator of social interaction with the goal of social action for program improvement.
Language Education Approaches to Evaluation

The argument by Worthen et al. (1997) against a single model for program evaluation would seem to point in the direction of content-specific models in education—or, at this point in time, content-specific approaches to evaluation. Actually, Worthen and Sanders argued much earlier (1984) that content specialization and evaluation did not necessarily go hand in hand—at least not in an individual. Their position was based at the time on several factors of academic (university) life, primarily the fact that a content specialist undertaking only evaluation projects would be doomed to career failure because of lack of direct involvement in the progress of the content field. Further, they stated that a content specialist might resist input from individuals or groups in other fields, perhaps even those in evaluation. They concluded that professional evaluators were usually the best choice for evaluating educational programs, but that the evaluators should at least be sensitive to the need for content expertise for certain delicate aspects of an evaluation. The current project is an effort to make the evaluation process more understandable to mathematics specialists, and in cases where appropriate or where there is no qualified evaluator, for mathematicians to undertake evaluation themselves.

The examination of available evaluation materials in mathematics education in Chapter 1 showed that existing literature provides only instruments rather than approaches. (One minor exception is an early suggestion by Alkin [1968], in which he proposes a nonmathematical “model” that he characterizes as mathematical due to its emphasis on educational inputs and outputs in a function-like fashion.) A logical next step, then, is to examine other content areas to
determine if content-specific approaches to program evaluation have been attempted. One area in which progress has been made in this direction is that of applied linguistics and foreign/second language education.

The field of language education has a more extensive literature available on the subject of program evaluation than does mathematics education. One of the earliest examples is a volume on curriculum planning and implementation in which several chapters focus on program evaluation (Johnson, 1989). Hargreaves' (1989) contribution emphasizes formative evaluation as part of curriculum design, calling for a "cyclical, integrated view" of the three aspects of a project: design, implementation, and evaluation (p. 35, emphasis in original). He also argues for internal staff to be involved with program evaluation in order to make an evaluation more thorough, and provides a checklist for evaluation procedures that is not based on any particular theory, but which has developed from his experiences with various language programs. The categories in his checklist are: target audience, purpose, focus, criteria, method, means/instruments, agents, resources, time factors, findings, presentation of results, and follow-up.

Brown (1989), in the same volume, discusses three dimensions of evaluation that he considers particularly relevant to language program evaluation: formative vs. summative evaluation, product vs. process evaluation, and quantitative vs. qualitative evaluation. Although normally seen as oppositions or choices, he views these dimensions as complementary. He also lists 24 data-gathering procedures for evaluations in this context, and discusses a model for curriculum design, implementation, and evaluation. The six overall
steps Brown suggests for evaluating a language program are (1) develop a framework, (2) determine theoretical foci, (3) formulate research questions, (4) select data-gathering procedures, (5) collect data, and (6) analyze and synthesize information. Although set in a language education context, the procedures and data-gathering techniques Brown suggests do not seem specific to the content and issues of language learning and teaching.

Similarly, Elley (1989) offers little that is language education specific, and presents a positivistic view in that he argues for independent (expert) evaluators, and proposes strictly experimental designs. This presentation is somewhat counter to the other literature on evaluation in language education, which tends to argue for a blend of quantitative and qualitative techniques, and recognizes the limitations of true experimental approaches in educational settings.

More promising is an entire volume titled *Evaluating Second Language Education* (Alderson & Beretta, 1992), which includes mostly case studies of various language program evaluations from around the world, but which also contains a final section in which Alderson brings together the basic principles put forth by the examples provided in the text. He guides the evaluator through the planning stages by providing some preliminary answers (or guidelines) to the questions: Why? Who for? Who (in terms of the evaluator[s])? What? How? When? He continues with advice for initial negotiations, implementation, interpretation, and reporting results. He emphasizes the need for evaluators to evaluate their evaluations in order to improve the evaluation process itself and the field in general. Alderson's synthesis, however, remains a fairly generic approach—a blend of many of the approaches listed in the general program
evaluation literature, including (at least) goal-oriented, user-oriented, and responsive evaluations—and adds little that is specific to the content and purpose of language education.

Perhaps the most promising of the approaches available in language education is offered by Lynch (1990a, 1992, 1996). Lynch presents what he calls the *Context Adaptive Model* that was used in an evaluation of an English as a Second Language (ESL) program. Although he admits that it may not be possible to design a single model for all language evaluation contexts, he nonetheless outlines what he believes will prove adaptable to “a variety of, if not all, settings” (1990a, p. 23). His model consists of seven steps, which are summarized in Table 2.2.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish audience and goals</td>
</tr>
<tr>
<td>2</td>
<td>Develop a context inventory and determine what is important in terms of the goals and audience</td>
</tr>
<tr>
<td>3</td>
<td>Develop a preliminary thematic framework based on the important issues in the context</td>
</tr>
<tr>
<td>4</td>
<td>Develop a data collection system, focused by the thematic framework</td>
</tr>
<tr>
<td>5</td>
<td>Collect the data, perhaps revising steps 2, 3, and 4</td>
</tr>
<tr>
<td>6</td>
<td>Analyze the data and revise steps 3 and 4, as necessary</td>
</tr>
<tr>
<td>7</td>
<td>Create a report of results/conclusions</td>
</tr>
</tbody>
</table>

Table 2.2: The Steps in Lynch’s *Context Adaptive Model* (1990a)
Key to this process is step 2, the context inventory, in which all aspects of the program to be evaluated are recorded, ranging from availability of measures of language skills to timing of the evaluation, descriptions of staff and students, lists of resources and instructional materials, and even the social and political climate of the program (1990a).

Lynch (1992), like Greene (1994), also reviews the paradigm shift that has occurred in program evaluation, and uses the Context Adaptive Model to make the point that program evaluation greatly benefits from the use of both quantitative and qualitative methodology. Lynch argues that quantitative data can clear up ambiguities in qualitative data—for example, by linking student performance on tests and other measures with student (and/or teacher) perceptions of a program or curriculum. Indeed, Lynch expands on this aspect in his 1996 book, explaining that each type of data can be used to explain ambiguities in the other, and emphasizing this conclusion as one of the major contributions of the Context Adaptive Model. Knight and Kuleck (1999) offer further support for the use of both qualitative and quantitative methods in program evaluation through their example of an evaluation of a literacy program, where small numbers of participants made qualitative data necessary to support positive quantitative results.

Whereas some (Guba & Lincoln, 1989) argue that the choice of methods represents and is indeed deeply rooted in allegiance to a specific research paradigm, Patton (1988) advocates the freedom to choose methods out of pragmatic reasoning. He refers to his approach as the “paradigm of choices” (p. 119), and explains that it is the very situational nature of a specific program
evaluation that necessitates the freedom to choose any variety of methods. Greene (1998) argues that it is the use of qualitative approaches to evaluation that has helped move evaluation from focusing on program outcomes to including the processes involved with the program. This point is interesting to pair with Lynch’s argument, since processes affect outcomes—that is, program outcomes described by quantitative data can be explained by program processes described by qualitative data.

Lynch’s Context Adaptive Model has received criticism for being too internally focused (that is, not taking into account external data from comparable programs and settings) (Swales, 1990). Lynch (1990b) responds with the argument that external concerns can be built into the evaluation design as part of the third step (developing a thematic framework). A second criticism is that the Context Adaptive Model is repetitive of program evaluation models in general (Meinke, 1990). To this, Lynch counters that although program evaluation is an ongoing process in many academic disciplines, “there is a need to tailor it to the specific concerns of specific disciplines” (1990b, pp. 764-765).

With this call for content-specificity, we return to the problem at hand for the current study: the need for a model, or as Worthen et al. (1997) would say, an approach, that is specific to mathematics education.

Some Examples of Mathematics and Science Education Program Evaluations

One way to work toward an approach for evaluation within the field of mathematics is to examine some evaluation reports of mathematics programs to see what kinds of patterns might emerge. Because the Problem-Solving
Approach was piloted with an integrated mathematics and science course, this section will review a number of recent evaluations of mathematics or mathematics and science programs. The focus of the review will be on methods and data collection procedures in order to show the types of evaluation that have been conducted in mathematics and science programs.

*Curriculum Evaluations*

**College Level Examples**

At the college level, Project FOCUS (Greenes, 1994), sponsored by the Fund for the Improvement of Postsecondary Education was aimed at those needing remediation in mathematics. Seven instructional modules with instructor's guides on basic topics such as integers, fractions, ratio/proportion/percent, graphs, and variables were developed by mathematics education faculty at Boston University. In this case, evaluation consisted of a self-evaluation instrument that was used by both teachers and developers, and the solicitation of comments from students. This is an example of a completely internal evaluation, and surprisingly, did not involve any form of testing.

**K-12 Comprehensive Examples**

An example of an entire public school system's evaluation of its mathematics program is provided in a report from the Des Moines Public Schools (Armstrong, Drake, Cunningham, & Arevalo, 1996). The evaluation is designed around four essential educational components: (a) context, including beliefs, standards, history, a description of the current program, and a list of needs identified in mathematics; (b) input, including budgets, revenue,
instructional materials, and human resource expenditures; (c) process, including goals, staff development activities, management systems, course revisions, etc.; and (d) product, including standardized test results, grades, awards, and identified strengths and weaknesses. This evaluation was more status-oriented—it was apparently not intended to render judgment on the overall effectiveness of the mathematics program for the school system, but rather to provide information about the current status of the program in an organized fashion.

In Hartford, Connecticut, a National Science Foundation (NSF)-sponsored project, the Comprehensive Partnerships for Mathematics and Science Achievement (CPMSA) (Borrelli & Cimochowski, 1999), aimed to increase opportunity for minority students to pursue mathematics and science courses considered gatekeeper courses for college entrance and continuation in higher level mathematics and science. The project involved 14 schools and encompassed 52% of the district’s 22,531 students, and incorporated professional development as well as classroom interventions in gatekeeper courses. Data collection for the evaluation of this program consisted mainly of interviews with the project director, department chairpersons, principals and teachers at the project schools, the superintendent, representatives of involved businesses and universities, parents, and providers of professional development activities. The evaluators also collected and reviewed various related reports, curriculum materials, and documents produced by various committees and working groups in the project. The evaluation revealed an improved implementation of the project in its second (and final) year, due mainly to the hiring of a new project
director, but showed that the project was unable to overcome difficulties in the school system that were beyond its control (the rapid changeover in superintendents, for example). This evaluation represents the only example located through this review of literature of a large-scale program that was evaluated primarily through qualitative methods (namely, interviews).

High School Level Examples

An example at the high school level is the evaluation of the Interactive Mathematics Program (IMP) (Alper, Fendel, Fraser, & Resek, 1997). The IMP, and NSF-sponsored project, was a four-year program for all secondary students and constituted a single curriculum. The IMP aimed at providing both skills and mathematical power for a full range of students. Concepts were introduced with concrete approaches in real-life situations; small group work was incorporated in heterogeneous classrooms; and families were involved with mathematics through problem-of-the-week assignments. Students were encouraged to solve problems in their own way as the project followed a constructivist philosophy.

Evaluation of the program consisted of several components. First, standardized test scores (such as the Scholastic Aptitude Test) were compared, revealing either no difference between the IMP students and those in traditional programs, or, on some comparisons, slightly higher scores for IMP students. Second, a university researcher conducted a five-year evaluation beginning in 1992. One component of this evaluation was a high school transcript analysis, revealing that a higher percentage of IMP students than traditional students took three or more years of college preparatory mathematics, and that IMP students had consistently higher GPAs both in mathematics and overall at both high school and college levels.
Another component of the evaluation, conducted in 1996, involved three studies of achievement using statistics items on the Second International Mathematics Study for ninth graders, two constructed response items from the Wisconsin Student Assessment System for tenth graders, and a quantitative reasoning test used at a well-known university for eleventh graders. In each case, IMP students scored better than students in traditional courses at the respective levels. This evaluation example shows a fairly traditional approach, using quantitative data derived primarily from standardized tests.

The Core-Plus Mathematics Project (CPMP) (Shoen, Hirsch, & Ziebarth, 1998), a comprehensive high school mathematics curriculum project funded by NSF, was evaluated as it was still in development (certain courses were still in revision at the time of this report). Achievement results were analyzed for 33 schools in 11 states where all the courses were field tested. Achievement was measured using standardized tests developed by Iowa Testing Programs and, for course 3, items from the National Assessment of Educational Progress; and using open-ended tests developed by the CPMP team for courses 1 and 2. In general, CPMP students' means were higher than those of students in traditional programs. Here again is an evaluation that relies heavily on standardized test data, but which does include an open-ended test component.

Worthy of mention at this point is Calhoun, Bohlin, Bohlin, and Tracz's (1997) work to develop an instrument to assess the degree of reform in secondary mathematics classrooms. Fifteen teachers at two high schools in California were assessed using a questionnaire about teacher beliefs and an observation inventory. Likert scales were used in each instrument. Analysis of
the beliefs questionnaire with the observation inventory showed correlations that supported the idea that reform could be measured through such a questionnaire. This evaluation is an example of a highly quantitative approach applied to affective and ambiguous components, such as beliefs and degrees of reform.

Middle School Level Examples

An example of an evaluation at the middle school level is that of the Lake Tahoe Watershed Project (Rohrer & Welsch, 1998), which was a summer program for female middle school students involving both mathematics and science. Funded by the Department of Energy, the project lasted two years, and created an environment that the researchers thought would be nonthreatening by putting females in all female groups and including all female scientists and science teachers. Anticipated outcomes included both enthusiasm and evidence of success in mathematics and science, positive interaction with peers, participation in reports for the Department of Energy and other agencies, and continued interest in mathematics and science as evidenced by continued enrollment in courses and by career choices. The curriculum for the summer program included activities such as testing water for pH levels, building a solar cooker, testing soil and rocks from the Watershed using real equipment in real laboratories, and conducting a problem-based research project. The format for instruction was self-paced student choice with small group discussion of the nongraded hands-on activities. With this project, the evaluation consisted of survey forms and interviews which showed that all planned outcomes were met and revealed some unexpected outcomes: The girls developed insights about
themselves, the environment, and the fields of science and mathematics. Some practical aspects of the program were cited as positives, including a one-week break in the middle of the program; the all-female staff; and the fact that there was one van for transport, which encouraged bonding among the girls when on field trips. Key to this evaluation process was the fact that the data from the first summer were analyzed before the second summer got underway. As a result, project directors were able to implement some changes for the second summer, including the fact that more students were selected from lower grade levels (eighth and ninth grade levels) in order to have an earlier impact on career paths, and an earlier discussion of potential research projects, which gave the girls more time to work on them. This example shows an evaluation with an orientation that was open to program improvement—even if unintended or unexpected—because data were analyzed and judgments made while the program was in progress. This evaluation, where the ability to solve problems was an important aspect of program development and improvement, provides an important insight for the Problem-Solving Approach, which uses the problem-solving process as a metaphor for evaluation.

**Elementary Level Examples**

At the elementary level, Wood and Sellers (1996) present an evaluation of a problem-centered mathematics program for third graders. Six classes received problem-centered mathematics instruction for two years (in second and third grades). Of the 19 total classes (417 students) involved, others received problem-centered instruction only in third grade and some were nonproject comparison groups. Evaluation consisted first of comparisons of standardized achievement.
test scores among the three groups. Results showed that the students in problem-centered instruction for two years had significantly higher scores on the achievement tests. A second component of evaluation was an instructor rating of conceptual development in arithmetic, for which those in two-year problem centered classes also scored higher. The final component was the instructors' examination of students' personal goals and beliefs about their reasons for success in mathematics; students in classes receiving problem-centered instruction for two years had stronger beliefs in this area. This evaluation used standardized test scores backed up by instructor-provided assessments of student understanding, thus adding a qualitative component and focusing attention on the perspective of the classroom teacher.

MathWings, a constructivist approach to teaching mathematics designed for use in upper elementary classrooms (grades 3-5), was evaluated in three contexts (Madden, Slavin, & Simons, 1999): four rural Maryland schools, an impoverished urban school in Texas, and an impoverished urban school in Florida. Standardized tests specific to each state were used to compare the MathWings groups to state averages, and results showed significant gains for all three grade levels (3-5) in Maryland, all three grade levels in Texas, and fourth and fifth grades (but not third) in Florida. In the conclusions of the report, it is noted that control groups were not used in this comparison, and should be used in future evaluations. This evaluation serves as an example of how existing standardized test scores can be used to determine success of a program.
The Sisters in Science Program, an NSF-sponsored project, was an after-school program for fourth grade girls in Philadelphia (Hammrich, 1997). It was intended to increase interest, achievement, attitude, and awareness of girls in science and mathematics. Evaluation consisted of a pretest/posttest format using a questionnaire to measure attitudes, interest, and awareness levels and separate instruments to measure science process skills and mathematics skills. Analysis of the strictly quantitative data revealed no change in science process skills, no statistically significant changes in mathematics skills (but notable gains in fractional equivalence), and positive changes in attitudes. This evaluation demonstrates use of the popular pretest/posttest method.

Another supplemental program at the elementary level was Project SEED, a nationwide program that involved professional mathematicians and scientists teaching abstract and conceptually-oriented mathematics (taken from high school and college algebra courses) using a Socratic questioning format to elementary students as a supplement to their daily schedules. Evaluation of the project as implemented in the Detroit (Michigan) Public Schools (Webster, Dryden, Leddick, & Green, 1999) involved use of a test of abstract algebra developed by the evaluator, a standardized test used to measure achievement, a student attitude scale, and surveys of teachers, principals, and parents. Students showed higher abilities in algebra and higher levels of mathematics achievement; student attitudes were highly positive; and teacher, principal, and student opinions of the program were positive. Exact numbers of student participants varied on each type of data collected, ranging from 300 to 500 students; while 25 teachers, 9 principals, and 267 parents responded to their respective surveys.
This evaluation represents a very comprehensive look at all stakeholders—all those who have interest in the program, whether they are directly involved or not—in the program, and is an example where the total numbers of participants varied widely from one measure to another. Notably, the same evaluation process was used for the overall evaluation of the nationwide program (Webster, 1998) as implemented in five urban school districts in five states, with similar results, reflecting an evaluation design that was effective and feasible on both small (300-500 students) and large (1400-1800 students) scales.

**Professional Development and Systemic Initiative Projects**

Because there was a professional development component for the teachers of the integrated mathematics and science laboratory course on which the Problem-Solving Approach was piloted, evaluations of professional development projects in mathematics and science education were also examined. One such professional development project focused on evaluation itself as a process that can promote reform in mathematics and science education. The Regional Math/Science Collaborative (RMSC) of Southwestern Pennsylvania (Tananis, 1998) was a representative group of stakeholders including educators, students, parents, university faculty, and business and community members in the region. A literature review of evaluation procedures and approaches helped them narrow their preferences to three forms of participatory evaluation, from which they designed a pattern of dialogue and action that they termed *discursive practice*. A metaevaluation of this process determined that it was successful in supporting development of the organization and in promoting utilization of
evaluation. Of importance to the current project is the fact that the process they developed involved a participatory form of evaluation, and included some form of action based on stakeholder interactions.

Several evaluations of NSF-sponsored Statewide Systemic Initiatives (SSI) were identified for the purpose of this literature review. The Connecticut SSI, Project CONNSTRUCT, contracted with a professional evaluation service to conduct a primarily qualitative evaluation of its operations in the seventh year of funding (Bruckerhoff & Bruckerhoff, 1998). Data collection included interviews with project administrators and school system representatives across the state, analysis of student achievement data, surveys, observations of events such as board and committee meetings, and document collection. Data analysis revealed five themes (continued success with the SSI’s fiscal agent, the Connecticut Academy for Education in Mathematics, Science, & Technology; enhanced focus on urban centers in the state; the development of high quality products such as assessment instruments; accountability for personnel and operations; and relatively few changes in leadership and staff positions) and a number of issues deserving attention (including the development of a data management system, the creation of partnerships with institutions of higher education in order to influence undergraduate mathematics and science teaching for the benefit of preservice teachers, the difficulty of social context issues in making gains in achievement, and attention to many other urban areas and towns in need beyond four main urban areas focused on to date). This evaluation is an
example of a large-scale project that utilized primarily qualitative methods to identify issues and clarify potential changes, improvements, and future directions for the project.

The evaluation of the New Jersey Statewide Systemic Initiative, an NSF project designed to increase student learning in science and mathematics education through such means as hands-on approaches, real-life examples, and the use of technology in the classroom, focused on three questions (Fenster, 1998): whether students learn more because of the SSI, whether students were better equipped to apply what they were learning to everyday problems, and whether inequities in performance among certain student groups were being reduced. Data were almost strictly from state-mandated standardized tests. Results showed that there was no additional learning for students exposed to SSI-influenced teaching, and that inequities in performance had worsened since the project's inception. The second question, concerning application to everyday problems, did not have much evidence, and could not be answered. Overall, this evaluation shows use of the questions-based or objectives-based approach to evaluation, but is an example of a large-scale program review that apparently attempted to answer questions of different natures with a single type of data.

Past experiences with such evaluations led NSF to conduct a metaevaluation process (Bruckerhoff, 1997) that attempted to identify the most important aspects of SSIs to evaluate and ways to evaluate those aspects. This metaevaluation was done by means of telephone interviews with 19 participants, including evaluators, SSI directors, and NSF officials. Participants were allowed to prepare responses to questions before the telephone call. Analysis of the
interview data showed that participants saw SSIs as significantly different from traditional programs in part because they are more fluid (activities and sequences of events are more flexible and interchangeable), they have changeable goals, and they have multiple sources of influence. These distinctive features of SSIs necessitate significantly different approaches to evaluation, which participants claimed, may eventually lead to a separate field of evaluation devoted to large-scale projects such as SSIs. This metaevaluation shows yet another need for new approaches to program evaluation, and since the SSIs are mathematics, science, and technology related, an approach specific to one or more of the content areas may aid the process.

Nontraditional Approaches in Mathematics and Science

In 1995, NSF published a monograph of papers centered around the theme of identifying new approaches to evaluation that could be used for NSF-sponsored programs. Frechtling (1995) explains that the project was necessitated by several difficulties with traditional approaches to evaluation. First, traditional evaluations have tended to assign responsibility for results to only one source. Second, such evaluations have used quantitative data almost exclusively. Third, traditional approaches to evaluation have placed primary importance on student achievement. These factors have proven problematic because many NSF projects, which are expected to be cutting-edge work, are far more complicated than those in the past, making it impossible to credit any one factor for success, requiring more than numerical data, and affecting components other than just student achievement.
The model proposed in this monograph is based on the metaphor of the footprint (Sharp & Frechtling, 1995). In keeping with the aim of NSF projects, the footprint metaphor was chosen to represent the idea that such projects should leave lasting impact in ways that follow one after another—that is, one project leads to another, which influences another, and so on, in the same way that footprints lead toward a goal. One of the main concerns of the contributors to this monograph is ways to evaluate this dissemination aspect of funded projects. Although this is an aspect that is unique to NSF projects that required a special view of the evaluation process, it still does not provide a mathematics-specific way of approaching program evaluation.

Romberg (1992) states that program evaluations use an increasingly nontraditional, convergent strategy—a combination of several different evaluation approaches in the evaluation of a single program. He cites an evaluation with which he was involved (the Individually Guided Education Evaluation Study), where mathematics and reading were studied in several phases, beginning with standardized tests and quantitative surveys, continuing with ethnographic case studies of a small number of exemplary schools, and concluding with observations and criterion-referenced tests. This combination of quantitative and qualitative methods is representative of such convergent strategies, and is similar in structure to the Problem-Solving Approach used in the current study.
Development of the Problem-Solving Approach

This section will explore the way that mathematics was applied to the field of evaluation in an attempt to create a mathematically sound model for evaluation of mathematics and/or mathematics education programs.

The Metaphor in Mathematics

Both Lynch (1996) and Worthen et al. (1997) mention the contributions of Smith (1981, 1988), who writes of various metaphors that have or can be applied to the field of evaluation. For instance, the adversary-oriented approach described earlier was developed using the metaphor of a trial or hearing, with advocates for multiple perspectives presenting their cases to the evaluator (the judge) and/or representatives of other participant groups (the jury). Other metaphors abound, including ideas from fields such as investigative reporting, architecture, geography, philosophy, literary and film criticism, and watercolor painting (Smith, 1981). Indeed, Payne (1994) points out that all approaches to evaluation are essentially invocations of metaphors, and that "the value of [a] metaphor is to help us think through the entirety of the evaluation task" (p. 90). The goal of the current project was to find a mathematical way to think through the evaluation task.

Problem Solving

In the search for an evaluation approach suitable to the field of mathematics, it would seem appropriate to identify a mathematical metaphor for evaluation. The metaphor used in this project was that of problem solving—a
nonalgorithmic approach to mathematical tasks, and one that seems to have relevance to evaluation functions. Certainly, this metaphor is both accessible and credible to mathematicians.

Problem-solving literature in mathematics education generally points to one common conception of problem solving: the procedure offered by Polya (1957) in the oft-cited volume How to Solve It. Specifically, Polya identifies four stages of problem solving: (a) understanding the problem, (b) devising a plan, (c) carrying out the plan, and (d) looking back to review the procedure and outcome (pp. xvi-xvii). These stages have direct application to the program evaluation process, and thereby imply that the evaluation of a program is a problem to be solved.

**Understanding the problem.** First, the problem of program evaluation must be understood. The tradition of program evaluation dictates that the evaluation report gives a pronouncement as to whether a program is successful or not, and makes some recommendations for improvement based on that fact. The focus of the evaluation process is on determining how well the program is achieving its goals—the recommendations come as a result of that assessment. It seems, then, that the problem being addressed by typical program evaluations can be phrased as a question: "Is this program good?" The Problem-Solving Approach proposes that the question be changed, and that the evaluation instead be focused on a question that presents a real problem that needs a solution: "How can this program be improved?" This question gets more directly at the point of program evaluation in the first place, but it goes beyond just that by streamlining the evaluation process. That is, rather than first determining the
worth of a program and then developing recommendations for improvement, the application of the problem-solving process to this new question puts the evaluator and participants in a new position, with a different stance. Indeed, a focus on improvement automatically shifts the evaluator and participants into a problem-solving mode.

Devising a plan. Many mathematicians and educators have developed problem-solving strategies that lead the problem solver through Polya's stages. Posamentier and Krulik (1998), for example, present the following heuristics for devising a plan: working backwards, finding a pattern, adopting a different point of view, solving a simpler (analogous) problem, considering extreme cases, intelligent guessing and testing, organizing data, and using logical reasoning. Such strategies are frequently mentioned in the mathematics education reform documents (including the NCTM Standards [1989] and Principles and Standards [2000]) that call for a problem-solving emphasis in the teaching and learning of mathematics. The problem-solving vocabulary is a part of the mathematics educator's everyday language because problem-solving skills are what we strive to address in our classrooms every day, as well. In terms of devising a plan, then, it seems logical to apply these very strategies (with which we are so familiar) to the problem of program improvement.

There are several ways in which these heuristics can be applied in almost any evaluation context. Table 2.3 presents the problem-solving heuristics listed by Posamentier and Krulik (1998) with their possible applications to program evaluation. Some problem-solving strategies such as organizing data and working backwards are fairly obvious aspects of any evaluation, since any
evaluator would create an organization system for managing data, and the analysis of data usually requires some degree of backtracking to reason through results. However, the Problem-Solving Approach assumes that the direct application and open acknowledgement of these processes in ways similar to those mentioned in Table 2.3 create a more active view of program evaluation, a fact that has the potential to streamline the evaluation process.
<table>
<thead>
<tr>
<th>Problem-Solving Strategies</th>
<th>Application to Program Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working backwards</td>
<td>Examine goals: what does it take to reach them?</td>
</tr>
<tr>
<td>Finding a pattern</td>
<td>Examine data sources from different participant groups (students and teachers), or different data sources within a group (tests and surveys from students)</td>
</tr>
<tr>
<td>Adopting a different point of view</td>
<td>Take the viewpoint of a student: What do I want from this program?</td>
</tr>
<tr>
<td>Solving a simpler (analogous) problem</td>
<td>Consider curricular issue on small scale—try an approach for one day and evaluate, then apply to entire program if successful</td>
</tr>
<tr>
<td>Considering extreme cases</td>
<td>What happens to high-achievers as they progress through this program? What happens to low-achievers as they progress? How/Why is experience different?</td>
</tr>
<tr>
<td>Making a drawing (visual representation)</td>
<td>Create a flow chart of objectives and goals to show relationships and conceptual and/or chronological order—can be used to see areas of overload or gaps</td>
</tr>
<tr>
<td>Intelligent guessing and testing</td>
<td>Use to help determine cause/effect relationships</td>
</tr>
<tr>
<td>(including approximating)</td>
<td></td>
</tr>
<tr>
<td>Accounting for all possibilities</td>
<td>Consider all possible reasons for dissatisfaction, then look to see if possible causes for each are present</td>
</tr>
<tr>
<td>(exhaustive listing)</td>
<td></td>
</tr>
<tr>
<td>Organizing data</td>
<td>Helps to make comparisons both between and among participants and/or instruments</td>
</tr>
<tr>
<td>Logical reasoning</td>
<td>Work backwards or forwards to determine cause/effect and reasons for other correlations</td>
</tr>
</tbody>
</table>

Table 2.3: Problem-Solving Strategies with Applications to Program Evaluation
Carrying out the plan. Another key aspect of the application of problem solving to contribute a more active view of the evaluation process is the fact that problem solving includes the act of solving a problem—that is, a plan is designed and it is carried out. In the problem-solving model proposed here, the carrying out of the plan means utilizing the heuristics in ways that will reveal useful information about the program and lead to ideas for program improvement. It is at this point that the role of the evaluator becomes key. The importance of a mathematics specialist as evaluator has already been emphasized, but the question remains as to how involved the evaluator should become in the program itself. Should the use of the heuristics—that is, the application of the working backwards procedure, or the adoption of a different point of view—be entirely up to the evaluator, entirely up to the program administrators, or some combination of these two? On the one hand, if the evaluator alone takes on this part of the process, the evaluation would end up looking more like one using the expertise-oriented approach. On the other hand, if left solely to the program administrators, the evaluation would lose the valued insight of the mathematics specialist in the evaluator role—a point for which this paper has argued. Indeed, this is the aspect of evaluation where the interplay between Stake’s responsive approach, Eisner’s connoisseurship model, and Guba and Lincoln’s fourth generation approach became important.

Looking back to review procedure and outcome. This stage of Polya’s problem-solving process involves evaluation of the outcome and of the process used to reach that outcome. As applied to program evaluation, this involves looking to see if the use of problem-solving heuristics produced actual program
improvements. The evaluator and program administrators would review the designs to improve the program, asking questions about their efficiency and productivity, and scrutinizing the changes implemented to determine if improvements did in fact result. Further improvements might also be gleaned from this process, both by determining better ways to collect and/or analyze data for future use, and by identifying overarching program improvements that may result from this overall view of the evaluation.

Summary

This chapter has presented the argument that there is no approach specific to the evaluation of mathematics programs, and that evaluation would benefit from a reorientation from asking “Is a program good?” to asking “How do we improve this program?” The metaphor of problem solving and the use of specific heuristics common to the act of problem solving has been proposed as a way to approach the evaluation of a mathematics program. This approach was piloted with an integrated laboratory mathematics and science course at the high school level. The course and data collection procedures for the evaluation of the course are described in Chapter 3.
CHAPTER 3

METHODOLOGY

This chapter provides an overview of the project on which the Problem-Solving Approach to program evaluation, which uses problem solving as its basis, was applied. Descriptions of the site, the participants, the development and content of the integrated mathematics-science course, and the time span of the project are included. Procedures and instruments (along with their development) are explained. All data collection activities were chosen and developed from the perspective of a problem-solving approach. They were designed to illuminate problems throughout the time span of the course, and were collectively intended to refine the problem-solving approach to evaluation proposed in the current project. A report of the evaluation is provided in Chapter 4. Evidence of problems that arose during the course and an analysis of the use of problem-solving strategies by all participants will be discussed in Chapter 5.
The Integrated Laboratory Mathematics-Science Course

Site

The Integrated Laboratory Mathematics-Science Course was conducted at Valley Circle High School (all names for locations and people are pseudonyms), a high school in a district described by the two teachers of the course as "rural with an urban profile" because of the number of students who take advantage of free and reduced lunch programs. Valley Circle is the only high school in the small town, and has an enrollment of approximately 350 students in grades 9-12. The facility is approximately 25 years old, and sits in the middle of corn fields on the outskirts of town. Because it is a small town school district, most students have known each other through all their K-12 years.

Participants

Enrollment for the course consisted of 12 students, primarily those who were taking the course as an additional mathematics credit beyond Algebra 2. Junior or senior status was a requirement for enrollment. The teachers, one mathematics (Mrs. M) and one science (Mrs. S), shared responsibility for the two-hour block of class time.

Course Development

A long history of student observation, primarily by the science teacher, prompted the teachers' work together on this project. Specifically, the teachers saw that related topics, when addressed in both science and mathematics classes, did not seem to connect for the students—relationships were not being formed between the mathematics content and the science content. They also noted that it was often the case that they could identify students who would not pass
Algebra 2 based on their performance in Algebra 1 and Geometry, and knew that with more time on the concepts of those introductory courses, those students would have a greater possibility for success in Algebra 2. Thus, the course was developed out of a desire to help forge the mathematics-science connections that would strengthen the concepts in both areas and increase success rates in Algebra 2.

Mrs. M and Mrs. S were awarded a Jennings Foundation grant to assist in the development of the course. Funds were sufficient to purchase graphing calculators and appropriate data collection devices necessary for activities to connect mathematics and science. The grant also provided funds for the two teachers to attend a Texas Instruments Teachers Teaching with Technology (T³) Institute in Summer 1999. They participated in the week-long session entitled “Connecting Mathematics and Science.” The teachers were also granted time for curriculum planning by their school system, during which they developed a course of study that uses mathematics concepts as the core curricular concept and appropriate science activities to support the mathematics. Specifically, the course focused on nine performance objectives:

Given the appropriate laboratory investigation students will collect data that

1. has a linear relationship.
2. will produce an inverse or inverse square function.
3. will produce a quadratic function.
4. will produce an exponential or logarithmic function.
5. will produce a logistics function.
6. will produce a piecewise continuous function.
7. will produce a second-degeree curve.
8. will produce a statistical plot of the data.
9. will produce a periodic function.
It should be noted that although a primary purpose of the course was to foster connections in students' minds between mathematics and science, the teachers agreed that despite many approaches to integration (Huntley, 1998; Lonning & DeFranco, 1997; Meier, Cobbs, & Nicol, 1998; Roebuck & Warden, 1998) mathematics should be the primary focus of this course. As a result, they negotiated a mathematics credit for students who completed the course.

**Time Span**

The Integrated Laboratory Course was offered in the third trimester of the school year, which began March 6, 2000, and concluded June 1, 2000. The trimester schedule design allowed for a two-hour block class by joining two of the one-hour periods of the day, with the teachers assigned to one hour each, and the other hour as their respective designated “planning” times. The teachers sometimes spent the entire block together in the classroom, while other times one or the other of them was responsible for the two-hour time period. On rare occasions they divided the time equally.

**Data Collection**

The application of problem solving to the program evaluation process had several implications for data collection. First, it was unclear as to what problems might arise, so the data collection activities changed slightly over time. Some data collection activities were developed as the pilot course was carried out, and were responsive to the needs of the teachers, the students, or myself as evaluator as we identified specific problems that needed to be addressed. Second, the problem-solving approach demanded that I, as evaluator, maintain regular contact with all the participants in a way that would enable me to identify
problems and be ready to design data collection techniques that would help address them. Third, the problem-solving approach is not one of the conventional modes of program evaluation. Recognizing that fact meant that I had to attempt to incorporate more conventional data collection activities already outlined in the grant proposal or expected by certain stakeholders. Through the data collection activities undertaken, I attempted to account for all these considerations.

Classroom Observations

To address the first two implications—that problems would arise as the project unfolded and that regular contact was necessary—I observed the class period (the full two-hour block) approximately once every other week. Observations included interaction with the students as they participated in data collection activities, analyzed data using graphing calculators, and discussed and/or wrote up their conclusions. Additional observations were occasionally included at the teachers' request, at times that allowed observation of the students as they engaged in the full range of activities implied by the curriculum (including assessment), and involved in work related to as many as possible of the nine performance objectives outlined in the course of study. Field notes from the observations were analyzed informally but continuously in order to identify concerns as they arose.

Discussions with the Teachers

Along with the biweekly visits, there was time during class or after school to discuss progress with the teachers (sometimes individually; sometimes together). Topics of concern regularly addressed during these informal
discussions included the teachers' own feelings toward progress they had made in designing/implementing the course, their assessments of student progress (both formal and informal), and their concerns for upcoming topics and/or activities. The researcher, in keeping with the Problem-Solving Approach of using the problem-solving process for evaluation, suggested that the teachers use problem-solving strategies to address problems that surfaced along the way.

To address the third implication of the problem-solving approach (that some tasks were predefined by the grant requirements and prior expectations of program evaluation), the areas that were specifically targeted for evaluation in the grant proposal included the following:

1. the students' familiarity with the use and features of the latest graphing calculator technology
2. the students' familiarity with the use and features of the Calculator-Based Laboratory (CBL) System and the Calculator-Based Ranger (CBR)
3. the students' ability to explore the scientific aspects of the data collected using the CBL/CBR system as related to science
4. the students' ability to learn the mathematical concepts inherent in the data collected with the CBL/CBR system
5. the students' ability to utilize technology in a balanced program of reasoning, connections, and communications

These points provided a basis for the data collection instruments that were developed for the evaluation, the first of which was designed by the teachers themselves.
Pre- and Posttest

The teachers had chosen a design for data collection centering on points 3 and 4 (from the grant proposal): A pre- and posttest (see Appendix A) was administered to assess the students' ability to predict the type of motion required to produce a certain graph (for which a sketch is provided), and to predict the type of graph (by sketching) produced by a type of motion. The other points (1, 2, and 5) were addressed at least in part through the regular observations and informal conversations with the teachers described earlier.

Student Questionnaire

Along with this specific list of evaluation issues provided by the text of the grant proposal, the teachers voiced other concerns that they wished to be addressed by this evaluation. They proposed that a questionnaire be administered at the beginning and end of the course to address whether the students saw learning mathematics and science as enjoyable tasks, the degree to which they saw connections between mathematics and science, and their opinion of the importance of seeing connections between mathematics and science. Points 1 (familiarity with graphing calculators) and 2 (familiarity with data collection equipment) of the evaluation in the grant proposal were also addressed in the questionnaire. The Student Questionnaire (see Appendixes B and C), both early and final versions, consisted of open-ended questions.

Student Interviews

In addition, and to parallel the pre- and post-activities indicated thus far, audiotaped interviews were conducted with each student near the beginning and at the end of the course. The questions for the Student Interviews (see Appendix
D) somewhat paralleled those of the Student Questionnaire because the researcher thought that interviews would produce richer data, especially with high school students. Questions for the student interviews at the end of the course (see Appendix E) were modified based on progress of the course and because some data were already obtained through the focus group.

**Focus Group**

A focus group was conducted just past midway through the course during class time to assess their views of progress and to determine how their views of mathematics and science (and the relationships between the two) might be changing. The questions for the Student Focus Group (see Appendix F) included, among other issues, opportunities for the students to indicate to the teachers their opinions of the usefulness of the activities in which they had been engaged. Both teachers, the evaluator, and an assistant to the evaluator were present and participated throughout the one-hour focus group.

**Teacher Interviews**

In order to gain more specific insights into the teachers' impressions near the beginning and end of the course, audiotaped interviews were conducted with the teachers. The first interview was conducted with the teachers together; the final interview consisted of one portion with both teachers and another portion with each teacher individually. The Teacher Interviews (see Appendixes G and H) focused on goals for the course and perceived expectations of the students, as well as assessments of their own progress and that of the students.
**Principal Interviews**

Similarly, the perspective of the principal was included near the beginning and again at the end of the course through audiotaped interviews. The Principal Interviews (see Appendixes I and J) included questions concerning implementation of the course from an administrator's point of view, such as financial considerations and scheduling issues, and future plans for such a course in the school's curriculum.

**Summary**

Data collection activities are summarized in Table 3.1. The table includes the planned dates for all activities—actual dates varied slightly in some cases. Significant variations in these activities will be discussed in Chapter 4.
<table>
<thead>
<tr>
<th>Data Activity</th>
<th>Week #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Questionnaire</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Interviews</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction Assessment</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal Interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Focus Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Table 3.1: Data Collection Activities for Laboratory Mathematics-Science Pilot Course
CHAPTER 4

EVALUATION REPORT

This chapter presents the data from the evaluation of the Integrated Laboratory course that was used to pilot the Problem-Solving Approach to program evaluation, and essentially serves as an evaluation report. These data will be discussed in light of the proposed evaluation approach in Chapter 5. The current chapter begins with an introduction to the student participants, followed by a chronological account of the course using the data collection events as reference points. Informal analysis was conducted as the data were collected, so the influence of the findings on the progress of the course will be addressed as appropriate.

Student Participants

Table 4.1 presents each student’s name and grade level along with a list of the highest level courses known to have been taken in each subject area. (The list of courses is compiled from incidental data in observation notes and analysis of teacher interviews and conversations. Students were not asked to provide this information, and therefore it is not consistent for all students.) Also included for each student is a brief description with information that will help the reader
identify individual students named in the chapter. These descriptions may include reason(s) for taking the pilot course, notes on career plans, and personality assessments made by the researcher and/or the teachers.

<table>
<thead>
<tr>
<th>Student Pseudonym</th>
<th>Grade Level</th>
<th>Highest Level of Prior or Simultaneous Courses</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>12</td>
<td>Physics</td>
<td>In the course because she liked both subjects and both teachers. Not an outstanding mathematics/science student in the past. Planning to become an elementary level teacher.</td>
</tr>
<tr>
<td>Bryan</td>
<td>12</td>
<td>Physics, Precalculus</td>
<td>Took the course because he had taken all other math and science offered at the school, and he wanted to get into a computer engineering program in college.</td>
</tr>
<tr>
<td>Ed</td>
<td>11</td>
<td>Chemistry, Algebra 2</td>
<td>Enrolled because he liked both subjects enough to pursue an elective course. Had somewhat of a “class clown” reputation.</td>
</tr>
<tr>
<td>Elizabeth</td>
<td>12</td>
<td>Physics, Precalculus</td>
<td>Chose the course because of absence of homework and tests, but also because she really liked Mrs. M. Extremely introspective and reserved, and described as very religious.</td>
</tr>
<tr>
<td>Gabrielle</td>
<td>12</td>
<td>Physics, Precalculus</td>
<td>Chose the course because she plans to go into medicine. Probably the most knowledgeable of these students in terms of the two subject areas.</td>
</tr>
</tbody>
</table>

Table 4.1: Student Descriptions
<table>
<thead>
<tr>
<th>Student Pseudonym</th>
<th>Grade Level</th>
<th>Highest Level of Prior or Simultaneous Courses</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>James</td>
<td>11</td>
<td>Chemistry Algebra 2</td>
<td>This course was given to him as a schedule filler—would rather not have been in the course. Distant from other students; worked at maintaining &quot;cool&quot; image, even when it meant accepting grades that were lower than his ability.</td>
</tr>
<tr>
<td>Lindsay</td>
<td>11</td>
<td>Algebra 2</td>
<td>Scheduled this course to avoid wood technology. Not enthusiastic about mathematics or science. Very shy and quiet—described by the teachers as a recluse.</td>
</tr>
<tr>
<td>Marie</td>
<td>12</td>
<td>Physics</td>
<td>Enrolled because of fondness for the teachers and the fact that there would be no homework and no tests.</td>
</tr>
<tr>
<td>Sarah</td>
<td>12</td>
<td>Chemistry Algebra 2</td>
<td>Needed a course to fill her schedule, and the no homework and no tests aspect helped convince her to take this course to improve her GPA.</td>
</tr>
<tr>
<td>Tammy</td>
<td>12</td>
<td>Physics Precalculus</td>
<td>Took the course because of the teachers and because she thought it would help in college. Planning to become a veterinarian.</td>
</tr>
<tr>
<td>Tom</td>
<td>11</td>
<td>Algebra 2</td>
<td>Took the course at Mrs. M's suggestion. He was the least academically experienced student in the class.</td>
</tr>
<tr>
<td>Tyler</td>
<td>12</td>
<td>Physics</td>
<td>Took the course as a college preparatory exercise. He was a good friend of Ed (already described), and the two of them had a reputation for being off task.</td>
</tr>
</tbody>
</table>

Table 4.1: Student Descriptions
Chronology

Pretest

The course began on Monday, March 6, 2000, with two data collection activities. Students first took the pretest, designed to assess the students' ability to predict the type of motion required to produce a certain graph (for which a sketch is provided), and to predict (by sketching) the graph produced by a certain type of motion (see Appendix A). The teachers then introduced the evaluator, who explained his relationship with the teachers and the class. The teachers scored the pretests, which were not a part of the students' grade, and shared the scores with the evaluator. All scores are reported in Table 4.2.

<table>
<thead>
<tr>
<th>Score</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td># students</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.2: Pretest Scores

As explained in chapter 3, these tests were intended primarily as a means of comparison from the beginning of the class to the end; thus the scores from the pretest were not used for diagnostic purposes at this point by either the teachers or the researcher.
The Early Questionnaire, also administered on the first class day, focused on attitudes toward mathematics and science, relationships between mathematics and science, reasons for taking the course, and prior experiences with graphing calculators. Data were analyzed by the evaluator, and preliminary results were reported to the teachers. Students were told that they were not required to put their names on the questionnaires, and although some did, the questionnaire responses were not tracked or matched in any way to interview data and/or Late Questionnaires.

Data from the Early Questionnaire revealed that this group of 12 students basically viewed mathematics favorably. Positive responses included remarks such as “fun sometimes;” “I like most math classes;” “I think that they are pretty fun;” and “very enjoyable.” Only one student stated that mathematics was not enjoyable; one stated that mathematics was “not my favorite class”; and one said that mathematics was not enjoyable “unless I get to do only problems by myself.”

Science was portrayed even more positively. Students remarked that science was fun, primarily because laboratory-oriented classes were more “hands-on,” and two students stated that they liked to learn “how things work” and “why they occur.” One student declared science to be only “somewhat” enjoyable, explaining that “formulas are fine, but vocabulary is not very easy.” Another student stated that science class was “not very enjoyable,” explaining that the classes were “boring.”
The key question on the Early Questionnaire was "How are mathematics and science related?" This question was intended to serve as a point of comparison at the end of the course, in hopes of determining whether the students' perceptions of the mathematics-science relationship changed as a result of the course. The early responses fell into three categories: a completely reciprocal relationship; mathematics is used in science; and unsure. The categorized responses are presented in Table 4.3.
You can’t have science without math—they go hand in hand.

Graphing motion detectors involves math to graph it and physics to decide why it acted the way it did.

Well, in science there is always math from measuring units to figuring out genetics. And in math you can always use science to figure out questions.

Well, they both involve each other. Many of today’s problems have both some type of math or science in them.

You use mathematics to figure out scientific problems.

Science, especially physics, is math. You use it all the time.

Every reaction can be explained with numbers.

For many problems in science you must use math. There are many mathematical formulas involved in science.

Since I took physics, I found that to find out something in science there is always math involved (ex: velocity, gravity).

I feel mathematics helps explain science by giving proof and reasons why science stuff happens.

You use mathematics formulas & equations in the science.

I am not really sure.

Table 4.3: Early Questionnaire: Question 3. How are mathematics and science related?

The students consistently rated very high the importance of seeing and understanding relationships between mathematics and science. Several added that seeing the relationships can potentially increase understanding—either their own, or for people in general. Others reemphasized themes present in the previous question, including the reciprocal relationship between the two subjects and the view of mathematics as being an integral part of all sciences.
Reasons for taking the course varied widely, and some students gave several reasons for enrolling. Three students admitted they were taking the course because it would be easy—no homework and no tests. Four students thought the class would be fun; four also stated that they viewed the course as college preparation. Two students commented on their preference for these teachers. All these reasons were stated in at least fairly positive terms. Three students, however, gave neutral or negative reasons for their presence: as a schedule filler, because of a parent's decision, and simply that the course "was given to me without my request." The teachers had consulted with all the students prior to the beginning of the trimester, and were aware of all these reasons for enrollment.

According to students, they had limited experience using graphing calculators, even though they were consistently used in second year algebra, which all students had taken at some point prior to this course. Two students also reported some use in physics and precalculus. This level of experience was consistent with what the teachers expected, since they had also taught these previous courses. Seven of the students reported experience, although minimal, using data collection devices. These students had already taken physics, where they used only motion detectors, and had been taught by Mrs. S.

The Early Questionnaire was intended to establish baseline data on the students' perceptions of the connections between mathematics and science, their reasons for the course, and their prior experience with the calculators and equipment. The Early Interviews were conducted in order to gain further understanding of the students' questionnaire responses.
Student Early Interviews

The Early Interviews, which essentially paralleled the content of the Early Questionnaires, were held at the end of the second week of the course. The second week was chosen in order to give the students some time to adjust to the daily routine of the course, which meant that they could discuss their first impressions.

Although individual questionnaires were not matched on a one-to-one basis with interview data, it was clear that the students echoed their written responses in terms of their preference for mathematics and/or science, their ideas of the relationships between mathematics and science, and their reasons for taking the course. Their ideas of the relationships between mathematics and science were accented by comments such as “every time you turn around, you can’t have math without science...math describes science,” “they’re basically the same concept...they’re linked from the beginning,” and “math is kind of like proving what science finds out.”

Two students had interesting comments in which they expanded on their preferences for mathematics or science. Abby, a senior, explained that she liked having one answer in mathematics, and did not like science as much because she did not like having so many answers for one question. Tom, a junior, discussed the fact that in mathematics, he liked finding an exact answer, but “in science, for me, it looks like it might be the answer or it might not be—you can’t tell.”

First impressions of the course from several students included their perceptions that it was, so far, reminiscent of physics. One specific question in the interview asked the students to address whether they had been challenged
thus far. At this early stage, five students stated that they were not yet really facing any challenges—that either they had seen the concepts in physics, or were at least confident from prior mathematics experiences. Only three students admitted to being seriously challenged by the material to date. Most stated that they believed the course would become more difficult, either because of the natural progression of curriculum or because they just knew what these teachers were like.

One aspect of student impressions of the course to date was revealed in the interviews without a specific prompt: The students at this point had begun to see the class as first a science period, then a mathematics period. Despite the fact that they had just explained their own views of the relationship between the two subjects (with some students stating that the two were inextricably linked) and had consistently stated that this course was meeting their expectations, their comments showed that they were thinking of the extended block time as a science period followed by a mathematics period. The most revealing statement from the interviews was “at the end of class . . . we have to figure out the math part.” This perception by the students was reinforced by comments made in the class that the researcher heard on subsequent observation visits, including the day of the early interview with the teachers.

*Teachers’ Early Interview*

The Early Interview with Mrs. M and Mrs. S was conducted with both teachers together after school on the fourth class day following the early interviews with students. The time between the early student interviews and the early teacher interview allowed the researcher to conduct some preliminary
analysis of the student interviews and incorporate that data into the discussion with the teachers. It should be noted that even at this point, the researcher had begun a practice of sharing with the teachers information such as informal data analysis, interesting observations, and general impressions during each observation visit. This constant interaction, based on the influence of Stake’s (1975) responsive approach, proved particularly useful during this evaluation.

Goals for the Students

The discussion began with the teachers stating their goals for the students. Mrs. S explained that her goal was to show the students the relationship between mathematics and science, primarily because of many problems in the past with students who did not seem to bring mathematics knowledge with them into her science classes. Mrs. M joined in by saying:

I think that is absolutely the only goal we really have for them. I told them the other day that when they take a math class from me it’s okay for them not to bring science knowledge, but when they take a science class it’s not okay for them not to take their math knowledge.

The teacher’s comment shows a somewhat different perception of the relationship between the two subjects than had been presented by Mrs. S and the students. Mrs. M proceeded to explain a situation from a previous class in which the students did not immediately recognize a characteristic of a parabola that she thought should be very obvious to them—and would be if they were in a mathematics class. She expressed her disappointment, but Mrs. S gave a different perspective on the situation: Mrs. M had been out of the room when the data collection was conducted, and had not realized that the parabola itself was the result of deductive work on the part of the students. Mrs. S, then, was
completely elated that the students had been able to predict the parabola's shape, and was "extremely excited because that's more connection than they've ever made for me, so I was fine with it." In general, Mrs. S felt that students were starting to get/make connections, while at this point, situations like the incident with the parabola made Mrs. M feel less confident about the students' progress on this goal.

The issue of the students' perception of the class being first science, then mathematics, was presented to the teachers, and they were somewhat surprised. They stated that they had not intended for it to be set up that way, and explored several possible explanations for this perception. First, the setup of the laboratory activities was such that actions (data collection) always preceded any on-paper work (the mathematics involved in manipulating the data). Second, Mrs. M had generally been taking the first hour of the block as her planning period, so was absent while the "science" was going on. Although regrettable, her course load was such that the planning time was necessary, while Mrs. S had a lighter class schedule during this trimester, and was devoting all of her time—both hours of the block—to the pilot course. Mrs. S explained that although she can often help with the mathematics portion of the activities, she sometimes tells the students that they will have to wait until Mrs. M arrives for assistance with it. At the same time, Mrs. S admitted that she often puts off answering students' questions in this way just to promote more thinking on their part. Nonetheless, this approach added to the perception that once Mrs. M entered the room, it was time for mathematics.

79
Professional Development

The professional development aspect of this collaborative venture between teachers arose during this discussion. Specifically, Mrs. M described a situation that occurred in the classroom in which the activity required that an exponential function be transformed into a linear function (through Newton’s law of cooling), and told how Mrs. S became excited because she remembered having done that in college. Mrs. M said, “It’s kind of neat for both of us because we’re both learning things, or reremembering.” She also explained how they had used an entire lunch period (their only joint planning time for the course) just discussing the vocabulary that would be involved in a particular activity. They saw the need for such a discussion based on several experiences in the classroom where they had actually confused the students by using differing vocabularies (more mathematics oriented versus more science oriented) to talk about the same concept. “The kids thought we were arguing,” she added.

Early Student Progress

Both teachers were thus far pleased with the students’ progress overall, but went into detail about one student, Abby. Mrs. S said that she “just knows and sees” and is able to guide the other students through the activities. Mrs. M explained that Abby has generally had to work hard for the grades she gets (in Algebra 2, for example); however, students were already seeing that “if Abby can’t figure it out, then no one can.” They also indicated that they are encouraging her to teach mathematics and/or science, since she is planning to go into education as a career. For her part, Abby stated in her early interview that
"this is my favorite class ever; I can remember so much more in this class—I remember more than if I were taking all these courses separately." Her case will be discussed in more detail later.

**Pacing**

The teachers had already commented on prior visits that the activities were going faster than they had originally planned. That is, the original idea was to take a set of 30 activities as the basis for the course. Even before the course started, Mrs. S had a feeling that they would not be enough, but Mrs. M did not expect such a shortage at all. Mrs. S said she presented her concerns, but Mrs. M just replied, "My heavens, we have 30 labs!" Mrs. M expected to have to spend at least two days on each activity, but said, "There have been days when we've done two labs or one and a half in one day." They echoed the students' general feeling that the course had not yet been really challenging, and this impression was based primarily on the fact that both of them had prior experiences with these students—in Algebra 2 or physics, for example. The researcher then asked if the faster pace with the activities meant that the students would reach the goal—making connections—faster. Both responded that they thought it would take about 9 of the 12 weeks of the trimester to reach this point; based on this prediction, they were planning for the students to begin their individual projects (creating a laboratory activity in the style of those they had seen in the course) at that time.
Order of Topics

The teachers explained that there had been some changes from their original plans for the course—even this early (week three) in the trimester. First, their idea for teaching by the type of function (first linear, then exponential, etc.) changed in part because the students were going through the materials faster than predicted (and were asking when it would “get hard”), and because the teachers realized that the set of activities they had chosen were sequential not with respect to types of functions illustrated, but with respect to calculator features. That is, a calculator skill taught in one activity was expanded on in the next activity, and this progression was done in such a way that skipping an activity would mean having to backtrack to learn keystrokes on the calculator, or even to explore a background concept related to a new calculator function.

Student Grouping Procedures

A second change had to do with the way students were grouped for their activities and how they submitted their work. At first, the teachers allowed the students to choose their groups and turn in one group assignment for all group members. They found that some students were not participating fully under this system, and with no tests in the course, they had to find another way to make the students accountable for participation and true engagement with the activity. They chose to start assigning groups randomly (by drawing names on cards in a can), and to require that each student carry out the data collection and turn in his/her own laboratory report. These actions helped with accountability, and had the added benefit of providing more points of comparison in any given
activity—that is, rather than three or four sets of data (one from each group), they now always had 12 data sets (one from each individual) to compare and make predictions.

Effects of the Evaluation Process

Because such changes were occurring so early in the course, the teachers were asked whether the evaluation process was having an effect on them, both individually and as a pair. Both immediately denied that the researcher’s involvement was having any effect on their conduct. Mrs. M explained:

Your coming up here has not changed one iota what we’re doing. I mean [Mrs. S] and I would be doing everything we’re doing whether you were coming or not. . . . To me that’s how it should be, for you to get good data, you don’t want us to put a show on. . . . You’re seeing the raw product.

Mrs. S explained that “we don’t have critical friends in this area” and that they were excited that they would be getting another opinion about their work. Nonetheless, Mrs. M stated, “I don’t think your coming has motivated any of our decisions.” Mrs. S then revealed that they both had some experience with evaluation from a professional development course they had taken together through a nearby university one summer.

The Teachers’ Relationship With Each Other

The teachers then returned to a discussion of their students’ work and progress, and began to compare their own approaches and strengths as teachers. It was clear at this point that these two had been working with each other long enough in other capacities to know how they complement each other. Mrs. M stated, “I haven’t had the experience [Mrs. S] has at answering a question with a question—she’s the best I’ve ever seen,” then explained her own desire to
always work out every problem before assigning them to her mathematics students. Mrs. S jumped in and said, "I am the total opposite." Her philosophy is to just tackle a problem or question, and even if the students fall on their faces (because of an unforeseen or unpredictable circumstance), everyone should learn from the situation. For Mrs. S, then, problems are learning opportunities. This discussion of each other's strengths and weaknesses matched comments about the teachers made by the students, both in their interviews (both early and late) and in class observations.

Focus Group

Similar to the teachers' experience working with each other, all the students had experience working with these teachers in the past. This prior association is probably one reason the focus group was so productive. The focus group was originally intended to take place near the middle of the trimester, but because of scheduling conflicts it did not occur until the end of week 8 of the 12-week trimester. By this time, the teachers had completed their intended curriculum, and had begun to let the students do some exploration with the various data collection devices in order to develop ideas for their final projects. In a sense, then, they had come to the end of the "regular" portion of the course, and were in a different phase that was more of a culminating or capstone experience. Thus, the focus group had a much more retrospective tone than was originally anticipated, but it provided useful data.

To conduct the focus group, the researcher asked a co-researcher to accompany him to the site to take notes using a word processor on a laptop computer. Prior experience conducting focus groups had led the researcher to
believe that people tend to be more open in focus groups when they are not being audio- or videotaped, probably due in part to increased anonymity. In the case of this pilot course it was decided (by the researcher in consultation with the teachers) that the students would likely be just as open with the teachers actually present and participating in the focus group. Thus, the focus group consisted of all 12 students and the two teachers, with the researcher facilitating the conversation and the co-researcher taking notes. The participants’ comments were tracked by the co-researcher only by a number she assigned based on the seating arrangement. Therefore, although comments may be sorted and grouped by participant in the focus group, they cannot be matched by individual to any other piece of data. Although the questions designed in advance for the focus group were intended to parallel the early questionnaires and early interviews, the discussion took a direction of its own, which led to the following topical categorizations of the data.

**General Impressions of the Course**

Responses to opening questions basically established that the students were generally pleased with the course, and that the equipment (calculators and data collection devices) was helpful conceptually as well as necessary to the format of the course. Several students stated that it took anywhere from 2 to 6 weeks to become comfortable with the calculators and equipment, but once they did, they used the calculators elsewhere (in other courses) more frequently. Nine of the students participated in an extended series of remarks on the absence of tests in the course: The consensus was that tests should be added in order to give students clearer goals and to make them more responsible for their own
learning/understanding. Homework was not addressed in the same way, but one student added that although tests were necessary, homework was not, and no one disagreed. Other than these points, there were three main topics that received extensive attention in the discussion: the issue of repetition, the balance of mathematics and science, and one particular activity that seemed to be everyone's favorite memory from that portion of the course.

Repetition

The issue of repetition arose when the researcher asked what had been the most challenging aspect of the course (other than the equipment). The response was "staying focused; staying on task," and another student explained that every laboratory activity sheet was the same: "Every lab looks the same; it's boring." Most of the students agreed that even when the activity sheets came from different sources, the questions and formatting remained the same, creating a great deal of monotony. This topic was resurrected at several points during the focus group. Near the end, one student interrupted one topic of conversation to add, "Have you seen [the movie] Groundhog Day [Albert & Ramis, 1993]? Every time you step in here, it's Groundhog Day—same thing every day."

Balance of Mathematics and Science

The topic of balance between mathematics and science arose from a comment from one student that the mathematics portion of the activities took a long time. Mrs. M asked whether the students thought that the course should be a mathematics credit. Three students stated that they thought the balance between the two had been about even throughout the course; but the majority
(the remaining nine students) felt that the course had emphasized mathematics heavily. This emphasis on science is related to the students' earlier comments about the perceived separation of "doing science" (the laboratory portion of each class block) and "doing math" (the data analysis portion of each class block).

Favorite Activity

The third important point to come out of the focus group concerned a particular activity that the students found enjoyable and very memorable. On a Friday about four weeks into the trimester, the teachers sensed that the students were becoming frustrated, in part by the mathematics involved in some of the recent activities and in part by the monotony of the class activity routine. They decided it was time for something different, and Mrs. S had an idea: allow the students to use the pressure probe and any available materials to explore pressure readings. They turned the activity into a contest, and the group that achieved the highest pressure got to share a cake baked by Mrs. S on the following Monday. Students reported that this activity was the most fun of any they experienced, and it came up in the focus group discussion at several points. Not only was it a recurring topic, but the majority of students also labeled it as their favorite activity of the entire course. It is important to note that this activity was in stark contrast to the usual, highly structured activities the students were given on worksheets. Its success contributed to the teachers' idea for the group's final project: to create an activity like those on the usual worksheets by exploring their own concept with the probes and equipment. Indeed, the idea of
open discovery using the equipment became a dominant factor in the final portion of the class, which had already begun by the time the focus group was held.

Assessment of Group Progress

One important result of the teachers being involved with the students in the focus group was that the students had the opportunity to react to the teachers’ general assessment of their progress as a group. At one point, Mrs. S offered a lengthy comment on the fact that she had seen definite progression in terms of making connections between mathematics and science, and added that the students had come from a point in the beginning where they were constantly asking questions, always having the teachers come over to help, to now simply proceeding with great confidence through an activity. In fact, the students were now showing the teachers how to use the calculators for certain aspects of data collection. The group had become much more self-reliant, and the teachers had noticed. The researcher then asked the students if they had felt themselves progressing in the ways Mrs. S mentioned. Six students “kind of” agreed, four said “no,” and only two agreed that they could see the ways they had changed. The students were reluctant to comment further, which may have indicated they needed more time to process their teachers’ evaluation of their progress. This topic was addressed again in most of the student late interviews.

Effectiveness of the Focus Group

The focus group was a very productive means of obtaining information from all participants at one time, since obviously all were present in the room at once. But beyond being efficient, the format allowed for the observations and
impressions of one group of participants to reverberate off those of another group, in terms of both differing opinions among the students and student opinions compared to those of the teachers. Specifically, there had been no other opportunity to date where the teachers would have given such a summary statement regarding the students' progress and conceptual growth—indeed, Mrs. S later stated that she was prompted to do so by listening to some of the students' comments. The focus group was viewed by the teachers (as revealed in final interviews) as the single most effective event of the evaluation process, primarily because of this interactive feature of the focus group format.

The content of the focus group discussion guided all data collection events that followed. The influence of the focus group was in part because of the topics that came to light during the discussion, but also because the students and teachers, in interviews and/or questionnaires, referred many times to the focus group, and used their memories of it to guide their responses—even responses to questions that were not directed at focus group outcomes.

_Late Questionnaire_

The Late Questionnaire followed the three-week period during which students experimented with equipment, planned and created formal write-ups of their individual projects, and individually led the class in conducting the activities they developed. Nonetheless, the influence of the focus group discussion was obvious. In part, this influence was due to the fact that this portion of the trimester was qualitatively different in the way it was conducted, and the issues that were raised in the focus group (repetition, balance of the two subjects, and the preference for experimentation with the equipment) were, to varying
degrees, addressed by the change in format of the class. The students were no longer entering a classroom to do “the same old worksheet” activities; they were focusing on either science or mathematics or both to the degree necessary for formulating their own projects (rather than as directed by questions on an activity worksheet); and they were constantly experimenting to develop their own uses for the various sensors and detectors. Thus, when asked about the degree to which the course was “enjoyable,” the majority responded favorably, but with reference to the repetitious nature of the earlier portion of the trimester. Two students could not get past the monotony of the early part of the course, and therefore rated the course as a whole not enjoyable. When asked whether the course met expectations, seven students again mentioned the repetitious nature of the activities; seven students did find that it met their expectations, while three said that it did not, and two said they really did not have expectations going into the course.

**Relationship of Mathematics and Science**

Some questions were repeated on the late questionnaire to offer points of comparison from the beginning of the trimester, although individual student responses cannot be matched. The first such question regarded the nature of the relationship of mathematics and science. This time, the students were asked to provide an example of the relationship they described. Their responses fell into four main categories: a relationship through graphs, a relationship through equations/functions/formulas, the idea that mathematics proves or explains science, and the idea that mathematics is in science. Categorized responses for all students are provided in Table 4.4.
<table>
<thead>
<tr>
<th>Category</th>
<th>Student Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship through graphs</td>
<td>Mathematics is related to science in the main form of a graph. In science you experiment and test different things and using math you can find the relationship mathematically by using graphs. The main connections I see are through graphs. Most of the time we saw the relation in graphs. Example: We would walk in front of the motion detector and see what kind of graph it made.</td>
</tr>
<tr>
<td>Relationship through equations/functions/formulas</td>
<td>You use math stuff to figure out science stuff. You use equations to figure out things like speed, velocity, etc. You really don’t need science in math, but you need math in science. Mathematics equations show scientific relationships. If you make a graph of a balloon falling you can use the formula for slope to determine its terminal velocity. Almost everything in science has some math formula that controls it.</td>
</tr>
<tr>
<td>Mathematics proves or explains science</td>
<td>To prove anything dealing with science, you have to use math. If you are talking about slopes and velocities, you need to use math to figure out the answers. Everything in science can be shown or proven with math. pH levels and oxidation numbers are examples. Math explains what science shows.</td>
</tr>
<tr>
<td>Mathematics in science</td>
<td>It’s like 2 in 1. Where there is science you always have math. Like in the motion labs. You always have the science and then you can find out things like velocity. You use math in science all the time, for example finding the slope of a distance time graph and knowing that’s velocity. You use math to do science questions.</td>
</tr>
</tbody>
</table>

Table 4.4: Late Questionnaire: Question 2. How are mathematics and science related? Give an example of this relationship.
When compared to the responses to the parallel question in the Early Questionnaire, these statements are more specific in terms of the type of relationship between the two subjects. On the Early Questionnaire, seven students gave responses that indicated that mathematics is "in" science; one student mentioned graphs, two noted formulas, and one included the idea that mathematics proves science. In the later responses, however, these ideas are more widespread among the students, and are stated more directly. One reason may be that the students were further prompted in this case to "give an example." However, it may be in the case of graphing, for example, that the students have been taught through their use of graphing calculators to think in terms of graphs as the basis for seeing the mathematics in science. Similarly, the students frequently used the regression menu on the calculators to determine a function for the graph they had created through data collection, and thus may have come to think of this procedure as the link between the two subjects. Therefore, the responses may well be a function of the equipment used as the primary teaching tool.

Importance of the Relationship of Mathematics and Science

Just as they did early in the trimester, students consistently rated high the importance of seeing and understanding the relationships between mathematics and science. On this questionnaire, students were also asked how the course had affected their thoughts on this matter. One stated that the course "made me realize how often we use our skills in other areas of study," while another said it
had "helped me with some misconceptions I had." Others stated simply that the course had made a difference, and one indicated that the course had affected views of science more than views on mathematics.

**Greatest Benefit**

Despite these changes and apparent progress in terms of the original goal of developing relationships between mathematics and science, students’ responses to the final prompt on the questionnaire were very revealing in terms of their view of the primary goal. They were asked, “What was the greatest benefit for you from this course?” Only three students indicated that the mathematics/science relationship was the greatest benefit, and of those, one included the fact that calculator skills had greatly improved. The remaining nine students all responded in terms of their experience with the various data collection equipment and their improved calculator skills. For the students, then, the extensive use of the calculators and equipment—the tools and the ability to use them—seemed to be the most important outcome of the course.

**Student Late Interviews**

The late interviews with the students were conducted just after the questionnaires and across the last two days of the course. Although the prompts were worded and ordered differently, essential goals were the same as for the early questionnaires and interviews. Students were again called individually to an office area to the side of the classroom. As with the late questionnaires, many of the students’ responses included references to topics first discussed in the focus group. Indeed, all but two students repeated their concerns about the monotony of the laboratory activities when asked about recommended changes.
for the course, and several of them specifically referred back to the focus group discussion, saying "like we were talking about before," or "like they said at the focus group." One student said that such repetition could be good for some students, though, and maybe it was best after all—even if he did not like it himself.

**Student Grouping Procedures**

Students were also prompted to discuss group formulation techniques used by the teachers. While students were at first allowed to work with any partner(s) of their choice, eventually the teachers began assigning groups by random means. Although the students were reported to have disliked this method when it was first instituted, statements made in the interviews show that they understood why the teachers needed to change the system (for accountability purposes), and several stated that it really did not matter that much because everyone there knew everyone anyway and they could all work easily together.

**Course Assignments**

A third evaluative issue that was specifically prompted in the interviews was that of assignments related to the course. Students were asked if they thought it was best to conduct this class as it had been—without homework and tests. All stated that homework would be virtually impossible given the nature of the equipment and the activities themselves, and acknowledged that that decision was appropriate for this course. However, six students indicated that they believe there should be tests—or at least quizzes or some sort of assessments (even self-assessments)—in such a course. Elizabeth stated:
I definitely think there should have been a couple of tests. It would have forced us to learn and pay more attention, if you will. That's not to say—I know I learned a lot, but it would have been deeper in my memory if I had had to study it and remember.

Another student, Sarah, added a rather negative assessment of herself, and said:

I just found out my results of my pre- and posttest and I went down, so I think having homework or maybe a quiz or something would help you learn it more, because I obviously didn't learn much. Not like a bunch of them, but maybe one every 3 or 4 weeks. I think I would have tried to understand more. I really just kind of said I don't have to understand, I just need to get the labs done and I'll be okay for my grade in here.

Gabrielle called for a different type of accountability: rather than having tests, she suggested the possibility of having more projects like the final project, or at least having longer laboratory activities that would seem more like projects. At any rate, the idea of being held accountable and in some way having to prove what had been learned was a concern for these students.

Topics Studied

The opening questions of the interview were not related to issues that had been brought up previously. The students were first asked to simply list some of the topics they had worked with in this course, first mathematics, then science. The most interesting aspect of their responses is that, despite their own feelings, as expressed in the focus group and even in these interviews, a number of students had trouble distinguishing between mathematics and science topics. When prompted for mathematics topics, one student included acceleration, another stated “Velocity, pressure, we did like pH stuff—a lot of stuff that is like math and science,” and still another said, “Voltage and stuff? Motion, velocity, acceleration, speed and all that.” After additional prompting from the researcher, these students were able to list more strictly mathematical topics they
had dealt with, such as slope, linear regressions, and graphing in general. One student added that in terms of mathematics, there was distinct overlap with the precalculus course in which he was enrolled this same trimester, overlap “even to the point that the very same day we would be covering something in precalc and then come downstairs after lunch and do the same thing here.” Several students pointed out similarities to other mathematics courses, including Algebra 2, and many related the pilot course to physics (“It was pretty much physics with a calculator,” one student added), as they had in earlier data.

Student Confidence

To assess the levels of confidence in the course, students were asked whether they would be willing to serve as a lab assistant in this course if it were offered again. Six students responded positively, while three said they definitely would not assist. One of the positive respondents, Abby, wanted to go into teaching, and thought such a role would be good experience, and another, Tammy, said such a position would be a way to “learn the stuff I already learned even better.” Those who responded negatively explained that they did so mainly because they knew they did not want to go into mathematics or science related occupations, and would not want to spend time in this way. The remaining students were undecided as to whether they would take such a job, but nearly all the students described themselves as comfortable enough with the equipment and material to be qualified for such a position.

Indeed, nearly all of the students developed to some degree an ability to investigate and troubleshoot with the equipment when things did not work out the way they expected them to, or the way that another group’s data did on a
given activity. The need for such troubleshooting abilities came up on several occasions during the interviews. Elizabeth complained about the very structured nature of the laboratory activity worksheets, saying:

We just got the labs and it told us step by step what to do. . . . Once they gave us a packet . . . that led us through things, but it's kind of hard to see it unless you actually play with the programs.

James, when asked about suggested changes for the course, stated:

Something I don't think the teachers can do anything about is the way those labs are set up, it doesn't give you any information about the labs, it just says follow the instructions. So you really don't know what you need until you get into it. . . . Troubleshooting is worthless until you get into it [the lab activity]. . . . I don't want the lab holding our hands or anything, but a little bit of [background or conceptual] help is nice.

The students' ability to troubleshoot with the equipment and laboratory materials was witnessed on numerous observation visits, during the time that they were completing prearranged labs as well as when they were exploring with the equipment to plan their individual projects. Their ability to see and understand the need for troubleshooting skills shows that the students had gained great confidence in use of the equipment. It also demonstrates that they were looking beyond what the calculator told them was the answer—they were using true analytical skills, comparing what they predicted the data would look like to what they saw and reconciling any differences between the two.

Overall, the students' late interviews provided an opportunity for them to sum up their experience. It was clear that they were openly evaluative of the course, the teachers, and themselves. It also gave the researcher greater insight into the students' perspectives on the course, and eventually provided points for discussion in the Teachers' Late Interview.
Posttest

The posttest, a repeat of the pretest instrument, designed to assess the students’ ability to predict the type of motion required to produce a certain graph (for which a sketch is provided), and to predict (by sketching) the type of produced by a certain type of motion. The posttest was administered on the last day that all the students were together. (Seniors were dismissed a few days earlier than the other students.) It was hoped that overall scores would improve, and they did. The pretest mean was 11.3, while on the posttest the mean was 14.4 on a scale of 0 to 20, for an average score gain of +3.1. Pretest and posttest scores are listed for comparison in Table 4.5. Individual student scores with individual gains are presented in Table 4.6.
<table>
<thead>
<tr>
<th>Score</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td># students pretest</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># students posttest</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5: Pretest and Posttest Scores

<table>
<thead>
<tr>
<th>Student</th>
<th>Pretest Score</th>
<th>Posttest Score</th>
<th>Gain/Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>13</td>
<td>12</td>
<td>-1</td>
</tr>
<tr>
<td>Bryan</td>
<td>11</td>
<td>18</td>
<td>+7</td>
</tr>
<tr>
<td>Ed</td>
<td>11</td>
<td>16</td>
<td>+5</td>
</tr>
<tr>
<td>Elizabeth</td>
<td>14</td>
<td>16</td>
<td>+2</td>
</tr>
<tr>
<td>Gabrielle</td>
<td>13</td>
<td>17</td>
<td>+4</td>
</tr>
<tr>
<td>James</td>
<td>13</td>
<td>14</td>
<td>+1</td>
</tr>
<tr>
<td>Lindsay</td>
<td>8</td>
<td>13</td>
<td>+5</td>
</tr>
<tr>
<td>Marie</td>
<td>14</td>
<td>19</td>
<td>+5</td>
</tr>
<tr>
<td>Sarah</td>
<td>11</td>
<td>10</td>
<td>-1</td>
</tr>
<tr>
<td>Tammy</td>
<td>14</td>
<td>17</td>
<td>+3</td>
</tr>
<tr>
<td>Tom</td>
<td>5</td>
<td>6</td>
<td>+1</td>
</tr>
<tr>
<td>Tyler</td>
<td>9</td>
<td>15</td>
<td>+6</td>
</tr>
</tbody>
</table>

| Mean: 11.3 | Mean: 14.4 | Avg. Gain: +3.1 |

Table 4.6: Pretest to Posttest Gain by Student
From Table 4.5, it can be seen that posttest scores are higher overall, and that one student (Marie) actually scored 19 out of 20, or 95%. The comparison of scores by individual students (Table 4.6) shows that almost half of the students (5 of the 12: Bryan, Ed, Lindsay, Marie, and Tyler) gained 5 or more points, or 25% or more of the overall points available. The highest gain was 7 points (Bryan). When looking at overall success of the pilot course as defined by improvement in test scores, these students’ scores can represent such success.

However, several factors must be taken into account in analyzing these scores. First, students may have answered correctly on the pretest without truly understanding a concept. Indeed, in an early interview, one student indicated his confidence that he had answered all the questions on the pretest correctly. Clearly, however, no student had a perfect score at the beginning. Since there was no other point during the course where all the information studied had been put together in one place at one time as it was on this posttest, it may be that some students confused some of the concepts. Finally, it should be noted that the posttest was so near the end of the school year that many students, especially the seniors, might not have taken it seriously. The teachers reported that one student, upon receiving her copy of the test, even remarked, “As if I really care.” This type of attitude (colloquially referred to as senioritis or spring fever) may well have been a factor in the students’ approach to the test, and may have had an effect on the results.

For these reasons, the teachers and the researcher did not place high value on the test results themselves as an indicator of course or student success. Although the teachers were generally pleased that the class mean on the test
rose, they were satisfied with that as enough evidence of student improvement to report to their funding source. For themselves, they relied on other evidence, as will be seen in the discussion of the late interview with the teachers.

**Teachers' Late Interview**

The Teachers' Late Interview was conducted several days after the posttest and the Students' Late Interviews, and after the school year had officially ended. Almost two hours was spent with the teachers together, followed by 15 to 20 minutes with each teacher separately. Although some prompts were obviously intended to provide information for points of comparison from the beginning of the course, others provided information about the students that had been gleaned from earlier data and served as an opportunity to gauge their reactions to the students' evaluative comments.

**Achievement of Goals**

The interview opened with a question about the degree to which goals had been reached in the course. Although Mrs. S indicated that the students had gone "above and beyond what we expected," Mrs. M simply stated that they had achieved exactly what they wanted to with the course. Both agreed that this experience had convinced them that their next project would be to develop a similar course aimed at lower-level students. According to Mrs. S, "If it worked so well with these upper level kids, it ought to be fabulous with the kids that really need the hands-on stuff." But Mrs. M's comment that the course had not surpassed the set goals prompted Mrs. S to tell her why:

Mrs. S.: You didn't know where they were coming from. You got your eyes opened this year.
Mrs. M: Exactly, I got my eyes opened. When [you] said they didn't get those connections, I just thought [you] were crazy, I really did, I'm sorry. I just cannot, I still cannot believe some of the connections they weren't making...And the fact that they did not make some of the connections that I [could see]... It was just mindboggling. That's the only word I know to use—it was mindboggling.

Later they agreed that although they differed on whether overall goals had been met or surpassed, they both thought that the students had gained more than they (the teachers) had expected.

Since their main goal, according to the early interview, was for students to make mathematics-science connections, the researcher asked whether the students could now make those connections. Mrs. S indicated that they could, but in most cases throughout the course, "they almost always needed help to make the connection." Nonetheless, she said, "for me, their progress has been leaps and bounds" in terms of connections, and described how pleased she was to hear what she called the "aha! moments"—those times when the students indicated that they "got it." She explained that this happened especially near the end, when the students were working on their individual projects.

Mrs. M's response about the connections indicated a different result: "The biggest thing that this course did for those kids is it taught them how to think."

She went on to describe the type of thinking:

It was not black and white, which is what they're used to in a math class, they're used to black and white, and science has always been gray, but you throw in some gray math, and they're in trouble. And it just made them stop and think.

102

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Later comments on individual students showed that the teachers also thought the benefits of the course to the students included independent thinking, a newfound self-confidence in some cases, and in at least one instance, social gains that far outweighed any academic achievements.

**Individual Student Assessments**

The teachers were asked by the researcher to provide an assessment of each student, tracing that student’s progress, and summing up his/her performance and accomplishments. It was clear that they had already discussed individual students between themselves before being asked the question, and that they had very strong feelings about each student. Tears came to their eyes as they described several of the students, and their long-term connections from other courses with the students were evident.

**Independent thinking skills.** They spoke of gains in independent thinking skills for four students: Elizabeth, Sarah, Gabrielle, and Tom. Elizabeth was described as being formerly dependent on approval for everything. At the beginning of the course, she asked the teachers about every answer on her activity sheets—so frequently that the teachers did not have to examine it any further to assign a grade. Mrs. M characterized her as a memorizer in a previous math course. Both recounted an incident that they referred to as a *meltdown*—something they said happened with several of the students—that was a turning point for her. The researcher visited on the day just after this meltdown had occurred, and recorded additional information in observation notes at that point. The meltdown was an emotional breakdown that occurred because of frustration with the material. In this case, Elizabeth was working with
an activity that involved two temperature readings, referred to as T1 and T2, and two velocities, V1 and V2. All these values went into an equation that simplified to a point where the V2 value appeared in both the numerator and denominator of an expression, and therefore could be cancelled. Elizabeth, however, could not see the cancellation, and became stuck with a very complicated expression. Her frustration caused her to say a curse word—something she never does—and burst into tears. Mrs. S was working with Elizabeth at the time, and described her problem with the cancellation as one in which she had attached “too much meaning” to the values involved, and therefore could not see the cancellation possibility. In general, this was a very frustrating time for the class as a whole, because this was also the week that the students actually showed frustration with the laboratory activities that were provided. It was the same week that the teachers decided to use the pressure contest as a break from the usual tasks.

Sarah was another student who was described as gaining independent thinking skills—but she was said to do so out of frustration. Sarah was one of the students who did not want to be in the course in the first place, and in general, based on the researcher’s observations and teacher input, did not have a positive approach toward the class and activities during most of the trimester. She was also the student who mentioned in the late interview that her score on the posttest was lower than that on the pretest. Her attitude was evident as the teachers handed out the posttest: Sarah was reported to have taken the test in her hand and said, “As if I really care.” Mrs. S did say that she made some progress in the course, but Mrs. M described her in this interview as “my biggest disappointment.” This disappointment was primarily because Mrs. M knew
Sarah from a previous course (Algebra 2), and she knew that she had put hard work into that course. That was not the case here, and both teachers agreed on that point. However, Mrs. S added that “when [Sarah] was pushed to the edge, she would think and produce what she could,” and would do so independently. She did not think that Sarah was able to make connections yet in the way that the other students did, but her impression was that it was mainly because Sarah just did not want to do so.

The third student who gained independence in thinking was Gabrielle. Mrs. M described her as a “person pleaser, [who] gets very upset if she thinks she’s supposed to know the answer and she doesn’t.” She went on to describe a day when Gabrielle told her that she felt she had let her down because “there was some math that she knew she was supposed to know and she didn’t know it.” This too, Mrs. S added, was an instance when a student was in tears—another frustration meltdown—but that Gabrielle did eventually get to the point where she was okay with not knowing the answer ahead of time and having to think her way through a situation.

Finally, Tom was described as gaining independent thinking abilities. From both the teachers’ description (“Everybody in the class was a step ahead of [him]”) and the researcher’s observations, Tom was probably the least experienced student in the group and was academically less empowered than his peers. He was a junior, and had not taken some of the same courses (precalculus, physics, etc.) that others had. But, according to the teachers, “He doesn’t know how much he got out of this course.” Mrs. M said, “He gave us everything he had.” They spoke partly in terms of his work ethic and
personality, and Mrs. S added that she hopes “he can see that he can hang in with the best of them, and he may not be on top, but he’s there.” His final project on the period of a double-string pendulum was one of the best in the class—and Mrs. M ended this part of the discussion with “he got an independence that he doesn’t even realize he has yet.”

**Self-confidence.** The teachers added that Tom’s independence was similar to the self-confidence they had seen develop among some of the other students. Specifically, Tammy and Bryan were seen to have gained a great deal of self-confidence. Tammy was described as one who could think and use information and knowledge and put it together already, but she did not have the confidence to present that as a whole—to say that she knew something was right because she had thought it through. This course, however, according to Mrs. S, helped Tammy become able to say “yes, I can come up with the correct answer, and yes, I know this is right and here it is.” Mrs. M added that “this class made her realize that her best is right,” and that maybe she has much more intelligence than she had previously thought. They added that although Tammy had been saying for a long time that she wanted to become a veterinarian, they were not convinced that she was convinced she could do it until their experience with her in this course. Bryan was also described as having gained self-confidence, but his self-confidence was mainly because of his previous performance in mathematics courses.

**Social factors.** In two cases, the teachers mentioned social issues as major factors in the growth of individual students. For Lindsay, who was described as being a recluse when she came to the school as a freshman, the social benefits
she gained far outweighed any academic progress in the teachers' minds. They discussed the changes they had witnessed in her behavior from the beginning of the trimester, when she would sit alone at the back of the room, to the end, when they witnessed her laughing and joking around with Ed, who was possibly the most outgoing student in the group. Mrs. M stated, "I don't care if [Lindsay] made any math or science connections or not," and Mrs. S explained that Lindsay had really "blossomed," and needed that more than any academic achievements. The teachers believed that it was the nature of the class, namely the structured group time with constantly changing partners, that helped Lindsay develop in this way.

Social factors were also key in James' progress, but in a different way. James was described by Mrs. M as "practically a genius," and she deliberately waited to discuss him last. She described him as someone who would never admit what he got out of this course—perhaps not even to himself. For James, the focus was on trying to fit in, but on his own terms, even to the point where he would "not work in a class and deliberately get a B just to be like everybody else." For him, then, to really accomplish something in this course required finding something that excited him. Not everything in the course did excite him, and even when it came to the final project, he procrastinated for 7 of the 10 days they were given for preparation because he was not able to find a topic that truly interested him. When he did, however, he even took it home, which was where the social issues came into play. James did not go to the prom that weekend, nor on the junior-senior trip to a nearby amusement park the next day, since, as the teachers described him, he was simply too detached for that kind of social
activity. Instead, he worked all that weekend on his project: a test to determine the difference of temperature created by light on two different colors of fabric. Further, the mathematics behind his work was quite complicated, perhaps the most complicated of all the projects in the course. He showed that he was truly making those mathematics-science connections when he realized that there was not a function already programmed into the calculator to match what was happening in the situations he created. Finding a function to fit his data became a puzzle for him, and that excited him. He explained in his late interview that he basically had three sets of data, and he worked with different combinations of the values until he created a graph that matched the data. Once he found one for one set of data, he tested it out on the other sets, and it worked for all three. The teachers explained how excited they were that he had taken such a skilled approach to the problem, and lamented that, although James had hoped to see whether the function he had created would work for all the data collected by the other students when he presented his project, time did not allow them to confirm his work. Nonetheless, James was clearly proud—and so were the teachers.

*Remarkable students.* Through analysis of the teacher interview data, the researcher found that James was one of four students who seemed truly remarkable—students whom the teachers described in great detail, with much emotion, and/or with surprise or admiration. Two other such students have already been mentioned: Tom, who held his own with more academically
capable students; and Abby, who showed great promise as a teacher through her remarkable ability to see and make the connections and help others in troubleshooting the equipment and materials.

The teachers also went into great detail about one other student: Ed, who has already been mentioned as the very outgoing one. Mrs. M described him as a student who is extremely intelligent, but who does not apply himself—a description that the researcher expected, based on observations. However, Mrs. S revealed a surprising fact: “Basically we had to design the class around him in order to force him to work independently.”

Mrs. M: Right. If we made a change in this class, we made it so that [Ed] would do what he was supposed to do. It really aggravated us. We started out with the kids turning in one lab per [pair or group of] partners. Well, we couldn’t do that because [Ed] would clown around because he knew his partner would write it up. Then we went to individual lab write-ups.

Mrs. S: But they shared the data [via calculator links]. And so he just copied the shared data. So then we said everybody has to [generate individual data].

So, it became apparent that while the changes made in procedure were on the surface aimed at a group of students who were not behaving appropriately, they were really made to accommodate the actions of this one student. Apparently, although there were similar behaviors in other students, none were so severe as with Ed. They continued by describing him as a “grade-driven” student. Mrs. S admitted that the lengths they went to in order to get this one student to behave, while unfair, did not hurt the other students in any way. Indeed, Mrs. M concluded:

In the long run...it could be that [Ed] was doing them all a favor. Because...it made them all independent, and that might not have happened if it had stayed the way [we] had planned.

109
Repetition

The teachers referred several times to the ideas of repetition in the course and the break in that monotony that happened on the day of the pressure contest. More than once, they indicated that they just wished they had known earlier how the students felt—they thought that maybe knowing that could have helped them make changes. Mrs. M admitted that their complaining, when it did happen, was probably taken a bit personally because from their point of view, it was not the same old thing every day. For the teachers, every day was a new science concept or task and a new mathematics function or analysis. They explained in the interview that they did eventually understand the students' point of view, however:

Mrs. S: We didn't see that [early on] because it was different keystrokes [to us]—well, to them it was keystrokes. There was a different data collection device, but it was data collection.
Mrs. M: It was a different algebraic function, but to them it was a function. I mean even though [we] weren't seeing it as monotonous at all, they were.

They mentioned several ways the monotony could be broken, including the use of different sources for activities and more open-ended activities, such as the pressure contest. They discussed the way that they had come up with that contest. When they were planning it, Mrs. M wanted to set up more rules and procedures than Mrs. S did, and even as the students began, Mrs. M had concerns about the many ways that the students attacked the problem. Mrs. S was very happy with the situation, and saw that the students were making connections with former science activities, linking gas, force, and depth in water to get higher pressure readings. She was convinced, then, that such an activity
was better without rules. The teachers continued by discussing various formats for exploring with the different probes and equipment, such as having every Tuesday as an exploration day. In general, the conversations showed that they had begun to think toward more open-ended activities, in part due to the success of the contest day and of the individual final projects.

Effects of the Evaluation

One question that had been asked in the early interview was repeated here for specific comparison. The researcher was interested in knowing the teachers’ perceptions of how the fact that this pilot course was being evaluated affected their actions, choices, and approaches to teaching. In the early interview, they were insistent that the researcher’s presence was having no influence—that nothing was being adjusted for the researcher’s benefit or to accommodate him in any special way. In the late interview, their opinions were different: Both stated that the evaluation process had helped tremendously. Mrs. S then said:

You know I almost feel guilty because our first answer to this question was “just because you’re here [we’re not letting that change the fact that]...we’re going to help the kids” but to have the opportunity to have those reflections and know what was going on in their heads—things they would never tell us.

They explained that without the interviews, questionnaires, and the focus group, they would have had to guess what the students were thinking. But, Mrs. M added, it also helped that the students knew they were thinking as they were. They agreed that the researcher’s presence, the questions asked, the approach to the focus group and all aspects of the evaluation had helped to pull the students together in a way that does not happen in a regular class. Overall, they felt that
the evaluation was the single most important factor in what they considered to be the success of the course. It is important to note, however, that a program evaluation is not always as welcomed as it was in this situation. If the success of the course was due in part to the evaluation itself, then it must be acknowledged that the success of the evaluation was due largely to the openness of the teachers to such professional scrutiny. Their eagerness to engage in an evaluation may have resulted from their lack of “critical friends,” as Mrs. S stated early in the course. More importantly, the teachers clearly demonstrated their willingness to be risk-takers through all that they did to develop and carry out this course; credit for success should go to the teachers for their willingness to step out of the typical molds of science and mathematics teachers.

Individual Teacher Interviews

Immediately following the interview with both teachers, each one was interviewed individually. The primary purpose of the separate interviews was to determine how the teachers had influenced each other in their work together. Each teacher was prompted to analyze personal gains and insights, to name the most surprising aspect of the course, to discuss the influence of the other teacher’s approach to teaching and learning, and to describe the ways that this experience will affect teaching in other courses.

Mrs. M. Mrs. M indicated that what she had gained personally was a new respect and value for hands-on learning. Although she was not sure how she would approach it, she indicated a desire to make her mathematics courses more activity oriented with hands-on explorations. She said, “I somehow have got to quit worrying about quantity and start worrying about quality...somehow I
have got to become more hands on.” The most surprising aspect of the course for her was the final projects: “They were much better than I ever thought they would be.” She feared that the students did not have enough guidelines to complete the projects, but Mrs. S assured her they would be fine, and they were. “I never thought they would come up with the quality of work that they did—that was the biggest surprise and so rewarding to me.” This reaction to the final projects illustrates what she felt would be the greatest influence of Mrs. S on her approach to teaching: “I learned to quit giving so much and start asking—instead of telling the kids how to do it, I learned to say okay, what do you think?” Mrs. M indicated that she had already been marking spots in her Algebra 1 and 2 texts where she would at least demonstrate next year some of the activities they had pursued in this course. She concluded the interview by emphasizing her desire to create a course similar to this one but oriented toward lower-achieving students. After all, she said, “If it can do this for the good students, I can’t imagine what it might be able to do for the kids that struggle with math and science.”

Mrs. S. Mrs. S stated that her greatest personal gain was learning that “the kids need to be given the time to make the connections. It’s very important that they’re allowed to sit there and think and share...that’s more important than what I thought.” She used the metaphor of an onion peel to describe the students’ growth—there is the core, and each layer adds depth to the overall understanding that makes up the whole onion. This impression of student growth was different from what she used to think, in that she now realizes that the structure of the courses she had taught traditionally, with state objectives and
a tight curriculum, had not allowed for the time to make those connections. In fact, she seemed to have changed her mind about being “disappointed” that the students were not previously making the connections she had expected. Now she sees that an experience such as this one is necessary. Her reaction to the evaluation process was that she only wished some feedback had come earlier—especially on the topic of repetition and monotony of the tasks. The information that did come helped her make decisions about what to do next and “how to tweak the course,” but earlier feedback could have improved that greatly. Mrs. S was most surprised by the number of mathematics-science connections the students made, and cited as an example the students’ skills by the end of the course at reading and interpreting meaning from graphs, which was a distinct weakness at the beginning of the trimester. She admitted to her reliance on Mrs. M for calculator expertise, and the fact that the way students were required to explain and clearly display steps in their mathematics work would affect the way she approaches teaching chemistry in the future. She is hoping that the task of explanation and inclusion of detail behind thought processes will help with “ownership” of the material—something she feels was a definite plus in the pilot course. She will also use the calculators and equipment for activities, some of which will come from the pilot course, in both chemistry and physics. Mrs. S wrapped up her interview with additional comments on the benefits of working with Mrs. M.:
Being able to hear [Mrs. M’s] insights, too, not only the kids’...she and I would sit, we wouldn’t know the questions to ask ourselves to be able to sit and converse, you know, we’d converse about the course, but I think maybe we get tunnel vision sometimes, and it’s really helped me with thinking about where the direction is and going, and just hearing what [she] had to say and different points of view. We still have opposite ways of viewing the universe that we’re not going to change, and I think it’s good for the kids to see that, that there isn’t just one right way—I tend to do the guess and check method, but it’s good to be able to sit down and pick things through, and that’s something else that has been brought out that was good.

With this comment, Mrs. S brought out specific problem-solving approaches: the guess and check method, and what she called being able to “pick things through.” This comment—with a direct reference to problem-solving approaches—leads to the next chapter, in which the problem-solving process is examined from the perspectives of the participants and in relation to the Problem-Solving Approach to program evaluation.

Evaluation Summary

The evaluation of the Integrated Laboratory Course showed that the program had some degree of success in achieving the goal of helping the students to grasp connections between mathematics and science. This conclusion is supported by the overall increase in the scores from pretest to posttest (although some individual cases did not necessarily show increased understanding), by the teachers’ assertion (and the evaluator’s observation) that students were able to see and make more connections at the end of the course than at the beginning, and by the fact that at least some students were initially unable to distinguish between mathematics and science topics in their late interviews. The evaluation also showed that the course helped students to grow
in ways not necessarily related to the mathematics and science being studied: Some students grew socially, and some gained academic confidence that will serve them well in future studies. Finally, the evaluation helped show—and to some extent helped the teachers see—that the collaborative work involved in such a course can provide substantial professional development experiences.
CHAPTER 5

EVIDENCE OF PROBLEM SOLVING

This chapter will reexamine the data from the evaluation in light of the problem-solving process. Specifically, the researcher will combine various data points from multiple participants to identify both the problems that each group of participants (students, teachers, and evaluator) faced and the problem-solving heuristics that are represented by their solution processes. This process will demonstrate the value of problem solving, the basis for the Problem-Solving Approach, in this evaluation project.

The Students as Problem Solvers

The students faced several kinds of problems in this course. This section reviews the way the students addressed problems as daily mathematical tasks, problems as the focus for their final projects, the problem of repetition in the course itself, and the issue of the relevance of this course in their overall academic careers. The problem-solving strategy the students used is explained in each case.


Daily Problem-Solving Tasks

First and foremost, the students were tackling problems on a daily basis from the tasks assigned. Although the activity sheets, which were published materials from a major graphing calculator manufacturer, were very structured in the way that questions were posed, it seemed that the real problems came in trying to obtain data that actually looked like what could be expected. It was not always clear from the activity sheet itself what the data should look like, but the students developed the habit of comparing data between and among themselves once the teachers began requiring that each student collect data individually. Thus, the students came to know that because the activities were well designed with specific questions that there should be a common appearance to the data tables and/or graphs.

When students found differences in the appearance of their data, they would investigate by trying the activity again, then by changing certain aspects of the task, such as distance to the detector, length of a string, or angle of a ramp. Often one or two students (but not always the same ones) would end up repeating the task many times, even well beyond the time that others had finished. Some would end up exchanging equipment, deciding that their probes or detectors “just weren’t working right”—and sometimes they were correct.

The Balloon Drop

One such instance occurred with an activity in which students had to determine the velocity of a balloon as it dropped to the floor. One group simply could not get their data to “look right,” and they tried making numerous changes to replicate the data that others were getting. First, they attempted to
add a small weight to help steady the drop. Then they considered the possibility
of a draft or breeze and tried to accommodate by moving their equipment away
from possible air sources. They thought perhaps a body part (arm, finger, knee,
etc.) or other object might be obstructing the motion detector, so they all moved
themselves a good distance away. Finally, they used another group’s
equipment—and still did not get a good data set. Although unsuccessful, the
researcher was very impressed with their problem-solving skills.

Problem-Solving Strategies Used

Overall, the students developed skill at identifying differences and
changing parameters through a process of elimination. The problem of
differences in data that should appear similar, then, was often solved by a
combination of guessing and testing, by accounting for all possibilities of difference,
and by logical reasoning and working backward from interpretations of graphs and
data points.

The Final Projects

The second set of problems faced by the students in this course was that
of finding a topic and problem to work up as an activity for the class as their final
projects. The task was very open-ended, and the students were open to all
options in terms of science topics and underlying mathematics concepts. No
direction was given to any of the students—they were simply told that they all
knew how all the probes and detectors worked and that they had to come up
with something new.
As was seen from the teacher comments in interviews, the results were remarkable. Projects ranged from James' comparison of temperatures created by light on two colors of fabric, to Abby's work with the acceleration and velocity of a homemade hovercraft, to Tom's activity comparing the effect of string length and angle on the period of a double-string pendulum. The concepts were heavy-duty, and as in the case of James where he had to develop a function to match the situation, often very original—at least compared to the content of the activities they had used in class. The co-researcher who took notes during the focus group also observed the class for the remainder of that day, and provided the researcher with comments on the class in which she wrote:

Impressive labs they are creating; they don't realize the high level of concepts (which may mean they see the concepts in simpler form so they understand the concepts behind the procedures).

Problem-Solving Strategy Used

The process of developing the final projects began with a period of time during which the students were expected to simply experiment with the equipment, in the same way that they had with the pressure contest. Equipment and materials were made available for students to play with, and from that play came genuine problems. Although the task was presented in a very open-ended fashion, it was open ended only to the extent that the equipment allowed it to be, and the students had only certain types of equipment available to them through which to focus their problems. Thus, in terms of solving the problem of finding a problem, it would seem that the students worked backward from their knowledge of the probes and detectors to frame a problem that could at least be interpreted through the available equipment and its limitations.
Repetition

A third problem area for the students was one that resulted from the fact that the course was being evaluated, which meant they were being asked for feedback. Because the students so often mentioned that the class had become monotonous, with the same activity worksheets in the same format with even the same font and illustrations being used each and every day, they were asked in several different formats (written questionnaires, the focus group, and the individual interviews) for their suggestions as to how that aspect of the course might be improved. Many suggested that a different publisher be used so that at least the formatting and font would change. Others indicated that a different approach to the activities altogether, such as doing more of the contest-oriented tasks where they were free to explore with the equipment, would be the way to improve. One suggested the possibility of more projects like the final one, but admitted that the other work (the repetitious activities they had complained about) was probably necessary in order to pursue such projects.

Problem-Solving Strategy Used

All in all, it seemed that the students put thought into their responses when asked to improve upon the course—which is not a question often posed to high school students. Indeed, the fact that they were being asked to assist in curricular decision-making presented itself as a new problem. The students' problem-solving process included the fact that they found patterns by labeling exactly what it was that made the class seem so monotonous, and then listed numerous possibilities for alternatives.
Relevance of the Course

The fourth area identified as a problem approached by the students was that of trying to determine where such a course "fits into" the overall scheme of their academic plans, and indeed, their lives and future. As described earlier, several students (Lindsay, James, and Sarah) were not in this course by choice, having been placed in it either by default (scheduling problems) or by a parent's insistence. Others had chosen to take the course because of the teachers themselves, based on positive prior experiences with one or both of them. Still others chose to take the course because of their future career plans, knowing that an additional mathematics and/or science course could only help them progress toward their goals.

Although the issue of where this course falls in the overall scheme of life was not approached directly with the students, it was clear that by the time of the late interviews they were pigeonholing its purpose in their lives. Some admitted that they had not really expected much from it, and that they had been enticed by the no homework and no tests end of the deal that the teachers made with this pilot group. Although some of them still did not see the course contributing to their future in any way, others had had a change of heart. James, for instance, who had been placed in the course because of scheduling problems, said he was really glad he had taken it after all, and expected his experiences would help immensely with physics and precalculus. Abby saw her work in the course as beneficial in her decision to become a teacher—she realized that she has the ability to help people see and understand concepts (a talent confirmed by
the teachers in their comments about her). Tammy thought the course was good preparation for becoming a veterinarian, even though the coursework she would pursue in college would be more life science oriented.

**Problem-Solving Strategy Used**

The most common rationalization for the course among the students, however, was that the benefit of learning to use the calculators and equipment was by far the most important outcome, one that would prepare them for many future endeavors. Only one student, Elizabeth, specifically pointed out that she would carry the type of thinking she had learned into other parts of life, including college. Overall, most of the students were attempting to see how to integrate this experience with the others of their lives and make it purposeful, and they seemed to be doing so at least in part by taking a different point of view: from their imagined vantage point 5 to 10 years down the road.

**The Teachers as Problem Solvers**

The teachers faced many problems throughout the course as well as during the planning and design stages. Many of the teachers' problems were of a very practical nature, and to some extent were solved by the students themselves. These lower level problems will be discussed first. They will be followed by those problems that were much larger in scope and difficulty, and basic to the nature of teaching itself—perhaps particularly in terms of teaching mathematics and science.
Repetition

The practical problems were largely related directly to the reactions of students to the course and/or the interactions of students and the group dynamics that are part of teaching. One problem was the student perception that the course was repetitious in nature, and that it focused more heavily on mathematics than on science. The idea of monotony came directly from student feedback, and indeed the solutions the teachers offered to this specific problem grew directly out of student comments, as well. However, in this case, it is perhaps the first of Polya’s problem-solving steps, understanding the problem, that is most important. The teachers saw the problem of monotony as being related to the perception that the class time was heavily oriented toward mathematics. Their understanding of the situation was that the students were interpreting the content of the course according to what happened; that is, the students saw science as being the act of collecting data by carrying out the instructions provided for the activity, while mathematics was the manipulation of data thereafter. Therefore, the students saw the worksheets as always being a little science (one page of instructions for data collection) followed by a lot of mathematics (two or three pages of questions), and always in precisely that order, with similar questions occurring each time.
Problem-Solving Strategy Used

Looking back on the course during the late interview, the teachers came to see this as a problem of presentation—even of diplomacy, in terms of giving equal billing to the mathematics and science involved in each activity. Their approach to solving the problem, then, involved adopting a different point of view: that of the student(s).

Student Grouping Procedures

The second very practical problem the teachers had to address was that of making all students accountable for the work involved in the course. As seen earlier in the late teacher interview, it was apparent that the changes in design of assignment, work submission, and grouping were primarily aimed at one student, Ed. Although the teachers worked hard at making it seem that the changes were aimed at at least a certain group of students (and succeeded in fooling even the researcher, an experienced high school teacher himself) and felt that it was unfair both to them and the other students, they admit that Ed might very well have done everyone a favor by creating a need for these changes.

One way in which the changes benefited the class was the fact that individual data collection, instituted because Ed had a tendency to merely copy other data reports and/or simply share data via the linking capacity of the calculators, increased the number of data sets in the room from which to make comparisons and generalizations. The researcher observed the students on many occasions applying their problem-solving skills by comparing their data
sets and trying to troubleshoot the equipment and/or the procedure and materials. Thus, a behaviorally-oriented decision made by the teachers had a definite pedagogical benefit.

Another positive change was the way in which groups were formulated. Although Ed was the instigator, he did seem to have greater influence on some students than others, and he and those students tended to congregate together whenever allowed to choose their own partners. When the teachers began assigning groups randomly to avoid having that common “problem group” to form each time, they effectively increased the exposure the students had to many different ways of thinking and problem solving, and increased the amount of social contact. As seen with students like James and Lindsay, this social exposure was at least as important to their development at this point in their lives as was any mathematics and/or science concept.

Problem-Solving Strategy Used

To solve these problems, then, the teachers had to first understand that it was indeed rooted in one person, and that in order to deal with that one person they had to create reasonable but fair consequences. Indeed, the teachers were forced to consider a single extreme case (Ed) to develop their plan for a solution.

Helping Promising Students Excel

One successful outcome of the pilot course was a solution, at least in some cases, to the age-old, often unspoken, problem that teachers face: how to get those students who are intelligent and promising to excel. Evidence of success is seen in at least four students whose experiences have already been recounted in some detail. Abby, who had never been an excellent student in science or
mathematics, stood out as a problem solver in this class, perhaps primarily because the two areas were brought together in a context that she could understand better. Tom, who entered the class a bit behind the others in terms of academic preparation, held his own when it came to the academic challenges put before the class, even though most were conceptually beyond his level. Lindsay, who began high school as a virtual recluse, blossomed socially in this class, perhaps because of the structured activities that required extensive interaction in groups with constantly changing members. Finally, there was James, who worked so hard to maintain a cool image, but ended up with perhaps the most sophisticated mathematical challenge in the class with his final project.

Problem-Solving Strategy Used

Although they never stated this as one of the goals they were working toward in this pilot course, the teachers were creating an environment for new successes through the integration of the two subject areas. Whether for academic purposes or social effects, the course was a success for Abby, Tom, Lindsay, and James. In retrospect, it seems that the very creation of this course served as a solution to that problem of getting those students who are in the shadows to shine. The teachers’ attention to the problem of integration was perhaps a way of looking at a simpler problem to solve that more difficult problem of helping promising students excel.
Establishing Relationships Between Mathematics and Science

Altogether, however, this was a side benefit to the teachers’ work on the larger problem of getting students to see and understand the relationships between mathematics and science that presented themselves so frequently in the separate courses that they had been teaching for years. They had been witnessing generations of students who compartmentalized their knowledge into the two distinct subject areas, and expressed constant amazement by the students’ inability to use even simple concepts of slope, for example, in their science courses.

Problem-Solving Strategy Used

After several years, the teachers decided through their discussions of the problem that one way to address the mismatch of subjects would be to teach them together in a single, integrated course. Their solution represents perhaps two problem-solving approaches. First, they were reasoning logically by thinking that if concepts remain as disjoint in the students’ minds as the courses are in the students’ daily schedule, then the concepts would be joined in their minds in an integrated course. Also, their attempt to think like the students in order to determine why the two subjects are so distinct for them was an example of adopting a different viewpoint.

Creating the Course

Creating the integrated course was not an easy task. There were several problems the teachers had to address, including acquiring appropriate equipment and calculators, convincing the school’s administration to allow an unorthodox team-teaching situation, and scheduling issues (both for teachers
and students). As discussed earlier, the teachers applied for a grant to cover the
cost of calculators and for several professional development sessions they
attended prior to the year the course was offered. Convincing the principal and
district to allow the team-teaching situation was not that difficult, since both
teachers were experienced and trusted. Also, the principal knew that the two
worked well together from previous collaborations that were not necessarily
content-related, and knew that their personalities complemented each other
quite well. The real problem here, though, was scheduling.

Although there was a logical time to offer the course from the students’
perspective, the school did not have enough staff resources to allow both a
common teaching block and a common planning time for both teachers.
Although this was a stumbling block and a point of contention, the teachers
decided that they could at least take advantage of the only time they had
together: a 30-minute lunch. Mrs. S took the point of view, as she stated in the
interviews, that even that small amount of time was more than they had done
before, and therefore it was worth it. Mrs. M seemed more upset about the lack
of planning time than did Mrs. S, and talked at length in the interviews and in
casual conversations with the researcher on observation visits about the need for
both a common planning time during the trimester that the course was offered
and common planning time in the trimester preceding the course in order to
gather materials and activities, and simply to be more organized. This concern
was perhaps greater for her because the course was offered as a block only by
means of scheduling one teacher for the course for one period with the other on
one-hour plan, and vice-versa for the other period. Although the original intent
was for both teachers to be there for the entire two-period block, Mrs. M ended up needing at least most, if not all, of the first class hour for her planning time, while Mrs. S had a less demanding daily schedule and was in fact able to be with the students at all times for the full two-hour block. Mrs. M often felt left out on activities, had to be “caught up” every time she entered the classroom, and often demonstrated misunderstandings of what the students had actually gained or learned in a given situation—a fact that was mentioned by both teachers in the late interview.

**Problem-Solving Strategy Used**

Overall, the issue of scheduling, particularly scheduling common planning time, was the primary problem faced in creation of the course. The teachers’ solution is perhaps an instance of working from an extreme case to solve a problem: the lack of a common planning period forced them to look for a way to take advantage of any, even the smallest, amount of time together (their common lunch period) to develop the course.

**Professional Development**

An overarching consideration in the creation of this pilot course was perhaps a problem these teachers had, and that many teachers experience: finding a way to obtain very targeted professional development so that the specific classroom issue these teachers were tackling would be addressed. In a sense, these teachers had informally been conducting their own form of action research through the years by noticing their students’ inability to make connections between their two fields. Although they sought professional development workshops with the use of the calculators and data collection
equipment, these experiences were not satisfactory for addressing the problem that their students were having, which was overcompartmentalization of the topics in subject areas.

Problem-Solving Strategy Used

The teachers had larger concerns with their own growth and development, as well. Mrs. M knew of Mrs. S’s talent in making learning hands-on and activity based. Mrs. S knew that she needed a refresher on some of the mathematics behind the science she was teaching, and needed more opportunity to determine exactly what it was that prevented students from making connections that seemed so obvious. It seems that the teachers came up with a solution to all these problems through one method: solving a simpler problem. While professional development is usually thought of as a task that requires time beyond the school day in course work that is outside of the school building itself, they looked to the students’ main problem, compartmentalization of subjects, and incorporated all these other aspects of professional development into their solution to that one issue: implementing an integrated course. Mrs. M had the opportunity to learn some of Mrs. S’s hands-on techniques, while Mrs. S had the reminders she needed of mathematics she had seen and studied long ago. Both witnessed the students as they began to make the desired connections, and learned that perhaps the key element, as Mrs. S stated in the late interview, is time.

In summary, the teachers were taking advantage of their years of teaching experience and were solving problems left and right, often without even thinking about the fact that they were doing it or how they were doing it.
Despite the researcher's requests for them to consider and even document problem-solving approaches used as the course progressed, the teachers did not follow through on this task. Perhaps now it can be seen that the problem-solving process had become so natural to them that they found it difficult to ferret out reasons and approaches for each solution—perhaps as difficult as separating the very mathematics and science concepts they wanted their students to link together.

The Evaluator as Problem Solver

The evaluator dealt with problems on at least two levels. First, as evaluator, there were problems related to the ongoing data collection and analysis. Levels of engagement and issues of role with respect to the various participants were among these issues. Second, as a researcher examining the program evaluation process, there were problems in terms of the overall conduct of the evaluation: The intent was for problem solving to be a deliberate and open process that was part of all the participants' approach to the course, but the reality was that problem solving itself was rarely discussed openly during the trimester. Although the researcher attempted to ask directly about problems and strategies to solve them during the early visits, this practice was eventually abandoned when it became clear that such prompts did not encourage open discussion of such issues.
Varying Roles of the Researcher

In the various interactions with the teachers and students during the course, the researcher found himself playing a range of roles. This was due mainly to the participant-observer mode of observation that was employed throughout the evaluation. This type of observation required the researcher not to sit in the back of the classroom and take notes, but rather to engage in activities with the students and teachers just as any participant would do. The levels of engagement with each group and with individuals became a problem that the researcher solved using several strategies.

With Students

The researcher's problem of the level or degree of engagement occurred on a frequent basis while working with the students. During the times that questionnaires were being administered, in interviews, and while the focus group was being conducted, it was clear to all those involved that the researcher was present in the role of evaluator. However, on days when the researcher visited purely for the purpose of observations, he was sometimes to the side of the classroom taking notes and/or engaging in discussion with the teachers, but was sometimes directly involved with the students. When with the students, he was alternately acting in the role of evaluator simply by continuing to observe and write notes later, as a teacher by fielding questions and making judgments as to how much of an answer (if any) should be given, and as a fellow student and group member by helping to hold or position a detector or troubleshoot a calculator. Although the teachers believed, according to indications in their interviews, that engagement with the groups was part of what made the
students so comfortable and open with the feedback they provided, the researcher often felt himself having to resist the urge to simply be a student. After all, such an experience in integration was almost as new to the researcher as to the students—he was learning, too. Further, the researcher's years of experience as a teacher created a natural tendency to become a teacher around almost any teenager. He was at times just as likely as the teachers to answer a question posed by a student as he was to reprimand them for not being on task.

Problem-solving strategy used. In general, the researcher found himself on a very tenuous line between evaluator, teacher, and student in his relationship with the students during observation visits. In order to deal with this fluid relationship, the researcher was in fact constantly using the problem-solving strategy of adopting a different point of view—that of the students or of yet another outside observer—to weave in and out of the various roles.

With Teachers: Researcher as Superior

The researcher's role with the teachers was similarly fluid. At times, but more toward the beginning of the project, the teachers seemed to treat the researcher as a superior. The researcher felt repeatedly that the teachers were seeking his approval. Mrs. S's statement in the early interview concerning the fact that they did not have "critical friends" nearby may have been part of the reason for such actions. Because they seemed to have only each other to sound off ideas and obtain professional feedback on the decisions that teachers face daily, a new face with new opinions and new perspectives was more than
welcomed. Furthermore, the new face was that of a fellow teacher who had pursued degrees and coursework beyond what they had, thus lending a sense of expertise and experience that went beyond their own.

Problem-solving strategy used. The solution to this problem rested on the fact that the researcher noticed a pattern. The researcher realized that he had experienced this type of treatment from teachers during his involvement with several other research projects that took him into public school classrooms. To avoid perpetuating the superior role of an expert, the researcher began referring more frequently to his own experiences as a classroom teacher, comparing certain classroom situations to those of his own past, and empathizing with the teachers in their dealings with students, parents, and administrators.

With Teachers: Researcher as Fledgling Educator

At the same time that the teachers treated the researcher as an academic superior, they seemed also to act as mentors. They were constantly concerned about what more I might "need" from them—in terms of data, background information, copies of materials, more time with students for interviews, etc. Worry about what my advisor would think of my work was also expressed repeatedly, and the teachers offered to adjust whatever they needed to in order to help me. On several visits, one or both of them made comments along the lines of "We just want to do what we can to help you get this PhD."

Problem-solving strategy used. Although this aspect of the researcher's relationship with the teachers was quite different from that of being viewed as the expert, the researcher realized that it was nonetheless part of a similar pattern: In situations with older peers in the education profession, he often finds
himself in a mentor-student relationship. These teachers had certainly been in education for many more years than the researcher, and they apparently saw themselves as responsible for helping a younger colleague pursue his career. Further, just as the researcher's natural tendency to become a teacher around teenagers made him likely to assume that role in this classroom, the teachers were likely displaying similar tendencies with a younger professional. The solution to the problem of the teachers treating the researcher as an expert served here, as well: Subtle but frequent reminders of past teaching experience and empathy for their situation seemed to help decrease the offers for mentor-like assistance, and the teachers came to treat the researcher more as a colleague.

*Explicit Use of the Problem-Solving Metaphor*

Despite the teachers' stated willingness to help the researcher in whatever way possible, it has already been mentioned that repeated requests to explicitly employ problem-solving strategies in a forward-looking manner seemed to be misunderstood. The teachers did not openly discuss problems and their solutions with the researcher as the course went along, but rather referred to them indirectly and primarily in the final interview. Nonetheless, the researcher felt that the problem-solving metaphor would still prove valuable in some way.

*Problem-Solving Strategy Used*

Indeed, it seems that the solution to the problem of the lack of direct use of the problem-solving process was to simply wait until the end of the project, when it became evident that the problem was actually a simpler one. The researcher's work to fully analyze the data helped him see that in fact the
problem was not so much how to adopt a problem-solving stance in approaching the evaluation, but rather how to apply problem solving as a tool for analyzing data and as a framework for the evaluation itself.

The basic data analysis, as reported in Chapter 4, revealed answers to some essential questions relevant to the evaluation. It is clear, for instance, that students progressed in their abilities to make mathematics and science connections, and that the teachers learned some basic principles for helping students make those connections. But the application of problem solving, the second part of this data report, revealed underlying issues such as the teachers' consideration of the single problem student as an extreme case and their conceptualization of the integrated course as a simpler solution to the problem of their professional development needs. These underlying issues became important both to the evaluation itself and to the researcher's study of the process of program evaluation.

Summary of Problem-Solving Strategy Use

The use of problem-solving strategies during this program evaluation was different from what the researcher had originally intended. Although original expectations were that the teachers in particular, perhaps with input from or collaboration with the researcher, would explicitly apply various strategies as they faced various problems during the course of the trimester, the strategies were instead used indirectly. That is, rather than direct discussion with the teachers of the use of particular problem-solving strategies as applied to specific situations, the researcher began by looking at the obvious problem-solving
activities undertaken by the students. Upon seeing the various strategies employed at the student level, the researcher then looked at the problems the teachers faced, and was able to identify clear uses of the strategies in a number of situations. Finally, the researcher reexamined his own actions, defined problems he faced, and identified problem-solving strategies that were utilized. In each case, the problem-solving strategies were being applied without explicit discussion or even acknowledgement of that fact. A summary of the various problem-solving strategies employed during the project appears in Table 5.1. The strategies are displayed along with the participants who used them and the problems that were solved. Two strategies, making a drawing (visual representation) and organizing data, were not directly identified in the data analysis, and therefore appear in the table with blanks.
<table>
<thead>
<tr>
<th>Problem-Solving Strategy Employed</th>
<th>Problem Solver(s)</th>
<th>Problem(s) Solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working backwards</td>
<td>Students</td>
<td>Various daily lab activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Developing problems to use for final projects based on equipment capabilities and limitations</td>
</tr>
<tr>
<td>Finding a pattern</td>
<td>Students</td>
<td>Determining how to improve the course by noticing a pattern to the daily activities that they found boring</td>
</tr>
<tr>
<td></td>
<td>Researcher</td>
<td>Helping the teachers to see him as a colleague rather than an expert or student by relating this situation to other such relationships with teachers in prior research projects</td>
</tr>
<tr>
<td>Adopting a different point of view</td>
<td>Students</td>
<td>Determining where this course fit into the overall scheme of their lives by imagining themselves 5-10 years in the future</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>Determining ways to improve the course by taking the view of the students</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Determining why the students tend to separate mathematics and science by thinking of the separation of the two subjects in the students' daily schedules</td>
</tr>
<tr>
<td></td>
<td>Researcher</td>
<td>Determining when to assume the various roles he played, including student, teacher, researcher, colleague, and evaluation expert</td>
</tr>
<tr>
<td>Solving a simpler (analogous) problem</td>
<td>Teachers</td>
<td>Finding a way to further their own professional development by solving a problem for the students: how to integrate the two subjects</td>
</tr>
<tr>
<td></td>
<td>Researcher</td>
<td>Finding a way to use the problem-solving metaphor when it was not being directly applied by the participants. The simpler problem was how to use problem solving as a data analysis tool.</td>
</tr>
</tbody>
</table>

Table 5.1: Problem-Solving Strategies Employed in Pilot Course Evaluation
Table 5.1 continued

<table>
<thead>
<tr>
<th>Problem-Solving Strategy Employed</th>
<th>Problem Solver(s)</th>
<th>Problem(s) Solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Considering extreme cases</td>
<td>Teachers</td>
<td>Developing ways to get all students involved by focusing on one student (Ed)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Finding a way to integrate the two subjects by taking advantage of even the smallest of joint planning times: lunch</td>
</tr>
<tr>
<td>Making a drawing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(visual representation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligent guessing</td>
<td>Students</td>
<td>Various daily lab activities</td>
</tr>
<tr>
<td>and testing (including approximating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounting for all possibilities</td>
<td>Students</td>
<td>Various daily lab activities</td>
</tr>
<tr>
<td>(exhaustive listing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizing data</td>
<td>Students</td>
<td>Various daily lab activities</td>
</tr>
<tr>
<td>Logical reasoning</td>
<td>Teachers</td>
<td>Finding a way to get students to link mathematics and science by determining that if the areas are linked in the daily schedule they will be linked in their minds</td>
</tr>
</tbody>
</table>

Table 5.1: Problem-Solving Strategies Employed in Pilot Course Evaluation
This retrospective use of the problem-solving process—to understand the problems faced, reexamine the apparent plan for a solution, review the solution employed, and verify that the problem was solved—led to more in-depth understandings in terms of the course itself. It also demonstrates the use of the problem-solving metaphor as a data analysis tool. Indeed, the problem-solving process is applicable at several levels of the program evaluation process, and these will be examined in Chapter 6.
CHAPTER 6

CONCLUSIONS

This chapter is a discussion of the applicability of the problem-solving metaphor to the program evaluation process, and presents the last phase of the researcher’s development of the Problem-Solving Approach to program evaluation. The practicality of the use of problem solving in the evaluation of the Integrated Laboratory Course will be reviewed, and potential uses of problem solving will be discussed in terms of the proposed Problem-Solving Approach. Finally, the research questions will be addressed and implications for future research and for the field of mathematics education will be discussed.

Problem Solving as a Data Analysis Tool

In the case of the evaluation of the Integrated Laboratory Course, which was intended as a pilot for the Problem-Solving Approach to program evaluation, the direct application of problem solving as an evaluation technique occurred only through its use as a data analysis tool. It has been explained that the researcher intended for the teachers of the pilot course to identify problems as they arose in the course, to sift through the various problem-solving strategies (perhaps in consultation with the researcher) and choose one (or
more) appropriate to each situation, then apply the strategy, reflect on the situation, and proceed to the next problem. Although this did not happen, it was shown in Chapter 5 that the teachers, the students, and the researcher were all solving problems as the evaluation proceeded, and that in fact, specific problem-solving strategies were in use even though they were not explicitly stated. For the reader’s convenience, some specific examples for each participant group (students, teachers, evaluator) will be reviewed here.

The Students as Problem Solvers

The students demonstrated use of problem-solving strategies on several levels during the course. First, they were solving problems in their daily tasks when they encountered difficulties with the calculators and/or data collection equipment. In particular, the students used intelligent guessing and testing to troubleshoot equipment difficulties, accounting for all possibilities when trying to determine reasons for data discrepancies, and working backwards from the equipment and its capabilities to develop problems appropriate for their final projects. At the program level, the students were able to recommend changes in future offerings of the course because they recognized a pattern—they found the daily activities to be repetitious and boring. On still another level, the students adopted a different point of view to determine what place this course had and what purpose it would serve in their academic careers and their lives. Overall, the students showed skill in applying the problem-solving strategies in academic situations and beyond.
The Teachers as Problem Solvers

The teachers demonstrated use of several problem-solving strategies, as well. On a very practical level, they considered an extreme case (Ed, the student who would take advantage of other students doing work for him when in a group) to design grouping procedures and data reporting requirements that forced the students to maintain some independence. In this case, their attempts to prevent the one student from avoiding work made all the students more productive and self-reliant. The teachers adopted a different point of view—that of their students—in order to determine ways to improve the course. They noted, for instance, the students' boredom with the daily activities, and the students' excitement when they were allowed to explore the equipment independent of a prescribed activity with detailed instructions. Finally, the teachers solved a simpler problem, finding a way to work together for the sake of their own professional development, as a means of addressing a larger problem, that of helping their students to integrate ideas in science and mathematics. Here again, problem solving was clearly taking place, the strategies targeted for use in program evaluation through the approach proposed in this project were being applied, and program improvement was taking place.

The Evaluator as Problem Solver

The evaluator also employed problem-solving strategies through the course of the evaluation, although certainly not in the way originally intended. He noticed a pattern in this situation compared to others he had been in with practicing teachers, and used his experience to find ways to help the teachers see him as a colleague rather than an expert (due to his advanced degrees) or as a
student (due to his age). Also, the evaluator was constantly adopting different points of view to determine which role to play in various situations—with students, for example, he was sometimes a fellow group member, sometimes another teacher in the room, and sometimes an "outsider" doing research on their experience, and his cognizance of these roles helped him determine when it was appropriate to change character. Here again, data analysis revealed that problem solving was an ongoing task even for the evaluator, despite the fact that it was not noticed while it was actually occurring.

This indirect use of problem-solving strategies, both in the actual implementation of the program and in the data analysis phase of the evaluation process, argue the case for their direct use. That is, since problem solving and the use of specific problem-solving strategies are evident and already in use in such a program, the possibility of expanding and highlighting the use of a problem-solving orientation would seem productive, particularly in light of the proposed Problem-Solving Approach.

Potential Use of Problem Solving in Program Evaluation

The evaluation of the Integrated Laboratory Mathematics-Science Course was a pilot for the Problem-Solving Approach to program evaluation, which, as proposed in Chapter 2, is an application of problem solving as a metaphor for evaluation. The problems that arose during the implementation of the Integrated Laboratory Course and its evaluation, along with the problem-solving strategies that were employed, serve as examples of the potential uses of problem solving in the evaluation process.
Adopting a Different Point of View

In Chapter 2, potential applications for each of 10 problem-solving strategies was outlined in Table 2.3, and several of those potential applications arose in the process of this pilot use of the proposed approach (see Table 5.1). For example, it was projected that the strategy of adopting a different point of view might be useful if we take the viewpoint of a student, asking “What do I want from this program?” In this case, it was the teachers who used this strategy, by noting that the students were very excited when allowed to freely explore the equipment, and the teachers used that information to help formulate the idea for the students’ final project.

Considering Extreme Cases

Another example is the strategy of considering extreme cases. It was projected in Chapter 2 that this strategy might be used by asking, for example, “What happens to high-achievers as they progress through this program? What happens to low-achievers as they progress? How/why is the experience different?” In this case, it was again the teachers who used this problem-solving strategy, and the problem was one of motivation rather than achievement. Their consideration of the extreme case of Ed, who was all too willing to let others do his work for him, forced the teachers to apply stricter standards for data collection and individual accountability across the board. The result was more data sets, which allowed for further comparison and more opportunities for trouble-shooting speculation when the need arose. It also enhanced the student experience by providing a greater range of data, which in turn increased the need for data interpretation.
Although these problems, their solution strategies, and the problem-solving process itself were not directly addressed as the problems were occurring, the fact that situations arose in which the problem-solving strategies were employed in ways similar to those projected shows that there is promise for the application of problem solving in program evaluation. Indeed, it was shown in Chapter 5 and emphasized here that all participants (students, teachers, and evaluator) were encountering and solving problems at several different levels of the program. Students were solving problems as part of their daily assignments; the teachers were solving problems at a curricular level, with intent for change and improvement both within the trimester and for the next offering of the course; and the evaluator was solving problems having to do with his role as evaluator and the evaluation process itself. This shows at least three levels of application of the problem-solving process to program evaluation, which implies the need to refine the Problem-Solving Approach.

The Problem-Solving Approach

As described in Chapter 2, the Problem-Solving Approach assumed the application of the problem-solving process to various problems that would arise during the implementation of a program. As a result of the pilot experience, the approach has been refined to incorporate the possibility of levels of application, so that problem solving becomes, in effect, a recursive process that is employed repeatedly and cyclically throughout the evaluation. Figure 6.1 depicts the revised approach.
Figure 6.1: The Problem-Solving Approach to Program Evaluation
Problem Solving at the Evaluation Level

The first level of problem solving in the Problem-Solving Approach is at the evaluation level. At this level, the primary problem is that of designing and conducting an evaluation that incorporates data collection processes that are appropriate to the program itself. This design process parallels that of the overall evaluation process, and involves understanding the program, designing a data collection strategy, carrying out the data collection activities, and looking back to review the effectiveness of the activities and the appropriateness of the outcome. Although an evaluation always begins with a plan for data collection activities, it is understood, partly through the influence of evaluation approaches such as Stake's (1975) responsive approach and Guba and Lincoln's (1989) fourth-generation approach, that data needs often change. Ongoing data analysis may reveal the need for additional or different data, or it may eliminate the need for certain data that was originally thought necessary.

The fact that data needs may change is accounted for through the problem-solving process because the final step is to review the effectiveness of a strategy and the appropriateness of a solution. In a problem-solving program evaluation, this means examining data continually to identify and illuminate problems at the program level, and to determine ways that additional or different types of data might be collected in order to further understand a particular problem at the evaluation level. The various problem-solving strategies (see Table 2.3) would prove helpful to the evaluator in determining ways to modify the data collection tasks. It may be, for example, that to determine a specific question or set of questions for a focus group, the evaluator
must adopt the point of view of a specific group of participants. It is also important to note that the evaluator must design data collection activities in such a way that program-level problems will surface. These program-level problems shape and become part of the data collection process. This relationship is depicted in Figure 6.1: A problem at the program level branches off from the third step, carry out the plan, in the problem-solving process at the evaluation level. Since the carry out the plan step is where data collection is occurring, that is the point at which a problem at the program level would be revealed.

Problem Solving at the Program Level

During the course of a program's implementation, problems develop and issues arise that program administrators (such as teachers) must address. Such problems should be identified and illuminated by the data collection procedures, and direct attention to the various problem-solving strategies would in many cases streamline the solution process. In the case of the Integrated Laboratory Course evaluation, the teachers involved were experienced problem solvers, and although the evaluator hoped for explicit discussion of the problem-solving process and solution strategies employed, it appears that the teachers did not need to address the solution process so directly. This may be in part due to the fact that they were experienced mathematics and science teachers, respectively, and problem solving is part of what they do and teach on a daily basis.

It is this familiarity with problem solving that the Problem-Solving Approach uses as an advantage. In the case of less experienced mathematics teachers, then, the explicit consideration of various problem-solving strategies could ease and even streamline the solution process, because the teachers would
at least be familiar, through their studies in mathematics, with the process of problem solving and the heuristics available for use. Further, the streamlining of the solution process would allow for program improvement while the program is being implemented. This streamlining of the improvement process would accomplish one of the stated goals of a problem-solving orientation, which is to direct the goal of evaluation toward genuine program improvement rather than a declaration of the success or failure of a program. The constant attention to problems at the program level might have implications for revisions in the data collection process. Thus in Figure 6.1, the problem-solving process at the program level reconnects with the process at the evaluation level at the carry out the plan, or data collection, step.

With a focus on program improvement, it is necessary to consider the daily tasks involved in a mathematics program. Therefore, in Figure 6.1, another level of the model branches off from the third step, carry out the plan, of the problem-solving process at the program level. Since the carry out the plan step of the process at the program level is the actual implementation of the program or course, it is at this level that the daily problems would appear.

*Problem Solving at the Mathematical Task Level*

Since the focus of the current project was to develop a mathematical approach to program evaluation so that mathematics programs could be more effectively and productively evaluated, the approach was investigated using a program that involved mathematics. With mathematics as the content area, problem solving was included in a daily task/assignment fashion.

151
In the Integrated Laboratory Course, the daily mathematical tasks involved collecting and analyzing data mathematically through the context of science. The researcher's analysis of the students' dealings with these daily problem-solving tasks aided in the realization that the same type of problem solving was occurring at the program and evaluation levels. Specifically, the fact that the students were recognizing a pattern in their daily activities and were able to suggest ways that the activities could be improved helped the researcher realize that the teachers were having to adopt the students' point of view on this and other problems that arose. This connection led the researcher to explore the various problems involved in all levels of the program from a problem-solving perspective, and in turn to the conclusions herein.

As a result of the researcher's experience with the Integrated Laboratory Course, it is suggested that attention be given to the daily problem-solving activity of the students involved in the mathematics program being evaluated. There are several reasons that such attention would prove productive. First, the evaluator needs to be in touch with the daily events of the classroom, in part to obtain better information about students' progress with the mathematics content, but also to identify potential problems at the programmatic level. Difficulty with problem solving at the mathematical task level may point to curricular issues (for example, choice of activities, order of topics studied) that become problems for the program administrators to solve.
Second, because the problem solving that occurs at the mathematical task level is more explicit, it provides the evaluator with more frequent and direct practice in the identification and use of the problem-solving strategies. In turn, this helps the evaluator maintain a problem-solving mindset, and may make some other problems or the strategies to solve them more apparent.

Finally, because the reason for developing the Problem-Solving Approach was to more closely link the content material and the evaluation process, it is suggested that attention to problems, their solutions, and their solution processes at the mathematical task level might provide insight for problems at the program and evaluation levels. Although this did not occur in the evaluation of the Integrated Laboratory Course, it would seem possible that the type of problems experienced at the mathematical task level would parallel or perhaps otherwise structurally resemble those at the other levels. Attention to the daily activities of a program should naturally inform the evaluation process in a number of ways.

Because of the potential influence of mathematical task level problems, Figure 6.1 depicts the problem-solving process at that level influencing the carry out the plan, or program implementation, step of the problem-solving process at the program level. Since problem solving at the program level in turn influences the process at the evaluation level, the model in Figure 6.1 illustrates a recursive use of the problem-solving process at all levels of an evaluation utilizing the Problem-Solving Approach.

153
In the discussion of the development of the Problem-Solving Approach in Chapter 2, three approaches to evaluation were mentioned as having helped to shape the proposed approach. The degree of influence of these three approaches on the pilot application of the Problem-Solving Approach to program evaluation will be presented here.

Stake's Responsive Approach

Stake's (1975) responsive approach was invoked through the focus on program improvement, the involvement of and attention to all stakeholders, and the allowance for change in the overall data collection plan. The degree to which the pilot application of the Problem-Solving Approach was faithful to the ideas of responsive evaluation can be determined in part by the fact that the teachers did work to improve the course as it was being implemented—examples include their change in the individual accountability required for students in terms of data collection, and their design for the final project, which was based on student interest generated by free exploration of data collection equipment. Further, the data collection plan included means of collecting information from all stakeholders at several different points in time, as well as more frequent but less formal contact through the evaluator's regular observations. This aspect of the design allowed the evaluator to inform the teachers of evolving student perspectives along the way, and allowed free flow of information between the two groups at the focus group. Finally, although few changes to the data collection schedule were made, it would have been easily possible had the need arisen to add another data collection event with one or
more stakeholders at almost any necessary point. Although the participants were very cooperative and would have been willing to participate in additional data events, it was decided instead to allow ongoing data analysis from prior data collection events to inform those that followed. Together, these features enhance the responsive nature of the pilot application of the Problem-Solving Approach.

_Eisner's Connoisseurship Approach_

The connoisseurship approach proposed by Eisner (1976) was described as influencing the Problem-Solving Approach through its emphasis on content expertise for the evaluator and the use of processes familiar to the evaluator. The faithfulness of the pilot application of the proposed approach to this idea is clear: The evaluator was a professional in mathematics education who was applying a mathematical process (problem solving) to the evaluation task. Further, the mathematical process of problem solving mirrored the problem solving that was occurring at the curricular center of the course through the students' daily tasks. The use of a content specialist as evaluator and a process that is inherent in the content itself together are in keeping with the connoisseurship approach to evaluation.

_Guba and Lincoln's Fourth Generation Approach_

The fourth generation approach proposed by Guba and Lincoln (1989) was described as influencing the proposed approach because of the focus on the role of the evaluator as one who facilitates social change. In the case of this application of the Problem-Solving Approach, the evaluator and the evaluation itself did seem to promote change both in the program and amongst the
participants. This change occurred primarily through the focus group, which was an event that allowed the students to know how much the teachers thought they had progressed, and provided the students a rare opportunity to offer constructive feedback.

On the whole, an occasion where the authoritative role of the teacher is erased as it was in the focus group is rare, particularly at the high school level, and the interchanges that took place during the session allowed both groups of participants to grow. The focus group was indeed a pivotal point in the evaluation of the course, and both teachers and students referred to it almost daily throughout the remainder of the term.

Although a single data event is not necessarily sufficient for orienting an evaluation toward promoting social change, the focus group was the aspect of this application of the Problem-Solving Approach that was most important in that respect. It should be noted that although the focus group questions and topics of discussion did not directly focus on problem solving, it is the case that a focus group is a data event that allows for problems to surface and perhaps work toward resolution through interaction. Through the data collection design, then, this specific evaluation facilitated some degree of social change through its orientation toward problem solving.

*Merit of the Problem-Solving Approach*

The Problem-Solving Approach was thus to some degree influenced by each of three distinct predecessors: the responsive approach, the connoisseurship approach, and the fourth-generation approach. It was mentioned in Chapter 2 that Stufflebeam (2001), using a rating scale for
adherence to the *Program Evaluation Standards* (JCSEE, 1994), performed a metaevaluation on the 22 distinct approaches that he has classified. He used this metaevaluation to judge the approaches, and to determine those most promising for use in the future. Of the nine approaches in his *most promising* list are descendants of the responsive approach and the fourth-generation approach. Although he does not include the connoisseurship approach on the most promising list (he considers the content area expert subject to bias), it is clear that at least two of the three major influences on the approach proposed herein are favored by one of the foremost evaluation experts.

**Addressing the Research Questions**

In this section, the research questions listed in Chapter 1 will be addressed (in the order they were presented) in light of the experience of its application to the Integrated Laboratory Course.

1. *Program Evaluation From a Mathematical Perspective*

The Problem-Solving Approach to program evaluation employs an essential mathematical skill, problem solving. Although the use of problem solving in the pilot application of the approach on the Integrated Laboratory Course was different from what the researcher had planned, a problem-solving orientation nonetheless proved useful as a data tool. Further, the problem-solving orientation helped to reveal underlying problem-solving processes that were guiding and influencing the development of the program all along, thus suggesting that direct application of problem solving (as originally intended) holds promise for the evaluation of mathematics programs.
2. Roles of Participants and Evaluator in the Problem-Solving Approach

The Problem-Solving Approach to evaluation is intended to acknowledge and bring to the forefront the fact that all participants in a mathematics program are problem solvers. In the case of the pilot application of the approach, the evaluator found himself assuming a number of varying roles—student, teacher, colleague, expert—which helped him in his ability, for example, to adopt the point of view of various participants. This fact alone enhances the evaluation process simply because it allows for a broader perspective on the program.

Additionally, however, it is predicted that the direct use of the problem-solving process and direct reference to problem-solving strategies would assist those who are relative novices at program implementation to orient themselves directly to program improvement through program evaluation—making improvement a process that works simultaneously with implementation.

Furthermore, in this pilot application of the approach, the students were able to offer suggestions for improvement based on their own recognition of problems in the course, and their ability to do so put them in a position that reduced the hierarchy normally present in the teacher-student relationship. Indeed, the students became program-level problem solvers alongside their teachers.

3. Effect of the Problem-Solving Approach on Evaluation Conclusions

To understand the effect that the problem-solving orientation of the Problem-Solving Approach had on the pilot application, it is best to compare the results discussed in Chapter 4 (the evaluation report) with those presented in

158
light of the approach in Chapter 5. While the conclusions reached in Chapter 4 are certainly valid and valuable, they are based on traditional views of the evaluation process and fairly standard data analysis procedures.

It is the case, for instance, that students showed general improvement in their understanding based on their test scores and from observations by and testimony from the teachers and the researcher, all of which is reported in Chapter 4. However, upon reexamining the data and the overall evaluation process with an orientation toward problem solving, which is presented in Chapter 5, it became more evident that the students progressed in their abilities to connect mathematics and science because they are seen as problem solvers—both in daily activities that had prescribed procedures and in their open-ended final project assignments. Thus, the problem-solving orientation led to enhanced understanding of conclusions already made on a traditional level.

Implications for Future Research

Future research on the use of the Problem-Solving Approach should include its use with larger problems. For the pilot application, it was used with a very small group of students in an individual class. In that situation, the approach seemed successful. But the applicability and practicality of the approach should be explored on a larger scale, perhaps with a curriculum program in use across an entire school or school district. While the Problem-Solving Approach may prove feasible only on a small scale, it may well be that it is more efficient on a larger scale because of its incorporation of program improvement as a part of the evaluation process.
Future attempts with the Problem-Solving Approach should also emphasize more direct use of the problem-solving stance. Specifically, the evaluator(s) should attempt to convey more accurately and completely the fact that problems and potential solution strategies should be openly addressed. The need for open discussion of problems could not be made clear to the teachers at the pilot stage of developing the approach, but may be facilitated in future situations by presenting some of the examples provided herein.

Another issue that might be addressed through future attempts at using the Problem-Solving Approach is that of whether some problem-solving strategies may be more useful when applied to program evaluation than others. In the case of the pilot for the approach, two of the problem-solving strategies (making a drawing and organizing data) were not used directly—or at least they were not identified in the data set analyzed by the researcher. Future investigations might inquire whether these two strategies are less useful in this context, or whether some strategies at least show more promise for their use in evaluation contexts.

Implications for the Field of Mathematics Education

The Problem-Solving Approach to program evaluation holds promise for the field of mathematics education for at least two reasons. First, this approach is understandable to mathematicians because it employs mathematics processes and mathematical ways of thinking. The Problem-Solving Approach would be easily adopted because problem solving is second nature for mathematicians. Most importantly, however, this approach might be more credible because its
processes mimic or at least parallel those going on in mathematics programs themselves. Although this did not occur in the pilot application of the approach, the study of problems at the mathematical task level could certainly aid the evaluator in solving similar problems at the evaluation level. Such an approach comes at time when we are seeking ways beyond just standardized tests to determine whether the Standards (NCTM, 2000) are being implemented, and indeed, this approach would allow for evaluation of the standards at many levels—for individual standards in particular courses, or for the entire set of standards in a system-wide mathematics program.

Second, the Problem-Solving Approach makes it possible for any mathematics teacher to undertake an evaluation of his or her program. Since program evaluation is applicable to any type or size of entity, it may well be performed with just an individual teacher’s class—indeed, that was essentially the case in the pilot application of the approach. At this level, teachers could choose to evaluate and thereby improve any individual courses they teach. In essence, the Problem-Solving Approach could foster participation by increased numbers of teachers in the evaluation process, involving them in a more reflective approach to their own teaching, and creating a focus on what could be linked to or possibly become a form of action research for classroom teaching improvement.
Appendix A

Pretest and Posttest
The instrument used for the pretest and posttest was adapted from materials used in various Texas Instruments Teachers Teaching with Technology Institutes. As participants in the Connecting Mathematics and Science Institute, the teachers of the Integrated Laboratory Mathematics-Science course were granted permission to use the materials with their students.
Sketch the distance (position) – time graph corresponding to each of the following descriptions of the motion of an object.

1. The object moves with a steady (constant) velocity away from the origin.

2. The object is standing still.

3. The object moves with a steady (constant) velocity toward the origin for 5 seconds and then stands still for 5 seconds.

4. The object moves away from the origin, starting slowly and speeding up.

Sketch the velocity-time graph corresponding to each of the following descriptions of the motion of an object.

5. The object is moving away from the origin at a steady (constant) velocity.

6. The object is standing still.

164
7. The object moves with a steady (constant) velocity toward the origin for 5 seconds and then stands still for 5 seconds.

8. The object moves with a steady velocity away from the origin for 5 seconds, then reverses direction and moves at the same speed toward the origin for 5 seconds.

9. The object moves away from the origin, starting slowly and speeding up.

10. Draw the position-time graph for an object moving toward the starting point for 5 seconds then turns around and moves away from the starting point for 5 seconds. Both velocities are numerically equal.

11. Draw the position time graph (the rest point is zero on the x-axis) of a weight that is bobbing up and down on a spring.
12. The volume of a gas is inversely related to the pressure when the temperature remains constant. Sketch the pressure versus volume graph for this situation.

13. The volume of a gas is directly related to the temperature when the pressure remains constant. Sketch the temperature versus volume graph for this situation.

14. Sketch the velocity-time graph for an object that increases velocity at a constant rate.

15. Sketch the velocity-time graph for an object that decreases velocity at a constant rate.

16. The velocity-time graph of an object is shown below. Figure out the total distance traveled by the object.

Distance = _________ meters
Draw the velocity graphs for an object whose motion produced the distance-time graphs shown below on the left. Distance is in meters and velocity in meters per second. (that is, the velocity is the number of meters the object would move in one second.)

Note: Unlike most real objects, you can assume these objects can change velocity so quickly that it looks instantaneous with this time scale.

17.

18.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Appendix B

Student Early Questionnaire
Integrated Math Science Evaluation

Student Questionnaire

Early questionnaire

Please answer and explain your answers to the following questions.

1. How enjoyable do you find mathematics classes?
   Why?

2. How enjoyable do you find science classes?
   Why?

3. How are mathematics and science related?

4. How important do you think it is to see and understand relationships between mathematics and science?

5. What are your reasons for taking this course?

6. What is your experience using graphing calculators?
   Where (School? Home? In a class?)
   How/For what purpose?

7. What is your experience using data collection devices with graphing calculators, such as the Calculator-Based Laboratory (CBL) System or the Calculator-Based Ranger (CBR)?
   Where (School? Home? In a class?)
   How/For what purpose?
Appendix C

Student Late Questionnaire
Integrated Math Science Evaluation

Student Questionnaire

Late questionnaire

Please answer and explain your answers to the following questions.

1. How enjoyable has this course been?

   Why?

2. How are mathematics and science related? Give an example of this relationship.

3. How important do you think it is to see and understand those relationships? Has this course affected your thoughts on this subject?

4. Did this course meet your expectations?

   How? or What was lacking?

5. What was the greatest benefit for you from this course?
Appendix D

Student Early Interview
Integrated Math Science Evaluation

Student early interview

1. How enjoyable do you find mathematics classes? Why?

2. How enjoyable do you find science classes? Why?

3. How are mathematics and science related? Could you give an example?

4. How important do you think it is to see and understand those relationships? Why?

1. What are your reasons for taking this course?

2. What is your experience using graphing calculators? Where (School? Home? In a class?) How/For what purpose?

3. What is your experience using data collection devices with graphing calculators, such as the Calculator-Based Laboratory (CBL) System or the Calculator-Based Ranger (CBR)? Where (School? Home? In a class?) How/For what purpose?
Appendix E

Student Late Interview
Integrated Math Science Evaluation

Student late interview

1. What were the mathematics topics that you studied in this course?

2. What were the science topics that you studied in this course?

3. What were the mathematics and science topics involved in your final project?

4. Has your idea of the relationships between mathematics and science changed as a result of this course? How important do you think it is to see and understand those relationships? Why?

5. Would you feel comfortable acting as a lab assistant for this course next year?

6. What improvements would you suggest for the next offering of this course?
Appendix F

Focus Group Questions
Integrated Math Science Evaluation

Focus Group

1. How enjoyable do you find this class?

2. Why?

3. How comfortable are you with the graphing calculators and data collection equipment?

4. How are mathematics and science related?

5. What aspect of this class helps you most in your understanding of mathematics and science concepts?

6. What types of activities would you like to do more frequently in this class? Why?

7. What types of activities should be done less frequently in this class? Why?

8. What other improvements would you suggest for this course as it continues?

9. If you could choose again, would you enroll in this course, or would you have chosen another option?
Appendix G

Teacher Early Interview
Integrated Math Science Evaluation

Teacher early interview

1. What are your goals in this course for your students? For yourselves as teachers?

2. Has the course had a successful beginning?

3. What problems/issues (immediate and long-range) are you facing as you progress?

4. What do you know/expect of your students at this point? Do you think they will reach the goals you have set for them?

5. How does this evaluation process fit into your overall plan for the course? What will be the role of the evaluation? What will be the role of the evaluator?
Appendix H

Teacher Late Interview
Teacher late interview

1. Did you attain your goals in this course for your students? For yourselves?

2. How have the students progressed individually? (Please discuss each student individually.) How have they progressed as a whole?

3. Were your early expectations of your students accurate? Did they reach the goals you set for them?

4. What problems/issues (immediate and long-range) did you face along the way?

5. How did the evaluation process affect the way you conducted the course?

6. What will you change the next time you offer this course? How will you improve the course?

7. What do you plan to share about this course? With whom?
Integrated Math Science Evaluation

Principal early interview

1. How did this course come into being? Your initiative? Teacher initiative?

2. What do you expect to be accomplished by this course?

3. What would you expect to see if you walked into the classroom where the teachers are teaching this course?

4. What did it take to bring this course into being?

5. (financial issues, scheduling, classroom availability, etc.)

6. What are your plans for this course in the future?
Appendix J

Principal Late Interview
Integrated Math Science Evaluation

Principal late interview

1. Did this course accomplish what you expected/hoped?

2. Describe what you saw when you visited the classroom where this course was offered.

3. What are your current plans for this course? Offer next year? Offer in two years? Never again?

4. What will it take (in terms of finances, scheduling, classroom availability) to offer this course again?

5. What changes (in students, teachers, department and/or school climate) have been brought about by this course?

6. What kinds of feedback on this course have you received from parents? From students? From other teachers?
LIST OF REFERENCES


188

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


191


