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. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
ELEMENTARY STUDENTS' USE
OF SCIENCE PROCESS SKILLS
IN PROBLEM SOLVING:
THE EFFECTS OF AN INQUIRY-BASED
INSTRUCTIONAL APPROACH

DISSERTATION

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

By

Douglas W. Smith, M.A.

*****

The Ohio State University
1997

Dissertation Committee:

Dr. Donald Haefele, Adviser

Dr. Susan Kent

Dr. Barbara Thomson

Approved by

Educational Theory and Practice
ABSTRACT

This research investigated the effects of inquiry-based instruction in elementary classrooms on students' use of selected science process skills. The study was designed to answer the question, "does the use of inquiry-based instruction in elementary school science increase the frequency and/or appropriateness of student's use of science process skills?" Also of interest in the study was whether the level of cooperation exhibited by students working in groups to solve problems would be higher for the students who had been taught with inquiry-based instructional strategies.

Experimental groups for the study were formed from students in two rural fifth grade classes whose teachers had received training in the use of inquiry-based instructional strategies in science. Control groups were comprised of similar classes of students from the same school systems whose teachers had not received the training.

Groups of students were videotaped performing a science-related and generic problem-solving tasks. These tapes were scored
by trained, blind raters. The instrument used was developed and field-tested for this study and is included in the appendices. It was used to evaluate the student groups' use of the science process skills of observing, measuring, predicting, communicating, forming hypotheses, experimenting, controlling variables, recording data, interpreting data, and applying and generalizing results. Scores for frequency and appropriateness of student use of these skills and for group cooperation were recorded by the raters.

Data related to the frequency of use of the skills consisted of tallies based on rater observations and were analyzed using chi-square tests for independence. Results of the analyses indicated that experimental groups used the science process skills more frequently overall than the control groups. Data related to the appropriateness of use of the science process skills consisted of ratings from a five-point Likert scale. These parametric data were analyzed using a repeated-measures ANOVA. The analysis revealed no significant difference between the groups in terms of overall appropriateness of use of the skills. Data related to level of group cooperation consisted of ratings on a five-point Likert scale. These data were analyzed using an ANOVA. Results indicated no difference between experimental and control groups.
To my wife, Gretchen
and my children,
Hannah and Phillip

Thank-you!
I love you!
ACKNOWLEDGMENTS

I wish to thank my adviser, Dr. Donald Haefele for his energetic, consistent, and invaluable guidance through this process. He has read and responded to many pages of my learning. From him I have learned much about being a researcher, a writer and an adviser.

I am indebted to Dr. Susan Kent for her meticulous, insightful and thoughtful input.

I wish to thank Dr. Barbara Thomson for being willing to serve on my committee at the last moment, and doing so with grace and professionalism.

My deepest thanks go to my wife and children who have patiently suffered through several years of the financial, physical and spiritual hardship that graduate school has been.

Finally, any praise, honor and glory which may come about as a result of this document or degree belong to my Lord and Savior, Jesus the Christ. It was accomplished by His grace and exists to shout His praises.
VITA

September 14, 1960........................Bom New Castle, PA

1982....................................................B.S. Psychology, Muskingum College

1983 - 1984........................................Science Teacher, Garfield Heights MS
                                          Garfield Heights, OH

1986 - 1992........................................Science Teacher, Southmoor MS
                                          Columbus, OH

1992 - 1995........................................Graduate Teaching Assistant
                                          The Ohio State University

1994....................................................M.A. Science Education,
                                          The Ohio State University

1995 - 1996........................................Assistant Professor of Education
                                          Bethany College, WV

1996 - present...................................Assistant Professor of Education
                                          Coastal Carolina University, SC

FIELDS OF STUDY

Major Field:  Teacher Education

Minor Field:  Science Education
### TABLE OF CONTENTS

Abstract .................................................................................................................... ii  
Dedication ................................................................................................................. iv  
Acknowledgments ................................................................................................... v  
Vita ............................................................................................................................ vi  
List of Tables ........................................................................................................... x  
List of Figures ........................................................................................................ xii  

Chapters:

1. Introduction .......................................................................................................... 1  
   Purpose of the Study ............................................................................................ 2  
   Justification and Significance of the Study .................................................... 3  
   Research Questions for the Study ..................................................................... 5  
   Background .......................................................................................................... 6  
      The Role of Inquiry .......................................................................................... 8  
   Limitations and Delimitations of the Study .................................................... 12  
   Issues of Internal Validity .................................................................................. 15  
   Summary ............................................................................................................ 18  

2. Review of Related Literature ........................................................................... 19  
   Science Curriculum Reform: 1957 to Present ............................................. 20  
      Introduction .................................................................................................. 20  
      Overview of Reform Efforts ......................................................................... 20  
      Overview of Learning Theory Involved .................................................... 28  
      Theoretical Basis of Current Instructional Reform .................................. 30  

vii
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Selected Science Curriculum Reform Programs</td>
</tr>
<tr>
<td>3.1</td>
<td>Description of Teacher Participants</td>
</tr>
<tr>
<td>3.2</td>
<td>Interrater Reliability Percent Agreement With Expert Rater: First Measurement</td>
</tr>
<tr>
<td>3.3</td>
<td>Interrater Reliability Percent Agreement With Expert Rater: Second Measurement</td>
</tr>
<tr>
<td>3.4</td>
<td>Interrater Reliability Percent Agreement With Expert Rater: Final Measurement</td>
</tr>
<tr>
<td>3.5</td>
<td>Interrater Reliability Percent Agreement With Expert Rater: Averages by Rater and Overall</td>
</tr>
<tr>
<td>4.1</td>
<td>Means and standard deviations of appropriateness-of-use of science process skills for treatment condition by problem type</td>
</tr>
<tr>
<td>4.2</td>
<td>Repeated measures ANOVA for appropriateness-of-use of the science process skills</td>
</tr>
<tr>
<td>4.3</td>
<td>Means and standard deviations for the appropriateness-of-use of individual science process skills by treatment and problem type</td>
</tr>
<tr>
<td>4.4</td>
<td>Results of the omnibus chi-square analysis of the frequency data for experimental versus control groups</td>
</tr>
</tbody>
</table>
4.5 Results of the omnibus chi-square analysis of the frequency data for science versus non-science..........................................................................................97

4.6 Results of the chi-square analysis of frequency data for experimental/science versus control/science..................................................................................98

4.7 Results of the chi-square analysis of frequency data for experimental/generic versus control/generic......................................................................................99
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Basic Science Process Skills</td>
<td>35</td>
</tr>
<tr>
<td>2.2</td>
<td>Integrated Science Process Skills</td>
<td>36</td>
</tr>
<tr>
<td>3.1</td>
<td>Derivation of Sample</td>
<td>50</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

There is a need to improve students' ability to solve problems (Aldridge, 1991; Berman, P., et al., 1988; Education Commission of the States, 1992; Hutto et al., 1990; Rutherford & Ahlgren, 1990; Timpane, 1984, 1982; Yong & Hickrod, 1985). In these and related works (AIMS, 1992; Bredderman, 1983; Padilla, 1986), a strong argument is made for the development of specific skills aimed at equipping students for problem solving (Padilla, Cronin, & Tweist, 1985; Padilla, Okey, & Garrard, 1984). Berliner and Biddle (1995) describe the well-educated high school graduate as being able to "...solve problems, identify problems, offer potential solutions from among alternatives, and implement actions that might reasonably be expected to solve the problems" (p. 301).

Science education reformers (mentioned in Chapter II) are clear about the desire to see students become effective in solving science-related problems using science process skills (Aldridge, 1991; American Association for the Advancement of Science, 1985; Kyle,
Shymansky, & Alport, 1982; Peterson, Bowyer, Butts, & Bybee, 1984; Rutherford & Ahlgren, 1990). There is evidence of what some researchers perceive as a crisis regarding the lack of students' ability to apply problem-solving skills in science (Berliner & Biddle, 1995). This is indicated by reportedly declining achievement (Berman, et al., 1988; David, 1992; National Alliance of Business, Inc., 1991; Shymansky, 1987). There appears to be no argument to the assertion that students would benefit from instruction where the goal is to develop problem-solving skills.

Substantial support for teaching students how to use science process skills is accompanied by the need to analyze and evaluate the instructional approaches used to present this information to the students. By studying instructional approaches and outcomes, teachers and teacher educators may learn about how effectively they are teaching the science process skills and how well their students are employing them (Cruickshank, Bainer, & Metcalf, 1995, p. 259).

Purpose of the Study

This study examined the effects of teacher instruction in science process skills on students' ability to solve science-based problems. The frequency and appropriateness with which students used the skills in different kinds of problem-solving situations were
examined. Teachers in experimental groups had agreed to teach their classes using an inquiry approach and to instruct their students in the use of science process skills when they engage in problem situations. The experimental groups for this study consisted of students who were taught using an inquiry-based approach.

Justification and Significance of the Study

As previously noted, increasing students' facility in the use of science process skills is a highly desirable outcome of science education. This study sought to examine students' use of science process skills in an environment where their teachers had been linked with environmental science specialists, trained in inquiry-based instruction of science, and were implementing this kind of instruction in their classrooms.

The training the teachers received was based upon the literature and research in science education reform. Appendix H contains the details of the training program and its goals. In summary, the training consisted of an examination of the processes involved in problem solving in the elementary science curriculum, and ways in which elementary teachers can organize instruction to promote development of problem solving process skills in their students. The training encouraged teachers to model an inquiry-
based approach to problem situations for their students and to instruct their students in the use of science process skills. This practice of teaching science using inquiry-based approaches, grounded in science education and child development literature, appears to display promise in addressing the concerns expressed by science education reformers for developing in students the ability to solve science-related problems.

If the ability to learn and apply science process skills is strongly formed, it is reasonable to expect that these habits may transfer the skills from the arena of science-related problem solving to generic problem solving (Aldridge, 1991, p.4). It is important that students are able to solve problem situations that are not specific to science (or any other isolated part of the curriculum). The problem situations children encounter in their world are rarely one-dimensional. According to Timpane (1984), knowledge about problem-solving in one area of the curriculum will enable the child to transfer problem solving skills to more generic problems.

Finally, demonstration of the level of effectiveness of inquiry-based instructional strategies at instilling problem-solving skills in elementary school students has implications for preservice elementary science teacher training. The effectiveness of this instructional approach at enabling elementary students to become
better problem solvers may have implications for elementary science teaching methods courses. The use of inquiry-based (hands-on) instruction is supported by theory (Piaget, 1974). It is important to demonstrate that this instructional method is supported by research.

Research Questions for the Study

In this section, research questions that were investigated in this study are presented. In these questions, the term "generic" is used to refer to problem situations where the solutions do not presuppose mastery of the knowledge in any specific area of the elementary school curriculum. More information about these questions is presented in Chapter 2, particularly at the close of the chapter. The questions are;

• With what frequencies do experimental and control group students use science process skills in science-related problem solving situations?

• With what frequencies do experimental and control group students use science process skills in generic problem solving situations?

• How appropriately do experimental and control group students use science process skills in science-related problem solving situations?
• How appropriately do experimental and control group students use science process skills in generic problem solving situations?
• Does the level of cooperation within groups vary between classes where students are taught using inquiry-based instructional strategies and classes where students are not taught using inquiry-based instructional strategies?

Background

Since the late 1950's, the quality of science instruction has increasingly become a matter of concern to both the education and business communities. Fueling these sentiments have been reports
• about the declining trend of achievement of American students versus the students of other industrialized countries (Berman, et al., 1988; David, 1992; National Alliance of Business, Inc., 1991; Shymansky, 1987).
• about the question of the relevancy of current science curricula considering the changing and increasingly technological nature of society (Berman, et al., 1988; Education Commission of the States, 1992; Yong & Hickrod,1985).
• about the under representation of women and minorities in upper level science classrooms (Aldridge, 1991; Rutherford & Ahlgren, 1990); and
that American students are not instructed in a way that encourages the development of problem-solving skills (AIMS, 1987) which are vital to learning the content of science (Rutherford & Ahlgren, 1990).

Recent reform efforts in science education have pointed to the importance of students developing an understanding of the physical and natural laws and principles governing their world. Authors in science recommend that students of all ages should first experience these laws as they are demonstrated in naturally occurring phenomena. Later the theories that address those observed phenomena should be presented and explained as they relate to the students' experiences (Aldridge, 1991; NSTA, 1987; Rutherford & Ahlgren, 1990). If students, particularly elementary students, are to be successful in the use of science process skills, they must be equipped differently than they are presently. The current trend in science education is to view the curriculum as isolated segments. For example, at the high school level, biology would be followed by chemistry, then physics, etc. (Aldridge, 1991). Additionally, the concepts within each discipline tend to be approached theoretically first, followed later by examples. As a result, limited hands-on or experientially based instruction occurs. Similarly, young students receive science instruction as a collection of loosely related facts
which are tied to other areas of the curriculum (Aldridge, 1991). So, the relationships between science, other areas of the school curriculum, and students' prior experiences are not regularly explored. Consequently, habits of inquiry and problem solving are not regularly fostered.

In order for young students to be able to confront the problem situations in their world and to understand how they might begin to contribute to their solutions, a firm grounding of students in science process skills has been called for (Aldridge, 1991; Rutherford & Ahlgren, 1990). A strong foundation of inquiry in science, emphasizing development of science process skills, which is laid in the elementary grades, may significantly affect the expansion and use of these skills at the high school level.

The Role of Inquiry in the Study

Inquiry, from the Latin quiare, means "to seek or search into" (Webster's Dictionary, 1978). Bruner (1968) said that inquiry is an attitude toward learning characterized by guessing and hunches, toward the possibility of solving certain problems on one's own. Because he does not specify skills or protocols for the process of inquiry, Bruner's definition can be interpreted in many ways.
Inquiry is also not a phenomenon that is peculiar to science. This study, however, focuses on the use of inquiry in elementary science.

Science educators vary in their approach to inquiry. The concept of inquiry appears to exist as a continuum where, at one extreme, elementary science teachers are centrally in charge of the investigation. Here, the teacher selects the problem, establishes methods of investigation, and guides students toward a solution. At the other extreme are those teachers who serve as a facilitator. In this case, the students establish the questions and problems for study along with the methodology of their investigation (Posner et al., 1986). In the current study, the term 'inquiry' is characterized as the kind of study which engages the student in the investigative nature of science (Haury, 1993, 1995). On the continuum mentioned at the beginning of this paragraph, Haury's notion of inquiry can be considered a compromise between the two extremes. Both the students and the teacher are central to the process of inquiry.

In classrooms where inquiry is emphasized, content is not the issue of primary importance. Of significance in the lesson is that students are engaged in scientific thinking and actions: asking questions, manipulating materials, observing phenomena, and constructing explanations to answer their questions (Wise & Okey, 1983). It is through encouraging students to identify and solve their
own problems (encouraging them to develop their facility with the process skills of science) that implementation of the inquiry approach to science instruction contributes to the development of scientifically literate students.

The reform efforts referred to above (Aldridge, 1991; NSTA. 1987; Rutherford & Ahlgren, 1990) support the use of inquiry-based instruction because they base their approach to science instruction in constructivism. The constructivist view of learning reflects the notion that experiences aid students in their development of conceptual frameworks from which they are able to construct more abstract interrelationships between the curriculum and their lives (Driver, 1986; Piaget, 1973; Posner, et al., 1986; Vygotsky, 1977). Piaget (in Woolfolk, 1990, p. 47) characterizes the developmental level of elementary school children (between six and ten years of age) as primarily concrete operational with a small representation of learners in both the pre-operational and formal operational stages. This suggests that learning in elementary students would best be addressed through the use of concrete, experience-based instructional strategies. The results of research related to instructional strategies conducted by Boulanger (1981) and Wise & Okey (1983) support the value of the use of Piagetian developmental theory to guide teachers' choices of instructional strategies. Both
studies reported higher student achievement in science when concrete instructional strategies (for example, manipulatives) were used than in cases where more abstract strategies (for example, text-driven lectures) were used. Relatedly, Posner et al. (1982) discussed the importance of accomplishing conceptual change in science students from their preconceptions about scientific phenomena to accepted scientific explanations. In this work they note the importance of a student's previous experience in the formation of preconceptions and the necessity of using experiences which conflict with the student's preconceptions (discrepant events) to break down those preconceptions. This process occurs best when the students confront the discrepant events in a setting of inquiry. These findings demonstrate that the use of experience- and inquiry-based approaches to science teaching have a greater likelihood of producing more desirable achievement and conceptual outcomes in elementary school age students than do more abstract approaches to instruction.

Rutherford & Ahlgren (1990), the Carnegie Corporation (1989), Goodlad (1983), Aldridge (1991), and the National Research Council (1989) agree that inquiry-based instruction is necessary in order for the American educational system to produce scientifically literate students who are capable of performing the problem-solving activities both extant and anticipated in society. Recent reform
efforts in the area of science education have argued for the implementation of a specific type of experience-based approach to instruction. Science curriculum reform movements (described in Chapter II) and the literature on inquiry-based instruction, indicate that this approach to instruction holds promise for addressing the concerns mentioned above related to production of scientifically literate students.

Limitations and Delimitations of the Study

This study was quasi-experimental. Because it occurred in a naturalistic setting, it was not possible to ensure random assignment of subjects or to control all the relevant variables. Kennedy & Bush (1985) characterize this situation as an experimental study...containing a manipulated variable but... lacking full control over potentially contaminating variables. Thus, in such situations, the researcher can never be sure that observed findings result from manipulation of the independent variable or from variation in uncontrolled variables—or, perhaps, both (p.4).

As is true with such a study, the findings of this study will not be widely generalizable beyond the sample of the subjects involved. At best, generalization may extend to like schools with populations and environments that mirror those features in these schools. In that this study was performed in a naturalistic setting, there were
extraneous variables that were not controllable. Examples of these are: attendance, pull-out programs, school functions, complications with subjects occurring during data gathering, and other situations and circumstances that are particular to these schools. As best the researcher can determine, none of these factors affected the outcome of the study.

The researcher went to the selected classes and was introduced to the students. The design of the study necessitated the use of video taping the students as they solved problems in groups. Blind raters were then used to evaluate the groups' use of science process skills in an effort to eliminate the internal validity threat of experimenter bias. The presence of technology and the researcher in the classroom presented a reactivity (Hawthorne) effect threat to the external validity of the study. The addition of video equipment further restricted the ability to generalize to other elementary-age subjects and other environments. Considering this, during the initial visit, the technology (cameras and microphones) to be used in the data collection was explained and demonstrated. Also, the set-up of the room was discussed and the students' role in the problem-solving activities was explained. Permission slips were distributed. The researcher returned about a week later and collected the data. The
video tapes were then rated by trained raters using the Science Process Skills Observation Form.

The testing content was controlled and was science-related in two cases and generic in the other two cases. Again, generalizability of the results is very narrow because the breadth of the content represented in the cases was not sufficient to represent the whole of the possible range of content in all science and generic situations. Another area of limitation was the process of rating the video tapes. Even though the raters of the videotapes were trained in their tasks, the subjective nature of the "appropriateness" and "group cooperation" scales on the rating form left room for potential differences between raters' interpretations of what they observed (see Chapter 3.) A measurement of inter-rater reliabilities on all three scales; frequency of use, appropriateness, and group cooperation are reported in Chapter 3.

Regarding the assignment of variables, the teachers who received the aforementioned training were self-selected --a selection by treatment effect concern. As this represented a considerable threat to the internal validity of the study, the researcher interviewed the teachers' principals in order to determine the extent to which this threat was real. All the school principals reported that the assignment of the students to the classes/teachers within grade
levels was randomly done in all cases. They reported that the overall characteristics of each class were similar and that within each class, there was a broad range of academic ability. The researcher did not examine the students' cumulative records for reasons related to student confidentiality. Although the teachers' participation in the training program was through self-selection, the principals all reported to the researcher that between the "experimental" and control teachers, they did not feel there existed significant differences in terms of general teaching ability or style. A detailed description of teachers and their students is reported in Chapter 3.

Issues of Internal Validity

Existence of a threat, that is a rival hypothesis, can invalidate or severely restrict the validity of the data. Campbell & Stanley (1971) have identified eight threats to the internal validity of the study. They are; history, reactivity (already discussed above), maturation, testing, instrumentation, statistical regression, differential selection of subjects, mortality, selection-maturation interaction, and experimenter bias. The potential relevance of these issues to this study is discussed below. This study is classified as a static group comparison (Gay, 1992). The design includes two or more groups where one half of the groups have received the
experimental treatment and the other half have not. In the current study, the experimental treatment was the implementation of inquiry-based science instruction in the classroom. Control groups were established where inquiry-based science instruction was not offered.

Two aspects to the design of the study were responsible in eliminating several of the issues of internal validity. The first was that the data were only collected once. This ruled out the history and testing effects. The other characteristic of the design contributing to the elimination of some issues of internal validity was that the subjects (students) had been assigned to their classes on a random basis. This eliminated the effects of statistical regression, differential selection of subjects, and mortality.

"Maturation refers to physical or mental changes that may occur in a subject over a period of time" (Gay, 1992, p. 304). Although this limitation was not controlled by the design of this study, an attempt to control for maturation was made by choosing control subjects of the same grade level as the subjects in the experimental groups and by scheduling all the data gathering within a two week period. The duration of the study eliminated a maturation effect with regard to development. As to boredom, fatigue, motivation, etc., these were not observed.
Instrumentation refers to invalidity, unreliability, or lack of credibility and consistency respectively, in measuring instruments which may result in an inaccurate assessment of performance (Gay, 1992). The instrumentation threat posed in this study was the use of raters who might be inconsistent about interpreting the meanings of what they observed. The use of trained, blind raters in the study was an attempt to control for this threat to internal validity. The results of the rater training and inter-rater reliability testing are reported in Chapter 3. Reliability of the raters was determined during the training and is reported in Chapter 3.

Selection interaction refers to the possibilities that exist for any of the groups which have been selected, to have an initial advantage over the other groups as a result of maturation, history, or testing factors (Gay, 1992, p. 306). For this study, the greatest likelihood for an interaction was between maturation and selection. Although this was a possible threat to validity, the examination and comparison of the control and experimental groups did not reveal initial differences between the groups.

The issues of external validity listed by Gay (1992) are pretest interaction and multiple treatment interaction. Since the study design contained no pretest and only one treatment, these issues of external validity were not relevant to this study.
Summary

Much support has been cited in the preceding pages for the instruction of students in the use of problem-solving skills. This study focused on the implementation of inquiry-based science instruction into several rural elementary school classrooms. Some of the teachers in these schools participated in a summer institute, during which they were grouped with other teachers in their geographical area and with professionals from the Ohio Department of Natural Resources. The participants devised action plans for implementing inquiry-based science instruction in the teachers' classrooms and committed to implement them throughout the following school year.

The purpose of this study was to examine the effects of the inquiry-based instruction as measured by the students' frequency and appropriateness of use of the science process skills when presented with both science-related and generic problem situations.
CHAPTER 2
REVIEW OF RELATED LITERATURE

The purpose of this study was to determine if the implementation of inquiry-based instruction resulted in an increase of students' use of the science process skills in problem solving situations which are science-related and generic. This chapter will provide a review of literature which supports inquiry into this phenomenon. The literature reviewed in this chapter relates to: a) a perspective of the various reform efforts in science education since 1957; b) research and other literature related to the nature and success of both partnerships in general and the specific partnership which serves as the venue for this study; and c) a review of the scholarly literature and research related to science process skills, inquiry-based instruction, and the relationship between the two.
Science Curriculum Reform: 1957 to Present

Introduction

Kyle (1986) describes the period of history that followed the advent of Sputnik I as, "the Golden Age of Science Education" (p. 4). Society began to perceive its needs from the fields of science to be much greater than they had been in the past. As a result, the financial support of the National Science Foundation (NSF) and other similar agencies for research into and reform of science education increased and many programs and approaches to science curricula were written, adopted by various schools, and put into action. Table 2.1 lists several of the better-known and more widely adopted science curriculum reform efforts as noted by Kyle (1986). During this time, the focus of the reform efforts was on the content of the science curriculum, while approaches to instruction of these new curricula remained essentially the same (Hofstein & Yager, 1982; Kyle, 1982; Peterson, Bowyer, Butts, & Bybee, 1984).

Overview of Reform Efforts

Although the new programs address a wide variety of aspects of the science curriculum, they have some common underpinnings. These curricula were the result of collaboration between scientists, educators, and members of the field of psychology (Kyle, Shymansky,
<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elementary Science Curricula</strong></td>
<td></td>
</tr>
<tr>
<td>Elementary Science Study</td>
<td>(ESS)</td>
</tr>
<tr>
<td>Science Curriculum Improvement Study</td>
<td>(SCIS)</td>
</tr>
<tr>
<td>Science- A Process Approach</td>
<td>(SAPA)</td>
</tr>
<tr>
<td><strong>Middle School/Junior High Curricula</strong></td>
<td></td>
</tr>
<tr>
<td>Earth Science Curriculum Project</td>
<td>(ESCP)</td>
</tr>
<tr>
<td>Individualized Science Instructional System</td>
<td>(ISIS)</td>
</tr>
<tr>
<td>Interaction Science Curriculum Project</td>
<td>(ISCP)</td>
</tr>
<tr>
<td>Intermediate Science Curriculum Study</td>
<td>(ISCS)</td>
</tr>
<tr>
<td>Introductory Physical Science</td>
<td>(IPS)</td>
</tr>
<tr>
<td><strong>High School Curricula</strong></td>
<td></td>
</tr>
<tr>
<td>Biological Science Curriculum Study</td>
<td>(BSCS)</td>
</tr>
<tr>
<td>- Yellow Version</td>
<td></td>
</tr>
<tr>
<td>- Blue Version</td>
<td></td>
</tr>
<tr>
<td>- Green Version</td>
<td></td>
</tr>
<tr>
<td>Chemical Bond Approach</td>
<td>(CBA)</td>
</tr>
<tr>
<td>Chemical Educational Materials Study</td>
<td>(CHEM Study)</td>
</tr>
<tr>
<td>Harvard Project Physics</td>
<td>(HPP)</td>
</tr>
<tr>
<td>Physical Science Study Committee</td>
<td>(PSSC)</td>
</tr>
</tbody>
</table>

Table 2.1: Selected Science Curriculum Reform Programs
Alport, 1982). Before this time, science programs and instructional materials did not typically represent a collaborative cross-section of the disciplines upon which practice in science teaching and learning rests (Kyle, 1982; Kyle, Shymansky, & Alport, 1982; Peterson, Bowyer, Butts, & Bybee, 1984). Table 2.1 provides a listing of several programs, conceived following the Sputnik era, which are considered by the abovementioned parties to be exemplary.

Among the characteristics of these exemplary programs is that they all represent collaborative efforts between scientists, science educators, and educational theorists (Kyle, 1982; Kyle, Shymansky, & Alport, 1982). The educational community must learn from this process to value the effectiveness of a collaborative effort in the creation of a product that is, by its nature, an amalgam from many disciplines.

Following the "alphabet soup" reform period, the science education community began to examine the requisite knowledge, skills, and attitudes to effectively teach these new curricula with a focus on learner satisfaction (Conrath, 1987). This gave rise to the use of new approaches to the delivery of the new curricula (primarily based upon a constructivist view of learning, described later) aimed at addressing several of the current major concerns which have been raised about science education, notably:
• the declining trend of achievement of American students versus the students of other industrialized countries (David, J.L., 1992; National Alliance of Business, Inc., 1991; Berman, P. et al., 1988; Shymansky, J.A., 1987);
• the question of the relevancy of current science curricula in light of the changing and increasingly technological nature of Society (Education Comission of the States, 1992; Berman, P. et al., 1988; Yong, R.S., Hickrod, G.A., 1985);
• the underrepresentation of women and minorities in upper level science classrooms (Aldridge, 1991; Rutherford & Ahlgren, 1990; Sivertsen, 1993); and
• student achievement in science curricula and the student's science attitude (Stofflett, 1991).

Regarding the first point in the above list, Berliner and Biddle (1995) state that the difference in the scores of students on an international scale is not what it appears. In fact, the differences between America and other countries, "particularly Japan, Korea, and Western Europe" (p. 51) can be more accurately accounted for by examining underlying factors such as national views of childhood and education (p. 52); "biased samples in the research" (p. 54); and curricular requirements (p. 55). Berliner and Biddle (1995) also assert that the differences between countries in terms of student
achievement are not as significant as reported by the media in general and specifically by reports like "A Nation At Risk" (p. 14). Berliner and Biddle also deny the existence of a decline in American students' achievement;

We confront these assertions with a novel approach; we review here the actual evidence generated by standardized tests. From that evidence, readers will learn that standardized test data reveal no recent drop in student achievement; indeed, many of the tests indicate modest recent gains in students' knowledge. In fact, we know of only one standardized test that has ever generated falling aggregate test scores—the Scholastic Achievement Test (SAT)—but, as we shall show, that decline had nothing to do with average student achievement. In fact, when analyzed correctly, the SAT data also reveal a pattern of achievement growth! In short, real evidence indicates that the myth of achievement decline is not only false—it is a hysterical fraud (p. 14).

Although there is controversy over the issue of declining student achievement, Berliner and Biddle's work does not specifically address the achievement of students in science. The fact remains that the issue of declining student achievement, regardless of its actual merit, was and continues to be both an area of interest and a driving influence in the reform of science education.

The path toward reform which science education has taken has been aggressive both in terms of the structure of the curriculum (Education Commission of the States, 1992; Berman et al., 1988;
and the approaches with which it is taught (Kyle, 1982; Kyle, Shymansky, & Alport, 1982; Peterson, Bowyer, Butts, & Bybee, 1984). Current prominent reform efforts in science education are Project 2061 (American Association for the Advancement of Science, 1985); Science for All Americans (SAA) (Rutherford & Ahlgren, 1990); and Scope, Sequence, and Coordination of Secondary School Science (SSC) (Aldridge, 1991). These reform efforts seek to provide for both the depth and breadth of experiences which students need to be well grounded in the facts and importance of science to their lives (Rutherford & Ahlgren, 1990). They propose to accomplish this by changing the structure of current K-12 science curricula and by reforming the instructional strategies and teaching techniques used in the K-12 classrooms.

There are many specific examples which can be cited in terms of comparison of the current reform efforts in science education. In general, four areas of commonality will be addressed here. They are:

- all children need and deserve an education in science and technology (Aldridge, 1991; Meyer & Fortner, 1990; Rutherford & Ahlgren, 1990; Sivertsen, 1993);
- the importance of a constructivist approach to the curriculum (Rutherford & Ahlgren, 1990; Aldridge, 1991; Boulanger, 1981;
students must understand and appreciate the interrelatedness within and among the various areas of science and technology, (Aldridge, 1991; Rutherford & Ahlgren, 1990; Meyer & Fortner, 1990, Koballa in Holdzkom and Lutz, 1986); and

- science should teach less but teach it better (Aldridge, 1991; Rutherford & Ahlgren, 1990; Meyer & Fortner, 1990).

SAA and SSC both emphasize the need to especially address the needs of girls/women and minorities (Rutherford & Ahlgren, 1990; Aldridge, 1991). SSC uses an integrated approach to the curriculum on all levels in order that all students get to experience all facets of the science curriculum. This is the result of Aldridge's (1991) contention that the "layer cake" approach is not adequately reaching the majority of students with all areas of science. He indicates that 80% of high school sophomores enroll in biology, 40% of high school juniors enroll in chemistry, and only 20% of high school seniors enroll in physics. This means that only 20% of the students in high school receive a comprehensive education in science (Aldridge, 1991).

Both efforts also emphasize the importance of a constructivist approach to the curriculum. SAA begins its list of the facets of science literacy with the notion that students should gain a familiarity with the natural world. In SSC, many statements are
made to the effect that the students should contact the phenomena of
their worlds experientially before addressing the theoretical
underpinnings of phenomena. As an example, in describing the
underlying questions of SSC, Aldridge states that the question of
"What do we mean?" can only be understood through experience
with a certain concept. He gives the example that gravity is just a
term until it is experienced and that outside experience with gravity,
one cannot understand it.

A third area of commonality is that students must understand
and appreciate the interrelatedness within and among the various
areas of science and technology. SAA supports the notion that along
with experiencing the world, students should understand that/how
various phenomena are interwoven with each other. SSC uses spaced
learning (Dempster, 1988) in order to accomplish the coordination
aspect of the program and is quite against the current "layer cake"
(as described by Aldridge, 1991) approach which isolates the various
parts of the science curriculum. This notion is supported by
Dempster's (1988) concept of spaced learning which can be described
as a process by which the forging of connections between parts of
science and, more globally, other areas of the school curriculum are
facilitated and encourged.
Finally, the current reform efforts champion the cause of teaching less but teaching it better. This is directly in line with the notion that the present science and mathematics curricula are overstuffed and undernourished (Rutherford & Ahlgren, 1990). Science reformers have chosen to decide what concepts are necessary for scientific literacy, and then to study them in depth rather than to attempt to address all areas within the science curricula, but not address any to significant depth (Rutherford & Ahlgren, 1990).

Overview of Learning Theory Involved

It has been mentioned above that the most appropriate approach to the presentation of science-related curricula is found in the notion of constructivism. Agwubike (1985), characterizes the constructivist-based approach to science education as consisting of experiences where the students can physically encounter the concepts and facts of science. This contrasts with the more traditional competency-based approach where knowledge, skills, and literacy are addressed in more text and lecture-driven classrooms. This the constructivist approach to science instruction supports the intentions of the previously mentioned science reformers.
Cognitive developmental theorists, Piaget (1973), Vygotsky (1977), Driver, (1984, 1989) and other constructivists describe the acquisition of knowledge as the product of a person's experiences. Piaget (1973) spoke of the importance of the student engaging actively with the world and developing schema to aid in understanding. When the child encounters situations that do not conform to his/her schema (ways of understanding the world), a condition of disequilibrium motivates initiation of a search for an adequate explanation. The search will either culminate in the reconciliation of the discrepant event with existing schema or will cause a new schema to be developed.

In addition to the foundational work of Piaget, the theory of Vygotsky is applicable to these reform efforts in that both support the use of heterogeneous grouping in the classroom (Aldridge, 1991; Rutherford & Ahlgren, 1990). Aldridge (1991) states that the SSC model encourages the grouping of students with different aptitudes in science and the SAA model speaks about being able to use scientific knowledge for personal and social purposes. Vygotsky's notions of the zone of proximal development and the process of scaffolding are useful in small group and peer instruction, both strategies inherent within the designs of both reform efforts. The aspect of both SSC and SAA which requires that the student's initial
contact with new information should happen on the phenomenological/experiential level rather than being introduced to the concepts via the presentation of scientific laws, allows the Conceptual Change Model (CCM) (Posner et al., 1982) to be used as a guide to practice regarding teacher's instructional strategies.

Theoretical Basis of Current Instructional Reform

The word "concept" has appeared repeatedly in educational literature and is a common term in everyday language. Bourne et al. (1986) say that most of what we know about the world involves concepts and relations among concepts.

According to Woolfolk (1990), concepts are categories we use to group similar objects, ideas, events, or people. Further, concepts do not exist in the real world, only individual examples of concepts exist in a concrete way (p. 264). That which is already known can be said to exist conceptually as an arrangement and internal categorization of previous experiences based upon common defining attributes (Bruner, 1973; Ausubel, 1968; Gagne & Driscoll, 1988). Rosch (1973, 1975) and Tennyson (1981) speak about prototypes as aides to childrens' grasping (via assimilation or accommodation, see Piaget, 1973) of concepts and the importance of adult involvement in pointing out the appropriateness of a particular prototype as an
example of a concept. Klausmeier (1976) also points out the importance of non-examples of concepts in facilitating the learner's acquisition of the knowledge with appropriate and correct understanding. It can be inferred from the above information that the acquisition of concepts by the learner happens via a process of interaction with the environment, and that the internalization and categorization of information into concepts is seen as a function of the learner's experiences. This relates closely to a Piagetian/constructivist view of learning.

According to Piaget's theory of cognitive development, children pass through several stages as they become more and more able to process incoming information. Although each of these stages has a discrete set of characteristics which separates it from the others, it is important to note that the stages are continuous rather than discrete (McGraw, 1987) and that a child's progress through them represents qualitative changes in thinking (Seigler, 1991).

Piaget (1954) uses the term "schema" to denote conceptions extant within a child. A schema is a collection of concepts with attributes which are functionally related to one another, such that their use in combination with each other facilitates a child's efficient and thorough understanding of and ability to function in his/her world. Piaget tells us that children deal with new information
through one of two processes, assimilation or accommodation. Operationally, Piaget suggests that the learner, when confronted with a new situation or new information, will first seek to categorize that information by attributes which are similar to those of an already existing schema (assimilation). Failing this, the learner will use the defining attributes of the new situation or information and will either create a new schema for it or will revise the definition of an existing schema to allow the new information to "fit" (accommodation). Essentially, Piaget believes that the child is forced to reconcile new and/or disparate information or circumstances with what is already known. A student who "experiences" a concept in a setting where the concept interacts with her/his life might be more likely to understand the connections between that concept and other concepts. The inquiry-based approach to science instruction provides opportunities for students to have these experiences.

Robert Sternberg's (1985) Triarchic Theory of Intelligence includes both provision for linking intelligence and the environment, and a mechanism for modifying intelligence through experience (Sternberg, 1986, 1988). The processing components deal specifically with the ways in which people solve problems. Key to the processing of information is the ability to acquire knowledge. Following acquisition, there is a performance component which is
where the problem is actually worked out and solved. The final processing component is the metacomponent which has as its function the organization and management of the other parts of intelligence. Sternberg (1988) explains that the three components actually interact with each other when the person is faced with a complex task, for example, writing a term paper.

The environmental aspect to Sternberg's theory demonstrates that intelligent people are able to use the environment to accomplish goals and tasks by utilizing one of three options. The person can adapt to the environment, change the environment, or select out of it. The final component of Sternberg's Triarchic Theory of Intelligence is closely analogous to the concepts of accommodation and assimilation in Piaget. It is the notion that the person will experience something and will have to either fit the new information or experience into an existing pattern, or will have to create a new pattern which will allow the person to make sense of the novel situation or information. As with Piaget, Sternberg's theory supports the use of inquiry-based instructional strategies as a means of providing experiences from which students can construct their understanding of concepts in science.
Science Process Skills

Knowledge is an internal phenomenon which is arrived at via a process of inquiry (O'Hair & Odell, 1995, p. 317). The role of the teacher in the classroom must be one of facilitation of an exploratory process by the students. Consequently, when a teacher seeks to teach science using an inquiry-based approach, the instructional objectives of the lesson must be arrived at via a process of hands-on investigation by the students (Matthews, 1992). In order to carry out these investigations, students must possess certain skills related to each step of the investigatory (experimental) process (Brotherton & Preece, 1995). Since it is the purpose of this study to examine how well elementary school children use science process skills following a course of inquiry-based instruction, it is necessary first to have an understanding of the nature and definitions of the skills.

The science process skills have been defined and researched by individuals in science education (for example: Brotherton & Preece, 1995; Padilla, 1986; Rakow, 1986) and were the focus of the American Association for the Advancement of Science (1965) curriculum reform program, Science- A Process Approach (SAPA). They are divided into two categories; Basic Process Skills and Integrated Process Skills. Although different authors vary in the verbage they use to define the skills, the meanings of the definitions
do not vary. Figures 2.1 and 2.2 provide definitions of eight of the
 ten skills related to this study. The other two skills are defined in
 the text following the figures.

These skills are described as a set of broadly transferrable
abilities, appropriate to many of the science disciplines and reflective
of the behaviors of scientists (Padilla, 1986). Although there are
process skills for other areas of the curriculum (mathematics and
social studies), they largely overlap (AIMS, 1992), and according to
Gagne (1965) the first discipline within which specific process skills
were defined and used was science. The American Association for
the Advancement of Science (AAAS) has defined the science process
skills in two categories. The basic process skills of observing,
predicting, inferring, measuring, communicating, and classifying, and
the integrated process skills of controlling variables, defining
operationally, formulating hypotheses, interpreting data,
experimenting, and formulating models.

Two skills pertinent to this study, but not mentioned by Padilla
were recording data and inferring/applying and generalizing results
(AIMS, 1992). Recording data is described as written expression of
facts, observations, measurements, etc. Inferring/applying and
generalizing is described as using the findings of an experiment to
Basic Science Process Skills:

Observing - using the senses to gather information about an object or event. Example: Noting a pencil is yellow.

Measuring - using both standard and nonstandard measures to describe the dimensions of an object or event. Example: Using a meter stick to measure the length of a table in centimeters.

Communicating - using words or graphic symbols to describe an action, object, or event. Example: Describing the change in height of a plant over time in writing or through a graph.

Predicting - stating the outcome of a future event based on a pattern of evidence. Example: Predicting the height of a plant in two weeks time based on a graph of its growth during the previous four weeks.

Figure 2.1: Basic Science Process Skills (Padilla, 1986)
**Integrated Science Process Skills:**

Controlling variables - being able to identify the variables that can affect an experimental outcome, keeping most constant while manipulating only the independent variable. Example: Realizing through past experience that amount of light and water need to be controlled when testing to see how the addition of organic matter affects the growth of beans.

Formulating hypotheses - stating the expected outcome of an experiment. Example: The greater the amount of organic matter added to the soil, the greater the bean growth.

Interpreting data - organizing data and drawing conclusions from it. Example: Recording data from the experiment on bean growth in a data table and forming a conclusion which related trends in the data to variables.

Experimenting - being able to conduct an experiment including asking an appropriate question, stating a hypothesis, and identifying controlling variables, operationally defining those variables, designing a "fair" experiment, conducting the experiment, and interpreting the results of the experiment. Example: The entire process of conducting the experiment on the affect of organic matter on the growth of bean plants.

Figure 2.2: Integrated Science Process Skills (Padilla, 1986)
answer experiment-related questions and to address situations that are similar but not directly related to the experiment.

**Acquisition of Science Process Skills**

Many researchers have investigated the issue of the acquisition of both basic and integrated process skills (Allen, 1973; McKenzie & Padilla, 1987; Padilla, Cronin, & Tweist, 1985; Padilla, Okey, & Garrard, 1984; Quinn & George, 1975; Theil & George, 1976; Tomera, 1974; Wright, 1981). Among the findings of this research were that there are three teaching strategies which proved to be most effective in terms of students acquisition of science process skills. According to Tomera (1974) the strategies were:

- applying a specific set of rules for predicting,
- using activities and pencil and paper simulations to teach graphing, and
- using a combination of explaining, practice with objects, discussions and feedback with observing.

Wellman (1978) revealed that direct manipulative experiences in science enhanced the development of science process skills in children grades K-3. She also reported a positive correlation with their success in beginning language and reading achievement. Wellman went on to discuss benefits from science instruction in the
areas of strengthening development of vocabulary, verbal fluency, ability to think logically, and improved concept formation and communication. Esler (1981) came to several of the same conclusions as Wellman, however, he found that scientific inquiry and activities also enhanced the performance of children in mathematics and other subjects as well as language arts. This notion is supported by Pauker and Roy (1993) via their claim that the basic science process skills can be subsumed by the category of critical thinking skills. Bredderman (1983) states that with the use of inquiry-based science programs, teachers can expect substantially improved science process and creativity; modestly increased performance on tests of perception, logic, language development, science content, and mathematics; modestly improved attitudes toward science and science class; and pronounced benefits to disadvantaged students. Bredderman concludes that students in inquiry-based science curricula clearly outperformed those in traditional read-recite classes especially in process skills (20 percentile points higher for the inquiry-based classes). Padilla, Okey, and Dillashaw (1983) found a +.73 correlation between the integrated process skills mentioned above and the formal thinking skills postulated by Piaget. Padilla (1986) points out that one of the tests Piaget used to determine
whether a child was in the formal operational stage was to ask him/her to design an experiment to solve a problem.

This literature clearly indicates that one should expect that the use of inquiry-based instructional strategies will produce an increase in the acquisition of science process skills in elementary school children. It is reasonable to expect that the students in this study will show increased use of the science process skills. It is necessary, at this point, to explore the literature relating to effective ways to measure students' acquisition/use of science process skills.

Assessment of Science Process Skills

Science education has accepted that the constructivist notion (as characterized in Driver, 1984, 1989) of hands-on activities in learning is the most effective approach to the curriculum in terms of conveying to the students more complete understandings of both the information in science and its relationships to their lives (Rutherford & Ahlgren, 1990). Many of the proposals for science reform reflecting a commitment to the hands-on approach have called for an improvement in the process of assessing science learning outcomes (AAAS, 1989; Brotherton & Preece, 1995; Brunkhorst, 1991, Sivertsen, 1992). Raizen and Kaser (1989) report that the current practice of using pencil-and-paper standardized tests of achievement
is not adequate to assess a student's ability to use problem solving
skills or to conduct scientific investigations and experiments.

Performance-based assessment approaches offer many
advantages in measuring different aspects of science achievement.
Unlike pencil-and-paper multiple choice tests, performance-based
assessments allow researchers to observe and assess students as
they perform process skills such as observing, classifying, measuring,
inferring, predicting, experimenting and formulating models (Smith,
Ryan, & Kuhs, 1993, p. 6). They also permit the measurement of
affective dimensions such as attitudes about science. This form of
assessment was supported by the National Association on Educational
Progress (NAEP, 1987) through a study it conducted by involving
hands-on activities to assess achievement in math and science. The
findings of the study indicated that the students responded well to
the tasks used in the assessment and that the teachers indicated a
willingness to engage in further hands-on testing and instructional
techniques. Included in the report was the statement that
"conducting hands-on assessment is feasible and extremely
worthwhile" (p. 9).

There are limitations to the method that must be addressed in
the research design in order to increase the level of generalizability
of the results. Shavelson et al. (1991) reported the phenomenon of
fifth grade students performing well on one task and not on another. Another concern about performance-based assessment deals with the concrete nature of the tasks. Baron (1991) expresses the concern that as the extent to which the tasks rely on the construction of new knowledge and new ways of thinking increases, the level of a student's previous experiences will prejudice the resultant data. Shavelson et al. (1991) also refer to this phenomenon in terms of a student's level of disadvantage coming into the assessment.

Smith, Ryan, & Kuhs (1993) list some other concerns about this method of assessment. Regarding reliability, they point out that types of tasks to be used must be standardized with regard to length, scoring scales, procedures, and materials provided and that the raters of the testing should all receive the same training and experiences in preparation for their service. Blank and Selden (1989) and Aschbacher (1991) present the issue of cost. The former mention three areas for performance-based assessment in which the costs will be incurred (designing activities, purchasing equipment, and administering and scoring) and the latter points to the testimony of several states where it was used claim that it is well worth the cost.
Backdrop for the Study

This study took place against the backdrop of a partnership between the Ohio State University - Mansfield, the Ohio Department of Natural Resources, and the Science and Math Network of Central Ohio. A summer institute was held during which partnership teams were formed consisting of teachers from rural school districts and professionals from the ODNR. Generally stated, the goal of this institute was to improve the quality of elementary science education in rural schools located in central Ohio. There is much support for the usefulness and effectiveness of partnerships for addressing this overall goal (Hutto, et al., 1990; Jinks & Lord, 1990; Johns, 1984; Ostlund, 1986; Otterbourg, 1990; Schoenberger & Russell, 1986; Tilgner, 1990; Weir, 1988; Weiss, 1987).

One of the ways in which the partnership program sought to address the quality of rural elementary science education was through the instruction of the Institute participants in science process skills and inquiry-based instructional strategies (see Appendix H). This study examined the effects this kind of instruction had on the elementary students' frequency and appropriateness of use of the science process skills.
Summary and Research Hypotheses

Many and varied reform movements have been put forth as attempts to address concerns regarding the instruction of science education. Public expectations are that students will be well grounded in a body of scientific knowledge. Chief among public expectations is a student's ability to solve problems and contribute to the mechanisms which drive society to a globally competitive level. Science reform has attempted to meet these expectations through the use of concept-based and process-based strategies.

Literature and research cited above demonstrate the high desirability of students' use of science process skills in elementary school classrooms and that acquisition of science process skills is most readily fostered by inquiry-based instructional strategies. Also, the importance of using hands-on assessment techniques to determine student achievement relative to the use of science process skills was emphasized. The transfer (Woolfolk, 1990) of the benefits of activity-based science programs to other areas of the curriculum has been attested to by several of the authors cited earlier in this review. The science education community has continued to endorse the instruction of science process skills. Students must develop the ability to think more critically and be better prepared to enter and contribute to society.
The combination of the preceding literature relating to science curriculum reform, science process skills, inquiry-based instruction, and partnerships supports the notion that they could be used in concert with one another to facilitate more effective instruction in science. If this is the case, the resulting science student performance should reflect the effects of this improved instruction. In this case, the literature relating to transfer of skills (Pauker & Roy, 1993; Woolfolk, 1990) suggests that it is likely that the science process skills taught in the science classroom would be demonstrated in science-related problem solving situations and that it is possible (however, somewhat less likely) that the same skills might be demonstrated in problem-solving situations which are generic.

Because this study was unique in both methodology and instrumentation used, it was important to limit the focus of the study and to examine variables which the literature suggested would contribute to improved practice in science. Throughout this chapter, problem solving has been emphasized. Science process skills have been identified as important tools for children to use in solving problems. Inquiry-based instructional strategies have been identified as preferable in the instruction of science. Frequency and appropriateness of use of the science process skills were chosen as variables for study because increases in both of these would suggest
improved problem-solving skills in students. Based upon the literature cited above, the first four hypotheses below are stated in directional terms. The last hypothesis, related to group cooperation, is posited because science reformers endorse the fact that in our society, science is done in groups (Aldridge, 1991; Rutherford & Ahlgren, 1992). That is, scientists do not typically work in isolation, rather, they collaboratively seek the input, advice, and wisdom of others. Inquiry-based instruction may have an effect on the level of student interaction. Were this to be the case, level of group cooperation might be considered a viable explanation for the results of the study.

Based upon this review of the literature, the following are the experimental hypotheses for this study.

H1 Students who are taught science using inquiry-based instructional strategies will demonstrate science process skills with greater frequency while solving science-related problems than will similar students who are not taught science with inquiry-based instructional strategies;

H2 Students who are taught science using inquiry-based instructional strategies will demonstrate process skills with greater frequency while solving generic problems than will similar students who are not taught science with inquiry-
based instructional strategies;

H3 Students who are taught science using inquiry-based instructional strategies will demonstrate science process skills with greater appropriateness while solving science-related problems than will similar students who are not taught science with inquiry-based instructional strategies; and

H4 Students who are taught science using inquiry-based instructional strategies will demonstrate process skills with greater appropriateness while solving generic problems than will similar students who are not taught science with inquiry-based instructional strategies.

H5 Students who are taught science using inquiry-based instructional strategies will demonstrate group cooperation differently than will similar students who are not taught science with inquiry-based instructional strategies.
CHAPTER 3
RESEARCH PROCEDURES

Introduction

This chapter will provide a description of the research procedures used in this study to address the research questions submitted in Chapter 1. The research procedures are presented under four major headings: a) sample; b) research design; c) instrumentation; and d) data analysis procedures. Reference is made throughout this chapter to a pilot study conducted to establish the usefulness of various aspects of the instrumentation and procedures. A report of this pilot study is found in Appendix G.

Sample

The population for this study was drawn from fifth grade students of the teachers who were involved in the training program described in Appendix H. The accessible population was fifth grade students in rural elementary schools whose teachers participated in a training institute during the 1994-1995 school year, whose parent(s) signed the "Parental Permission" release form (see Appendix A).
For the purposes of this study, a list of all the teachers of fifth grade classes involved in the partnership was obtained from the program's summer institute materials. Of the classes of the eight teachers (representing a total of 161 students) who were in the accessible population, two classes (representing a total of 48 students) were randomly selected using a table of random numbers. By randomly selecting two teachers from the list of fifth grade teachers to be in the experimental group, 29.81 percent of the total accessible population of students for the treatment group were subjects in the study. Since the teachers were randomly chosen and, as indicated below, the students had been randomly assigned to the classes, the sample was of sufficient size to generalize to the effects of inquiry-based instruction to the fifth grade classes whose teachers were involved in the training program (Gay, 1992, p. 137). Generalization of the results of this study beyond this population would not be valid.

Two more teachers were selected via the process described below to provide the control condition for the study. By dividing each of the teacher's classes into four groups of students, a total of 16 groups of students resulted (see Fig. 3.1). By choosing similar teachers who were not involved with the inquiry training, an equal number of groups for both treatment and control (eight for each) was
obtained. Based upon four classes with four groups per class (each group containing four or five students), and gathering data on every group performing all four tasks, this sample size of 16 groups (treatment and control combined) produced a total of 64 units for analysis (32 units of science-related activities and 32 units of generic activities).

Based upon the indications from the pilot study (Appendix G) it was determined that groups should not exceed five students in size. Groups which were larger than this in the pilot study tended to have students who showed hesitancy in asserting themselves in the group

<table>
<thead>
<tr>
<th>Experimental*</th>
<th>Control**</th>
</tr>
</thead>
<tbody>
<tr>
<td>E 1</td>
<td>E 2</td>
</tr>
<tr>
<td>Group 1</td>
<td>Group 1</td>
</tr>
<tr>
<td>Group 2</td>
<td>Group 2</td>
</tr>
<tr>
<td>Group 3</td>
<td>Group 3</td>
</tr>
<tr>
<td>Group 4</td>
<td>Group 4</td>
</tr>
</tbody>
</table>

* Two classes randomly selected from the accessible population of fifth grade teachers involved in the partnership program.

** Two classes selected at the suggestion of the building principals based upon similarity of teaching style and sameness of grade level.

Figure 3.1: Derivation of Sample
setting. That is, some were reluctant to participate in the group process while other members of the groups tended to dominate both the conversation and manipulation of the items with which the students were to solve the problems. For this reason, the maximum group size was limited to five students. The pilot study also revealed that how well the members of the group cooperated with each other could contribute to the group's ability to solve the problems and as such, could be a significant limitation in this study if it was not built into the design. For this reason, a dependent variable of overall group cooperation was included in the design. This variable was used only as a monitor of the ability of the group to work together when trying to solve the problems and did not address any of the experimental hypotheses.

Selection of Control Teachers

In order to obtain classrooms which were as similar as possible to the experimental classrooms to serve as the control group, it was first necessary to secure the cooperation of the experimental group teachers and to become informed about their teaching habits and experiences. Based upon the information gathered from the experimental teachers, the control teachers were chosen in the following way.
The principals of the school where the experimental group teachers work were contacted via a letter of introduction (Appendix B) and asked for permission to conduct research in the school building. In one case, the principal agreed without hesitation; in the other, the principal felt the superintendent should be consulted. The superintendent was contacted by the researcher, given a copy of the letter of introduction and a verbal explanation of the purpose and scope of the study. The superintendent approved the school's involvement in the study. After the principal was notified by the superintendent that involvement was approved and with the permission of the principal, the researcher contacted the experimental teacher in the building from the partnership program arranged and conducted an interview. The purpose of the interview was to gather information from the experimental teacher related to teaching experiences and styles (see Appendix C for interview questions). The information gathered in the interview related to teaching experiences and styles was verbally summarized for the principal. The principal was then asked to identify a fifth grade teacher either in his/her building or in the nearest elementary school in the same district who was not practicing inquiry-based instruction of science and whose teaching style was otherwise, in his/her estimation, similar to that of the teacher in the experimental group.
Upon the approval of the principal in the building where the recommended control group teachers taught, the recommended teachers were contacted and, in both cases, agreed to be involved. In one school system, the control teacher was in the same building as the experimental teacher but had not received the training in inquiry-based instruction. In the other school system, the control teacher was in a neighboring building to the experimental teacher. A comparison of the teacher characteristics can be found later in this chapter.

The Teacher Interview Process

The teachers selected for both control and experimental groups were contacted personally by the researcher and interviewed two weeks prior to data gathering. Following the data gathering, the teachers were interviewed a second time to gather information about their familiarity with the nature and use of science process skills or practices which approximate them. In all cases, the interviews were audiotaped and transcribed by the researcher. The first interview was transcribed prior to the data gathering and the second was transcribed following the gathering of the data but prior to the data analysis.
As discussed in Chapter 1, the fact that the teachers in the experimental group had self-selected for involvement in the summer training institute constituted a threat to the internal validity of the study. The interview data were gathered in an attempt to ascertain the level of this threat and attest to the similarity of the teachers involved in the study (regarding teaching experiences, style, and curriculum) to each other. The information gathered is summarized later in this chapter.

Description of the Setting

The study took place in four elementary schools in central Ohio. Central Ohio is defined as an 11-county area in and around the capital city of Columbus (A. Thompkins, Division of Computer Services and Statistical Reports, Ohio DOE, 1991, in Cullen, 1992). These 11 counties contain 89 school districts which account for about 13.5% of the state's total student population (Ohio Department of Education, 1994). The districts involved in this study were in rural areas of the geographic region described above. Two of the schools are located to the west and two are located to the north of Columbus.
Description of the Participants

All teachers involved in the study were fifth grade teachers and were the regular teachers of the students who were videotaped in the problem-solving sessions. The teachers worked in three schools. In Rural District 1 (RD1) the experimental and control teachers worked in separate buildings but had a common curriculum. In Rural District 2 (RD2) the experimental and control teachers were in the same building but were in self-contained classrooms. In each case (RD1 and RD2), one of the teachers had attended the summer training institute, the other had not and was not involved in any way with the instructional activities of the first. There were no "nontraditional" or "alternative" instructional approaches to the curriculum in place in any of the schools involved in the study (eg. literature-based, whole-language, etc.), except the inquiry-based instruction related to elementary science education in the case of the experimental classrooms. Table 3.1 presents a comparison of the experiences and assignments of the teachers involved in the study.

This information was gathered from the teachers via the first interview mentioned above. An analysis of the answers the teachers gave in the interviews relating to commonalities and differences between districts, schools, teachers, and students is reported in Chapter 4.

55
Table 3.1: Description of Teacher Participants

Research Design

There were three dependent variables in this study. "Appropriateness of use" of the science process skills, and "frequency of use" of the science process skills were the dependent variables which addressed the experimental hypotheses. Group cooperation was added as a dependent variable because the pilot study indicated it as a possible rival hypothesis.

Within each room there were four problem situations set up, two were science-related and two were generic. In groups of four
students each, the students randomly rotated through the centers until each group had experienced all four problem situations. Four video cameras were set up in the center of the room, one for recording each of the activities. At the end of the data gathering session for one class, each camera had recorded four groups doing the activity assigned to that camera. In all, there were 16 units for analysis gathered in each class. The data gathering was done the same way as indicated above for all four classes (two experimental, two control).

The data were collected by trained, blind raters as they observed video tapes of groups of students performing problem-solving tasks. For instance, as a tape was playing, the rater would watch for and record (on the observation form) each instance where a science process skill was used. On the observation form (Appendix D) there were rows for all ten skills, each divided into five two-minute segments. At the end of the tape (no longer than 10 minutes) the rater would add all the tallies for each skill and record the number in the "total" column for each skill. She/he would then reflect upon how appropriately each skill was used and would record this judgement on the Likert scale provided opposite each skill. In this way, each skill received a frequency and appropriateness score. The final task of the rater was to again reflect on the whole session.
observed and to rate the overall level of group cooperation on the Likert scale provided at the bottom of the observation form.

Problem Situations Employed

There were four problem-solving activities used in this study. Two were science related and required the students to work within the bounds of science concepts which are commonly found in fifth grade science curricula (AIMS, 1987, 1993). The other two problem-solving activities were generic. That is, their solutions did not presuppose mastery or facility with any particular curriculum content or conceptual frameworks.

The two science-related activities were obtained from the AIMS Education Foundation materials (1987, 1993) and were modified slightly for the purpose of facilitating the students' use of science process skills while working on the tasks without the aid of a teacher. The modifications were necessary for the activities because they were originally intended to be performed as teacher-directed (specifically, guided-discovery), laboratory-type activities. In order to use these in an inquiry-based setting, it was necessary to provide written directions for the students (Appendix E). These directions presented the problem and any conditions under which the students were to work in solving it. They also stated the amount of time
available to solve the problem and listed the materials students had been given to use for their attempts at solving the problems. The observations made in the pilot study revealed that although the groups were allowed 15 minutes to complete the tasks, they finished the tasks in eight to ten minutes in all cases. Based upon these observations, the time allowed for each group to work on each task for this study was ten minutes. The generic activities used in this study are addressed separately. The modified AIMS activities used in this study to assess students' ability to use science process skills follow.

All Wound Up (AIMS, 1993). The purpose of this activity was for students to use the materials provided them to construct a "tractor" that would travel at least four feet along a level surface under its own power. The science topic area addressed by this activity was friction. The science process skills listed in the AIMS literature which this activity addressed are; observing, measuring, making and testing hypotheses, data recording, interpreting data, and applying and generalizing results.

Fill Up (AIMS, 1987). The purpose of this activity was for students to discover how much space exists between particles of gravel. The science topic area addressed by this activity was space in matter. The science process skills listed in the AIMS literature
which this activity addressed are; measuring, estimating, data recording, interpreting data, and applying and generalizing results.

There were two activities used in this study that were not related to any content area. The first problem was modified from a problem presented in the *New Games Book* (Project Adventure, 1982). The second was taken directly from Beyer (1991). The problems required validation by a panel of science educators to verify that the science process skills represented in the science-related problems were applicable to the solutions of the generic problem situations as well. To accomplish this, three science educators (one high school biology teacher, one middle school science teacher, and one elementary teacher) were independently given the definitions of the science process skills addressed by the AIMS materials listed above, the goals of the study, and the generic problem situations. They were then asked to evaluate the generic activities based upon whether or not they believed students would have opportunities to demonstrate the specified science process skills while attempting to solve the problems. All members of the panel agreed independently that the problems stated below provide opportunity for students to demonstrate the science process skills of observing, making and testing hypotheses, recording data, interpreting data, and applying and generalizing results as these
skills are defined by Padilla (1986) for the AIMS problems. The problem situations the students were given follow.

**The Foxes and the Chickens** (Project Adventure, 1982). Five foxes and five chickens are traveling as a single group through the woods. There are five of each because if the foxes were ever to outnumber the chickens, they would eat the chickens immediately! The group happens upon a river. This is a problem since none of them can swim. Luckily, there is a small canoe on the shore but only three of them can be in it at any one time.

Your task is to come up with a way that the whole group can get across the river (a few at a time), using only the canoe for transportation. Remember, the foxes may never outnumber the chickens on either shore or in the canoe. (Yes, the chickens and the foxes can both paddle the canoe.)

**The Desert Problem** (Beyer, 1991). You are the only survivors of a plane crash in the Arizona desert. Just as engine trouble developed, you were told by the pilot and co-pilot (both killed in the crash) that the nearest town is 60 miles to the south. The area around you is flat and apparently barren except for some small (and pointy) desert plants. You know that the temperature where you are reaches 110 °F in the mid-afternoon and goes down to 35°F at night. Everyone is dressed in lightweight clothes and when you checked
your pockets, you found handkerchiefs, small change, matches, lipstick, and credit cards.

You have all agreed to stay together for survival. Your task is to rank the eleven items on your group's sheet from the most important to the least important using 1 for the most important and 11 for the least important. Make sure the whole group agrees on the rankings before you write them down. Your lives depend on it!

The items listed on the ranking sheet are: a cosmetic mirror, a pair of sunglasses per person, a flashlight (4 battery size), a hunting knife, a magnetic compass, one quart of water per person, one top coat per person, a parachute (red and white), a 45 cal. pistol (loaded), the pilot's map of the area, and a bottle of Tylenol.

Preparation for Data Gathering

Following the identification of the participating classrooms and initial interviews of the teachers, the researcher visited the classrooms to briefly meet the students, ascertain how the activities could best be set up and videotaped, and acclimate the students to the presence of both the researcher and the equipment used in the recording of the observations. To do this, the researcher spent about two hours in each classroom.
The time the researcher was in the classroom was primarily for the purpose of becoming acquainted with the students in an informal way via short conversations and response to their questions about the researcher's presence. During this time, the researcher also drew diagrams of the classroom, including the location of any immovable furniture/fixtures, estimated measurements of the open spaces in the room for group placement, and the location/availability of electrical outlets from which to operate the cameras. The camera placements were then determined and arrangements for gathering the data were made with the cooperating teacher.

In all four cases, after the researcher had been in the classroom approximately 90 minutes, the cooperating teachers formally introduced the researcher and allowed the researcher to address the class as a whole. The researcher first informed the students that they would be participating in some problem-solving activities in about a week that would be videotaped. They were further informed that their participation was optional and that in order to participate, they would have to return the parental permission form (Appendix E), signed by a parent, prior to the videotaping. The forms were passed out and in all cases, the teachers agreed to collect them as they were returned over the following week.
The students were then shown the video camera, microphone, and tripod setup to be used in the data gathering. The students were asked if the equipment was new to them. In all cases, all students had had experience with the equipment. When asked if there were any who had never been videotaped, in all cases, there were none. The students were then asked if they had any questions. There was one question that was asked in all four classes. The students wanted to know if they would be allowed to see the tapes. The researcher answered that there would not be that opportunity. Neither teachers nor students were specifically informed as to the purpose of the investigation being conducted (that is, their use of science process skills during the problem-solving sessions) until after the data had been collected. Details of the data collection process are discussed later in this chapter.

On the day of videotaping, four camera and microphone sets were used in the classroom, one for each problem situation, to videotape groups simultaneously as they rotated through all four problem solving tasks.

Instrumentation

The development and modification of instruments were critical to this study. A rating form for videotapes which included all science
process skills to be evaluated and measures for overall group cooperation, appropriateness of use of the skills, and frequency of use of the skills was designed using information about the skills referred to in Chapter II (AAAS, 1965; AIMS, 1987, 1993; Padilla, 1986). The format of the rating form (Appendix D) for this study followed suggestions by Bijou, et al. (1969) and the example of Flanders' Interaction Analysis (in Gay, 1996) for the observation of behaviors over a period of time. On the rating form, each science process skill was accompanied by a short description (Padilla, 1986), a space for the rater to record frequency of its use within five two minute intervals (0-10 minutes), five point Likert scales for rating the appropriateness of use for each science process skill, and another five point Likert scale for rating overall quality of group cooperation on the entire task. The resulting Science Process Skills Observation Instrument (Instrument) included the ten science process skills indicated by the AIMS (1987, 1993) materials as specifically applying to the problem situations used (Appendix E). These skills were; observing, measuring, predicting, communicating, formulating hypotheses, experimenting, controlling variables, recording data, interpreting data, and applying and generalizing results.

As was mentioned in Chapter II, the science process skills have been defined by several different authors or agencies (AAAS, 1965;
AIMS, 1987, 1993; Padilla, 1986) using different words. However, the meanings of the individual skills have remained essentially the same throughout the literature. In that the content of the Instrument comes directly from the literature related to science process skills and the terms and definitions were commonly agreed upon in the literature, the instrument can be said to be valid with regard to content. Construct validity of the instrument was established during the rater training process described below.

Rater Selection and Training

In order to avoid experimenter bias in the interpretation of the videotaped situations, it was necessary to select and train individuals who would then be qualified to watch the tapes and record their observations. Selection of the raters for the study was done via consideration of the background of several doctoral students in the College of Education at the Ohio State University. The four doctoral students chosen (asked) to be raters had either some teaching experience in science or were currently engaged in graduate study in science education. These people were approached individually by the researcher and, after receiving a brief explanation of the study and the role of the rater, were asked to participate in the study. All four agreed to participate and time for the rater training was established.
Kieren & Munro (1985) state that it is important that measurement be valid (correspond to the observed events), and reliable (that random error noise be reduced). Wilson, et al. (1988) state that it is not reasonable to assume that without training, all raters will interpret descriptions and instructions in a uniform manner. In other words, raters must be trained so that they arrive independently at a consensus of opinion on high-inference issues and so that they have a uniform and correct understanding of their tasks. Thus, training is indicated in situations where the variables to be observed involve varying degrees of judgement by the raters. The purpose of training the raters was to arrive at a high level (greater than 90% agreement with the researcher) of interrater reliability (Gay, 1996). If the raters are reliable about performing their tasks throughout the data collection, then the data gathered can be considered to be reliable. For this study, both the group cooperation and appropriateness of use dependent variables were considered high-inference (Wilson et al., 1988). It was also important in the case of the frequencies, that all raters demonstrate a common understanding of the meanings and characteristics of each of the science process skills and that this understanding was consistent with the definitions and characteristics of the skills portrayed in the related literature. Wilson et al. also suggest that when training raters,
copies of the rating instrument to be used should be distributed and explained relative to how raters are to use it and how the information contributes to the study. The above advice was followed for the training of the raters for this study.

Rater training for this study occurred during two sessions (Appendix F). At the first meeting, the raters were given a short introduction to the goals of the study and their part in it. They were then introduced to the concept of science process skills, and were given definitions and examples of each skill represented on the observation instrument. Following this, the raters were introduced to the format of the observation instrument and viewed several tapes from the pilot study as a group. Although the format of the study was outlined to the raters, they remained blind with respect to knowledge about whether they were rating control or experimental groups and science or generic problem situations. The tapes viewed during the rater training included the problem solving situations used in the study. This was done to familiarize the raters with the problem solving situations and to allow them to discuss possible interpretations of their observations. It was anticipated by the researcher that this dialogue might contribute to higher interrater reliability.
During the initial series of videotape observations, the instrument was viewed by all four raters simultaneously and the tape was stopped whenever a rater had a question or asked for clarification about marking the instrument or about a science process skill. The raters viewed five problem solving sessions as practice. When it seemed that the raters had developed a consistent and accurate understanding of the rating process (a judgment made by the researcher based upon results on the raters' Instruments and the quality of dialogue among the raters when discussing their observations) they were checked for interrater reliability in the following way.

The raters were asked to simultaneously view a tape of a problem solving situation and to use the instrument to rate the session. The researcher rated the tape along with the raters and the ratings of the researcher were considered the standard against which reliability of the raters was to be measured. The process for this viewing was the same as that which the raters used during rating of the tapes for the study. That is, the raters and the researcher first viewed the entire session without stopping the tape, recording any observations, or discussing anything in order to familiarize themselves with the session. The video tape was then rewound to the beginning of the activity to be rated. The video tape was started
a second time simultaneously with an audio tape on which the sound of a chime was recorded at two minute intervals. The raters then recorded their frequency observations in the two minute interval cells on the instrument. If there were no skills observed in a given interval, the cell was left empty. Empty cells were treated as frequencies of zero. At the end of the problem solving situation, the tape was turned off and the raters recorded their evaluations of the appropriateness of use for each skill and overall group cooperation. At the end of the training session, the researcher collected the rating sheets and analyzed rater responses to determine interrater reliability. The procedure and results of the investigation into interrater reliability are reported in the next section.

**Interrater Reliability**

It was necessary to ensure high rates of interrater reliability on the measurement of all three dependent variables in order to establish reliability of the data. It was important to know whether the students were intentionally using the appropriate skills at times when use of the skill was warranted or, whether they were using skills in unfocused or haphazard ways. Since the dependent variables for the study were highly subjective, reliability between raters had to be established.
Interrater reliability was initially addressed during the rater training sessions using a videotape from the pilot study and was reevaluated near the middle and near the end of the rating process by inserting problem situations from this study which were common to all four raters, then comparing those ratings with the researcher's own ratings of the same tapes.

The method used for analysis of interrater reliability was percent agreement with an expert rater (Gay, 1996). The expert rater for this study was the researcher. Regarding the acceptable level of reliability, Gay (1996) states, "...a coefficient over .90 would be acceptable for any test." It was mentioned previously that the instrument developed for this study was patterned after Flanders' Interaction Analysis which reports the acceptable level of interrater reliability to be above .85 (Gay, 1996). For this study, the acceptable level of interrater reliability for consistancy in reporting observations of the video tapes was 90% agreement with the expert rater (researcher).

Percent agreement was calculated by comparing the ratings of the expert and the individual raters. For the frequency of use dependent variable, the frequencies marked in each of the two-minute cells (described previously) were compared. If they were the same, a score of one was recorded in the corresponding cell on a
blank observation form. If they were not the same, a score of zero was recorded. These marks were added together across a skill. The resulting sum was recorded in the "total" column. A perfect match would have ones in all five cells, therefore, the perfect score (100 percent agreement) would have been five. The sum of the markings across the skill was divided by five and recorded as percent agreement. For example, if the sum was four, that number would have been divided by five. This calculation would have yielded a score of eight tenths. This decimal would then have been converted to a percentage by multiplying by 100.

The process was nearly the same for the other dependent variables with the only difference being that the appropriateness of use and group cooperation scores were not reported in two-minute intervals, but were reported over the full ten-minute period. The result of this was that the appropriateness and cooperation ratings which were compared to the expert rater's scores were not sums. They were taken directly from the Likert scales for each skill (in the case of appropriateness) or the overall session (cooperation).

Table 3.2 reports the results of the first measurement of interrater reliability as calculated via the above explanation. Note that all percent agreement values were above the preset level of acceptance (90%). It was based upon these values that the training
was ended and the raters were given the tapes from the study to take for observation and evaluation.

The assignment of the specific sessions to the raters was done randomly by assigning each session a number and using a random number table to match raters and sessions. The sessions were then recorded onto tapes in a random order with the common sessions (used for measurement of interrater reliability) occurring in the seventh position and the 14th position. Since the common sessions were randomly chosen from the available 64 units for analysis (sessions), this left 62 sessions to be divided between the raters. Consequently, with the addition of the two common tapes for measurement of interrater reliability, two of the raters received seventeen tapes to rate and the other two received sixteen. In view of the very high overall reliability of the raters (ranging from 93.34% to 100.00% and averaging 97.50%, see table 3.2), and the consistency of the reliability of the raters across all three dependent variables ( Appropriateness, 95%; Frequency, 97.5%; and Cooperation, 100%) duplication of rating for each tape was not indicated.

Table 3.3 presents the percent agreement with the researcher for each of the dependent variables for the second measurement of interrater reliability. This measurement was taken about the middle
Table 3.2: Interrater Reliability Percent Agreement With Expert Rater: First Measurement

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>Ave. percent agreement by dependent var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approp.</td>
<td>100.00</td>
<td>100.00</td>
<td>90.00</td>
<td>90.00</td>
<td>95.00</td>
</tr>
<tr>
<td>Frequency</td>
<td>96.00</td>
<td>100.00</td>
<td>100.00</td>
<td>94.00</td>
<td>97.50</td>
</tr>
<tr>
<td>Cooperation</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Ave. percent agreement by rater over all measurements</td>
<td>98.67</td>
<td>100.00</td>
<td>93.34</td>
<td>94.67</td>
<td>97.50</td>
</tr>
</tbody>
</table>

Note: This measurement was taken at the end of the rater training process.

of the rating process. The reader should note that although rater 4 dropped to 80.00% for the appropriateness of use in the second measurement, this rater's reliability rebounded to 100.00% in the third measurement (Table 3.4) and averaged 90.00% over all measurements of reliability for appropriateness of use of the science process skills. Again, the reliability of the raters (both individually and as a group) for all dependent variables remained above the acceptable level of agreement (ranging from 92.67% to 100.00% and averaging 96.83%). Table 3.3 presents the percent agreement with
Table 3.3: Interrater Reliability Percent Agreement With Expert Rater: Second Measurement

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>Ave. percent agreement by dependent var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approp.</td>
<td>90.00</td>
<td>100.00</td>
<td>100.00</td>
<td>80.00</td>
<td>92.50</td>
</tr>
<tr>
<td>Frequency</td>
<td>100.00</td>
<td>100.00</td>
<td>94.00</td>
<td>98.00</td>
<td>98.00</td>
</tr>
<tr>
<td>Cooperation</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Ave. percent agreement by rater over all measurements

96.67 100.00 98.00 92.67 96.83

Note: This measurement was taken near the middle of the rating process.

Table 3.3: Interrater Reliability Percent Agreement With Expert Rater: Second Measurement

the researcher for each of the dependent variables for the second measurement of interrater reliability. This measurement was taken near the end of the rating process. The reader should note that although rater 1 dropped to 80.00% for the appropriateness of use in the third measurement, this rater's reliability in the first and second measurements (Tables 3.2 and 3.3) was 100.00% and 90.00% respectively and averaged 90.00% over all measurements of reliability for appropriateness of use of the science process skills. Table 3.5 presents the average percent agreement with the
Table 3.4: Interrater Reliability Percent Agreement With Expert Rater: Final Measurement

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>Ave. percent agreement by dependent var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approp.</td>
<td>80.00</td>
<td>90.00</td>
<td>90.00</td>
<td>100.00</td>
<td>90.00</td>
</tr>
<tr>
<td>Frequency</td>
<td>100.00</td>
<td>96.00</td>
<td>100.00</td>
<td>98.00</td>
<td>98.50</td>
</tr>
<tr>
<td>Cooperation</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Ave. percent agreement by rater over all measurements</td>
<td>96.22</td>
<td>98.45</td>
<td>97.11</td>
<td>95.56</td>
<td>96.84</td>
</tr>
</tbody>
</table>

Note: This measurement was taken near the end of the rating process.

Table 3.5: Interrater Reliability Percent Agreement With Expert Rater: Averages by Rater and Overall

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>Ave. percent agreement by dependent var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approp. Ave.</td>
<td>90.00</td>
<td>96.67</td>
<td>93.33</td>
<td>90.00</td>
<td>92.50</td>
</tr>
<tr>
<td>Freq. Ave.</td>
<td>98.67</td>
<td>98.67</td>
<td>98.00</td>
<td>96.67</td>
<td>98.00</td>
</tr>
<tr>
<td>Coop. Ave.</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Ave. percent agreement by rater over all measurements</td>
<td>96.22</td>
<td>98.45</td>
<td>97.11</td>
<td>95.56</td>
<td>96.84</td>
</tr>
</tbody>
</table>
researcher for each of the dependent variables over all measurements of interrater reliability. The reader should note that in all cases, the reliability of the raters remained at or above the acceptable level of percent agreement. Based upon this and the above information, interrater reliability is claimed to be sufficient to claim reliability of the data presented in Chapter 5.

Data Collection Procedure

Since the researcher did not know the students well enough to make the determinations about which students should be grouped together and which should not, the cooperating teachers were asked to group the students in the way the teachers were most accustomed to. All four teachers voiced their preference to group students heterogeneously, being especially considerate of mixing gender and level of academic achievement. All teachers agreed to set the groups according to the above criteria on the day of the data collection. The steps of the process follow.

Students were given permission forms (on the day of the introductory visit) to be signed by a parent and returned to the cooperating teacher. Only those students who return signed forms were taped during the data-gathering sessions. In only one class were there students who did not return the form. This class had two
students who did not return the form. The reason given by both students was that they had forgotten the forms at home. The teacher had attempted to contact the parents for permission but was unable to. These two students were sent to the library during the videotaping and were supervised by another teacher whose class was in the library at the time.

The problem situations were set up at the stations (1-4) in the classroom as follows:

- Station 1  All Wound Up
- Station 2  Foxes and Chickens
- Station 3  Fill Up
- Station 4  The Desert Problem

The groups of four or five students were randomly assigned to starting positions in the room by the researcher (via use of a random number table). Groups traveled from station to station in a random order, again assigned by the researcher using a random number table in an effort to minimize any interference that a specific sequencing of the problem situations might have introduced. The session was finished when all groups had completed all tasks.

Ten minutes were used to introduce the procedure and to ensure that all activities and recording equipment were prepared. Since the researcher was the only one to operate the cameras, the
starting and stopping times were staggered by about 20 seconds to allow the researcher to turn on/off the cameras.

Each group was allowed 10 minutes for each activity. For each activity, the researcher activated the camera and instructed the group being recorded by that camera to begin. All groups working on all problems were videotaped. There was no researcher or teacher involvement in the problem-solving process, however, the researcher and teacher both circulated about the room to monitor and facilitate the process. Clarification of tasks, provision of materials, dealing with classroom management issues, etc. are examples of the facilitative behaviors in which the teacher and researcher engaged which did not impact the problem-solving process in the group.

At the end of the ten minute time period, the groups were instructed to stop, clean up, and move to the next activity. A seven minute time was given for the resetting of the activities, movement of the groups, clarification of any questions, and checking of the cameras/microphones. The activities proceeded in turn as described until all groups had been video-taped attempting all activities.
Data Analysis

This quasi-experimental (Gay, 1992) study contained three variables of interest. In order to arrange the data for analysis, each dependent variable was addressed separately. The dependent variable of group cooperation (Figure 3.2) was addressed using a one between (treatment/comparison), one within (science/generic) experimental design. The appropriateness of use dependent variable (Figure 3.3) was a one between (treatment/ comparison), two within (process skills nested within science/generic) experimental design. The dependent variable, frequency of use (Figure 3.4) was a one between (treatment/ comparison), two within (process skills nested within science/ generic) experimental design.

The unit of analysis for this study was each group performing each task. The analysis of video taped problem solving sessions by trained raters was used to gather data related to the appropriateness and frequency of the students use of science process skills and level of cooperation within the groups. All sessions from the problem solving situations which are generic and all sessions of the science-related problem solving situations were viewed and scored by raters trained in the area of science process skills (training described above). The scores were the measures of appropriateness and
frequency to be analyzed. The specific analyses of the data were conducted using repeated measures analyses of variance (ANOVA).

Summary

This study attempted to examine the effects the use of inquiry-based elementary science instruction had upon students' use of science process skills while solving both science-related and generic problem situations. Of particular interest in this study were the frequency and appropriateness of the students' use of the skills and the level of group cooperation within the groups as they performed the tasks.

Video technology, trained raters, and both science-related and generic problem solving situations were used in this study. Comparisons were made of the frequency and appropriateness of use of selected science process skills and of levels of group cooperation between classes whose teachers had been trained in the nature and use of both science process skills and inquiry-based instructional strategies and similar classes where the teachers had not received this training.
CHAPTER 4

RESULTS

Introduction

This study examined the effects of instruction in science process skills on students’ ability to solve problems. The frequency and appropriateness with which students used the skills in different kinds of problem-solving situations were examined. Of the science process skills reviewed earlier, the ten skills (identified by the writers of the problem situations employed, and verified by a panel of science teachers as applying to this study) were; communicating, measuring, predicting, stating hypotheses, making observations, recording data, interpreting data, controlling variables, applying and generalizing results, and experimenting.

Experimental teachers had agreed to teach their classes using an inquiry-based approach and to instruct their students in the use of science process skills when they engage in problem situations. The experimental groups for this study were made from two classes of fifth grade students who were taught using an inquiry-based approach. The initial research questions for the study were:
• With what frequencies do experimental and control group students use science process skills in science-related problem solving situations?

• With what frequencies do experimental and control group students use science process skills in generic problem solving situations?

• How appropriately do experimental and control group students use science process skills in science-related problem solving situations?

• How appropriately do experimental and control group students use science process skills in generic problem solving situations?

• Does the level of cooperation within groups vary between classes where students are taught using inquiry-based instructional strategies and classes where students are not taught using inquiry-based instructional strategies?

The data were gathered through the analysis of video-taped problem solving sessions. The tapes were viewed, scored and reported by trained, blind raters. The frequency-of-use (non-parametric) data were analyzed using a series of chi-square tests for independence. The appropriateness-of-use (parametric) data were analyzed using a repeated measures analysis of variance (ANOVA).
This chapter is a presentation of the results of the data analyses related to the frequency and appropriateness of use of the science process skills by 16 groups (eight experimental and eight control, four groups per class) and of the level of cooperation between students in those groups. The chapter is organized into five sections: (a) statistical hypotheses related to frequency-of-use data, (b) statistical hypotheses related to appropriateness-of-use data, (c) statistical hypothesis related to group cooperation data, (d) ancillary information for discussion, and (e) summary.

Statistical Hypotheses Related to Frequency-Of-Use Data

The statistical hypotheses related to students' frequency of use of the science process skills were:

H₁ Students who are taught science using inquiry-based instructional strategies will demonstrate science process skills with greater frequency while solving science-related problems than will similar students who are not taught science with inquiry-based instructional strategies; and

H₂ Students who are taught science using inquiry-based instructional strategies will demonstrate process skills with greater frequency while solving generic problems than will
similar students who are not taught science with inquiry-based instructional strategies.

In each of the first two experimental hypotheses, exploration of two prerequisite conditions was necessary in order to warrant exploration of the hypotheses. That is, both hypotheses were based upon the expectation that there would be differences in the use of science process skills between experimental and control groups and between science-related and generic problem types. If either of these tests were not statistically significant, neither of the hypotheses could be supported. In the following paragraphs, the results of the tests of these conditions and of the two hypotheses are reported.

Gravetter and Wallnau (1992, p. 523-524) report that the chi-square test for independence is appropriate when applied to nonparametric data to test whether or not there is a relationship between two variables. A chi-square test for independence was used to examine the frequency data from the experimental and control groups. The omnibus chi-square test for independence of the experimental and control groups considering all problems (science-related and generic) revealed a significant value ($X^2(9, \, n=64)=25.83$, $p<.05$). Because the critical value ($X^2(9, \, n=64)=16.92, \, p=.05$) was exceeded, a difference between experimental and control groups was
claimed, and the alternative hypothesis was accepted. These results, combined with the experimental design of the study suggest that this difference may be rooted in the use of inquiry-based instruction with the experimental groups.

The test of the second premise underlying both of the hypotheses related to the frequency-of-use data (that there would be a difference between students' use of science process skills in science versus generic problems) was also performed using a chi-square test for independence. The omnibus test of the science and generic problem situations considering all classes (both experimental and control) resulted in a significant chi-square value \( X^2(9, n=64)=271.88, p<.05 \). Because the critical value \( X^2(9, n=64)=16.92, p=.05 \) was exceeded, the alternative hypothesis was accepted. That is, the way the students performed in these two conditions differed significantly with respect to students' frequency-of-use of the science process skills.

Because the results of the omnibus tests for the underlying premises of the first two research hypotheses revealed significant relationships, further testing was warranted in order to more fully explore the meaning of the results as they related to frequency of student use of the science process skills.
Comparison of the experimental and control groups' use of the science process skills on the science-related questions was conducted using a chi-square test for independence. The resultant value $(X^2(9, n=32)=20.49, p<.05)$ exceeded the critical value $(X^2(9, n=32)=16.92, p=.05)$ the experimental and control groups were found to be different with respect to use of the science process skills on science-related problems.

The final investigation into the relationship of the experimental and control groups as they performed the generic problems (the second research hypothesis) was also conducted using a chi-square test for independence. In this case, the resultant value $(X^2(9, n=32)=13.45, p>.05)$ did not exceed the critical value $(X^2(9, n=32)=16.92, p=.05)$, therefore no difference between the experimental and control classes with regard to the generic problems was found.

Statistical Hypotheses Related to Appropriateness-Of-Use Data

Research hypotheses three and four dealt with the appropriateness of student use of the science process skills addressed in this study. The hypotheses related to the appropriateness of students' use of the science process skills were:

$H_3$ Students who are taught science using inquiry-based
instructional strategies will demonstrate science process skills with greater appropriateness while solving science-related problems than will similar students who are not taught science with inquiry-based instructional strategies; and

\textbf{H}_4 \quad \text{Students who are taught science using inquiry-based instructional strategies will demonstrate science process skills with greater appropriateness while solving generic problems than will similar students who are not taught science with inquiry-based instructional strategies.}

Table 4.1 presents means and standard deviations of the groups' appropriateness-of-use scores as measured and reported by blind, trained raters on a five-point Likert scale. The means depicted in Table 4.1 reveal a difference in appropriateness of use of science process skills between science-related and generic problems with overall skill use having been rated as more appropriate (signified by higher means) in the science-related problem situations. Overall differences between the experimental and control groups were small with slightly higher means for the control groups than for the experimental groups. That is, it appeared that the control groups used the science process skills more appropriately than did the experimental groups. These differences were particularly pronounced for the science-related problems.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Problem Type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Science</td>
<td>Generic</td>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>37.7500</td>
<td>28.6250</td>
<td>33.1875</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(16.5680)</td>
<td>(6.0460)</td>
<td>(11.3070)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=16</td>
<td>n=16</td>
<td>n=32</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>42.8750</td>
<td>25.5000</td>
<td>34.1875</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.7161)</td>
<td>(11.6128)</td>
<td>(11.6645)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=16</td>
<td>n=16</td>
<td>n=32</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>40.3125</td>
<td>27.0625</td>
<td>33.6875</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14.1421)</td>
<td>(8.8294)</td>
<td>(11.4858)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=32</td>
<td>n=32</td>
<td>n=64</td>
<td></td>
</tr>
</tbody>
</table>

Note. Values enclosed in parentheses are standard deviations.

Groups of students are treated as subjects in that measurements were taken of the entire group, not of the individuals in the group. Reported values are omnibus in that they are aggregated over all process skills used by experimental and control and over problem type.

Table 4.1: Means and standard deviations of appropriateness-of-use of science process skills for treatment condition by problem type.
In order to explore the magnitude of the differences between the means, the appropriateness data were analyzed using a two-way, repeated measures analysis of variance which examined the independent variables of science versus generic problem types and experimental versus control treatment conditions. The individual tasks performed by the groups were treated as a repeated measure. The results of this test are summarized in Table 4.2. The ANOVA indicated that the difference observed between science and generic problems was statistically significant ($F(1,63)=13.36, p<.001$) for appropriateness of use of science process skills. There was not a significant difference between the experimental and control groups ($F(1,63)=.839, p>.05$).

As indicated in Table 4.2, the test statistic value related to the treatment condition, inquiry-based instruction versus no inquiry-based instruction ($F(1,63)=.839, p>.05$), indicated that the difference between the two groups was not statistically significant for appropriateness of use of the science process skills. These results fail to support the hypotheses as stated for the appropriateness-of-use dependent variable. Thus, both $H_3$ and $H_4$ were rejected.
### Table 4.2: Repeated measures ANOVA for appropriateness-of-use of the science process skills.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups of Students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental/Control</td>
<td>1</td>
<td>0.839</td>
</tr>
<tr>
<td>G Within Treatments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>14</td>
<td>(186.49)</td>
</tr>
<tr>
<td>Within Groups of Students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science/Generic</td>
<td>1</td>
<td>13.36 *</td>
</tr>
<tr>
<td>Exp/Con x Sci/Gen</td>
<td>1</td>
<td>1.30</td>
</tr>
<tr>
<td>G Within Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>condition Error</td>
<td>14</td>
<td>(105.10)</td>
</tr>
</tbody>
</table>

Note. * indicates statistically significant difference at p< .01. Values enclosed in parentheses represent mean square errors. G= Groups of students. These groups are treated as subjects in that measurements were taken of the entire group, not of the individuals in the group. Exp= Experimental; Con= Control; Sci= Science; Gen= Generic.
Statistical Hypothesis Related to Group Cooperation Data

The final area in which data were collected was that of group cooperation. Because the data were gathered by videotaping students working in groups, the possibility of differences in how well individuals within the groups worked together was anticipated as a possible rival hypothesis. For this reason, the final statistical hypothesis related to group cooperation was included in the study. This hypothesis was:

\[ H_5 \] Students who are taught science using inquiry-based instructional strategies will demonstrate group cooperation to a greater degree than will similar students who are not taught science with inquiry-based instructional strategies.

This hypothesis was not supported by the analysis of the group cooperation data. A two-way analysis of variance performed on these data revealed that there was not a statistically significant overall difference \( (F(7,63)=0.96, p>.05) \). Further analysis of the data with respect to the components of the omnibus \( F \) value revealed that there was neither significant difference between experimental and control groups \( (F(1,56)=.63, p>.05) \) nor science and generic tasks \( (F(3,56)=.47, p>.05) \). The measure of interaction between the groups was also not significant \( (F(3,56)=1.45, p>.05) \). That is, the analysis of experimental versus control, science versus generic, and any
statistical interaction between any or all of these showed that the
groups did not behave differently from each other.

Ancillary Information for Discussion

Even though the results of the tests of the hypotheses for the
study only produced one statistically significant finding, other
information of interest was discovered. It is presented here as
support for the discussion and recommendations in Chapter 5.

A review of the means and standard deviations of the
appropriateness-of-use data, reported in Table 4.3, reveals
significantly more appropriate use of the skills of making
observations, measuring, and controlling variables for generic
problem-types. In contrast, the skills of recording data and
applying and generalizing results were used significantly more
appropriately with the science-related problems.

Other findings of interest regarding the appropriateness-of-use
data related to the individual science process skills. Of particular
interest was information describing which groups used which skills
more appropriately. Table 4.4 depicts the use of the science process
skills in order of magnitude of contribution to the overall chi-square.
The reader should note that the skills of recording data and
measuring contributed strongly to the overall chi-square value.
<table>
<thead>
<tr>
<th>Skill</th>
<th>Control Science</th>
<th>Control Generic</th>
<th>Experimental Science</th>
<th>Experimental Generic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs.</td>
<td>3.7321</td>
<td>4.4238</td>
<td>3.4142</td>
<td>4.5249</td>
</tr>
<tr>
<td></td>
<td>(0.8452)</td>
<td>(0.9797)</td>
<td>(0.7560)</td>
<td>(0.9978)</td>
</tr>
<tr>
<td>Meas.</td>
<td>* 3.1180</td>
<td>4.0917</td>
<td>2.7906</td>
<td>3.7678</td>
</tr>
<tr>
<td></td>
<td>(0.8763)</td>
<td>(0.8425)</td>
<td>(0.8839)</td>
<td>(1.4745)</td>
</tr>
<tr>
<td>Pred.</td>
<td>3.3693</td>
<td>4.1213</td>
<td>3.6956</td>
<td>3.8708</td>
</tr>
<tr>
<td></td>
<td>(1.1161)</td>
<td>(1.5353)</td>
<td>(0.9426)</td>
<td>(1.7728)</td>
</tr>
<tr>
<td>Comm.</td>
<td>4.6202</td>
<td>4.5739</td>
<td>4.5739</td>
<td>4.5981</td>
</tr>
<tr>
<td></td>
<td>(0.7289)</td>
<td>(0.6513)</td>
<td>(0.7040)</td>
<td>(0.7440)</td>
</tr>
<tr>
<td>Form Hyp.</td>
<td>4.6458</td>
<td>4.6927</td>
<td>4.6220</td>
<td>4.8062</td>
</tr>
<tr>
<td></td>
<td>(0.7071)</td>
<td>(0.7440)</td>
<td>(0.6781)</td>
<td>(0.6690)</td>
</tr>
<tr>
<td>Exp.</td>
<td>3.8371</td>
<td>4.7386</td>
<td>3.9685</td>
<td>4.8062</td>
</tr>
<tr>
<td></td>
<td>(0.7040)</td>
<td>(0.7071)</td>
<td>(0.5630)</td>
<td>(0.8839)</td>
</tr>
<tr>
<td>Cont. Var</td>
<td>* 0.0000</td>
<td>1.5000</td>
<td>1.7047</td>
<td>0.9354</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.9911)</td>
<td>(0.7071)</td>
<td>(0.8211)</td>
</tr>
<tr>
<td>Rec. Data</td>
<td>* 2.5000</td>
<td>2.1795</td>
<td>2.4238</td>
<td>2.1213</td>
</tr>
<tr>
<td></td>
<td>(0.9162)</td>
<td>(0.9162)</td>
<td>(0.7763)</td>
<td>(0.5976)</td>
</tr>
<tr>
<td>Int. Data</td>
<td>3.0616</td>
<td>3.3184</td>
<td>3.3452</td>
<td>3.0917</td>
</tr>
<tr>
<td></td>
<td>(1.3024)</td>
<td>(0.9978)</td>
<td>(0.9258)</td>
<td>(1.1631)</td>
</tr>
<tr>
<td>App./Gen.</td>
<td>* 2.2247</td>
<td>0.0000</td>
<td>2.6956</td>
<td>1.2041</td>
</tr>
<tr>
<td></td>
<td>(1.0691)</td>
<td>(0.0000)</td>
<td>(1.4501)</td>
<td>(0.7071)</td>
</tr>
</tbody>
</table>

Note. * indicates statistically significant difference between science and generic problem types at p< .01. There were no statistically significant differences between experimental and control groups.

Table 4.3: Means and standard deviations for the appropriateness-of-use of individual science process skills by treatment and problem type.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Skill</th>
<th>Group With</th>
<th>Observed</th>
<th>Expected</th>
<th>Contribution to Total $X^2$*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rec. Dat.</td>
<td>Exp.</td>
<td>165</td>
<td>136.7</td>
<td>11.912</td>
</tr>
<tr>
<td>2</td>
<td>Meas.</td>
<td>Cont.</td>
<td>54</td>
<td>40.3</td>
<td>9.106</td>
</tr>
<tr>
<td>3</td>
<td>Predict.</td>
<td>Cont.</td>
<td>43</td>
<td>36.4</td>
<td>2.359</td>
</tr>
<tr>
<td>4</td>
<td>Cont. Var.</td>
<td>Exp.</td>
<td>17</td>
<td>14.7</td>
<td>.707</td>
</tr>
<tr>
<td>5</td>
<td>App./Gen.</td>
<td>Exp.</td>
<td>12</td>
<td>10.2</td>
<td>.674</td>
</tr>
<tr>
<td>6</td>
<td>Int. Dat.</td>
<td>Cont.</td>
<td>117</td>
<td>112.1</td>
<td>.415</td>
</tr>
<tr>
<td>7</td>
<td>Comm.</td>
<td>Cont.</td>
<td>346</td>
<td>338.4</td>
<td>.338</td>
</tr>
<tr>
<td>8</td>
<td>Observ.</td>
<td>Exp.</td>
<td>104</td>
<td>101.1</td>
<td>.166</td>
</tr>
<tr>
<td>9</td>
<td>Hypot.</td>
<td>Exp.</td>
<td>287</td>
<td>291.2</td>
<td>.122</td>
</tr>
<tr>
<td>10</td>
<td>Exp.</td>
<td>Exp.</td>
<td>177</td>
<td>175.3</td>
<td>.033</td>
</tr>
</tbody>
</table>

Note. Rank is listed from highest to lowest contribution to the total chi-square.

* $X^2(9, n=64)=25.83, p<.05$

Table 4.4: Results of the omnibus chi-square analysis of the frequency data for experimental versus control groups.
The experimental groups demonstrated recording data more frequently than expected and the control groups demonstrated measuring more frequently than expected.

Table 4.5 depicts the use of the science process skills in order of magnitude of contribution to the overall chi-square test for independence of the science and generic problem conditions. The reader will note that the majority of the overall chi-square value in this case, was accounted for by the first four process skills (recording data, experimenting, making observations and measuring, respectively). Of these four skills, only the skill of recording data was used more frequently for the generic problems than the science-related problems.

Tables 4.6 and 4.7 present the findings of chi-square tests for independence focusing specifically on how the skills were used for the science-related and generic tasks respectively. As was true with the omnibus chi-square tests, the reader should note that in Table 4.6, the majority of the overall chi-square value can be accounted for by three of the process skills (recording data, measuring, and predicting). Again, it was the control groups which used the skills of measuring and predicting more frequently for the science tasks and the experimental groups which used the skill of recording data.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Skill</th>
<th>Group With</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
<th>Contribution to Total $X^2$*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rec. Dat.</td>
<td>Gen.</td>
<td>195</td>
<td>123.1</td>
<td>71.602</td>
</tr>
<tr>
<td>2</td>
<td>Exp.</td>
<td>Sci.</td>
<td>255</td>
<td>187.2</td>
<td>59.165</td>
</tr>
<tr>
<td>3</td>
<td>Observ.</td>
<td>Sci.</td>
<td>155</td>
<td>107.9</td>
<td>48.511</td>
</tr>
<tr>
<td>4</td>
<td>Meas.</td>
<td>Sci.</td>
<td>71</td>
<td>44.5</td>
<td>36.601</td>
</tr>
<tr>
<td>5</td>
<td>App./Gen.</td>
<td>Gen.</td>
<td>20</td>
<td>9.2</td>
<td>22.781</td>
</tr>
<tr>
<td>6</td>
<td>Comm.</td>
<td>Gen.</td>
<td>362</td>
<td>314.8</td>
<td>9.514</td>
</tr>
<tr>
<td>7</td>
<td>Predict.</td>
<td>Sci.</td>
<td>51</td>
<td>40.1</td>
<td>7.297</td>
</tr>
<tr>
<td>8</td>
<td>Hyp.</td>
<td>Gen.</td>
<td>274</td>
<td>262.2</td>
<td>6.793</td>
</tr>
<tr>
<td>9</td>
<td>Cont. Var.</td>
<td>Sci.</td>
<td>22</td>
<td>15.7</td>
<td>5.956</td>
</tr>
<tr>
<td>10</td>
<td>Int. Dat.</td>
<td>Gen.</td>
<td>121</td>
<td>104.3</td>
<td>3.660</td>
</tr>
</tbody>
</table>


* $X^2(9, n=64)=271.88, p<.05$

Table 4.5: Results of the omnibus chi-square analysis of the frequency data for science versus non-science.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Skill</th>
<th>Group With Highest Freq.</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
<th>Contribution to Total $X^2$ *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rec. Dat.</td>
<td>Exp./Sci.</td>
<td>49</td>
<td>37.6</td>
<td>6.411</td>
</tr>
<tr>
<td>2</td>
<td>Meas.</td>
<td>Cont./Sci.</td>
<td>49</td>
<td>39.0</td>
<td>5.587</td>
</tr>
<tr>
<td>3</td>
<td>Predict.</td>
<td>Cont./Sci.</td>
<td>38</td>
<td>32.0</td>
<td>2.487</td>
</tr>
<tr>
<td>4</td>
<td>App./Gen.</td>
<td>Exp./Sci.</td>
<td>2</td>
<td>.9</td>
<td>1.672</td>
</tr>
<tr>
<td>5</td>
<td>Observ.</td>
<td>Exp./Sci.</td>
<td>80</td>
<td>72.4</td>
<td>1.472</td>
</tr>
<tr>
<td>6</td>
<td>Comm.</td>
<td>Cont./Sci.</td>
<td>194</td>
<td>184.7</td>
<td>1.013</td>
</tr>
<tr>
<td>7</td>
<td>Hypot.</td>
<td>Cont./Sci.</td>
<td>138</td>
<td>145.7</td>
<td>.895</td>
</tr>
<tr>
<td>8</td>
<td>Int. Dat.</td>
<td>Cont./Sci.</td>
<td>58</td>
<td>53.6</td>
<td>.776</td>
</tr>
<tr>
<td>9</td>
<td>Cont. Var.</td>
<td>--------------</td>
<td>12</td>
<td>11/13</td>
<td>.169</td>
</tr>
<tr>
<td>10</td>
<td>Exp.</td>
<td>Cont./Sci.</td>
<td>133</td>
<td>133.8</td>
<td>.011</td>
</tr>
</tbody>
</table>

Note. Exp. = Experimental groups. Cont. = Control groups. Sci. = Science-related problems. Rank is listed from highest to lowest contribution to the total chi-square. Expected frequencies listed for controlling variables are experimental/control (there was no difference in the observed frequencies for this skill.)

$X^2(9, \text{n}=32)=20.49, \ p<.05$

Table 4.6: Results of the chi-square analysis of frequency data for experimental/science versus control/science
In Table 4.7, the majority of the overall chi-square value can be accounted for by the first five process skills (controlling variables, measuring, experimenting, stating hypotheses, and recording data, respectively). Also, the reader should note that with the exception of measuring, all skills were used more frequently by the experimental group than the control group.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Skill</th>
<th>Group With Highest Freq.</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
<th>Contribution to Total $X^2$ *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cont. Var.</td>
<td>Exp./Gen.</td>
<td>5</td>
<td>2.9</td>
<td>3.732</td>
</tr>
<tr>
<td>2</td>
<td>Meas.</td>
<td>Cont./Gen.</td>
<td>10</td>
<td>6.4</td>
<td>3.507</td>
</tr>
<tr>
<td>3</td>
<td>Exper.</td>
<td>Exp./Gen.</td>
<td>63</td>
<td>56.1</td>
<td>1.979</td>
</tr>
<tr>
<td>4</td>
<td>Hypot.</td>
<td>Exp./Gen.</td>
<td>157</td>
<td>168.9</td>
<td>1.962</td>
</tr>
<tr>
<td>5</td>
<td>Rec. Dat.</td>
<td>Exp./Gen.</td>
<td>116</td>
<td>107.6</td>
<td>1.519</td>
</tr>
<tr>
<td>6</td>
<td>Predict.</td>
<td>Exp./Gen.</td>
<td>10</td>
<td>8.6</td>
<td>.543</td>
</tr>
<tr>
<td>7</td>
<td>Comm.</td>
<td>Exp./Gen.</td>
<td>195</td>
<td>198.7</td>
<td>.159</td>
</tr>
<tr>
<td>8</td>
<td>Observ.</td>
<td>Exp./Gen.</td>
<td>24</td>
<td>23.5</td>
<td>.028</td>
</tr>
<tr>
<td>9</td>
<td>App./Gen.</td>
<td>Exp./Gen.</td>
<td>10</td>
<td>10.3</td>
<td>.021</td>
</tr>
<tr>
<td>10</td>
<td>Int. Dat.</td>
<td>Exp./Gen.</td>
<td>70</td>
<td>69.9</td>
<td>.001</td>
</tr>
</tbody>
</table>

Note. Exp.= Experimental groups. Cont. = Control groups. Gen.= Generic problems. Rank is listed from highest to lowest contribution to the total chi-square.

$* X^2(9, n=32)=13.45, p<.05$

Table 4.7: Results of the chi-square analysis of frequency data for experimental/generic versus control/generic.
Summary

Of the five experimental hypotheses examined in this study, one revealed significant omnibus results. Experimental groups demonstrated significantly greater frequency of use of the science process skills on the science-related problem-solving tasks than did control groups. Although no other omnibus test results were significant, much information was revealed about groups’ frequency and appropriateness of use of the individual skills. In addition, robust test statistics were associated with the comparisons of science and generic problem types. Finally, the results of the hypothesis related to group cooperation showed that groups did not differ in how well they worked together, thus, the level of group cooperation did not impact the results of the other tests. Chapter 5 contains discussion and implications of the results.
CHAPTER 5
SUMMARY AND DISCUSSION

Introduction

This chapter presents a summary of the study followed by a discussion of conclusions, implications, and recommendations based upon the findings of the investigation. This study was an investigation of the effects that implementation of inquiry-based instruction in elementary classrooms had on students' use of selected science process skills. More specifically, the study was designed to answer the question, "does the use of inquiry-based instruction in elementary school science increase either the frequency or appropriateness of student's use of science process skills?" Also of interest in the study was whether the level of cooperation exhibited by students as they worked in groups to solve problems would be higher for the students who had been taught with inquiry-based instructional strategies.

The experimental groups for the study included students in two rural fifth grade classes whose teachers had received training in the implementation of inquiry-based instructional strategies in
science. The control groups were comprised of two classes of fifth grade students from the same school systems whose teachers had not received the training in inquiry-based instructional strategies.

Videotapes were made of the groups of students performing a variety of problem-solving tasks. These tapes were scored by trained, blind raters. The instrument used, the Science Process Skills Observation form, was developed and field-tested for this study based upon the suggestions of Bijou, et al. (1969) and the discussion of Flander's Interaction Analysis in Gay (1992) for the observation of behaviors over a period of time. It was used to evaluate the student groups' use of the science process skills of observing, measuring, predicting, communicating, forming hypotheses, experimenting, controlling variables, recording data, interpreting data, and applying and generalizing results. Scores for frequency and appropriateness of student use of these science process skills were recorded. In addition, a rating of group cooperation was obtained.

The data related to the frequency of use of the science process skills consisted of tallies on the Science Process Skills Observation form representing each time use of a skill was observed. These nonparametric data were analyzed using several chi-square tests for independence. The results of the analyses demonstrated that groups of students who were taught with inquiry-based instructional
strategies used the science process skills more frequently overall than the comparison groups. The data related to the appropriateness of use of the science process skills consisted of ratings from a five-point Likert scale. These parametric data were analyzed using a repeated-measures analysis of variance (ANOVA). The analysis revealed there was no significant difference between the groups in terms of overall appropriateness of use of the skills.

Discussion

Discussion of Results Related to Frequency of Use Data

The research hypotheses for this study which related to frequency of use of science process skills were:

H₁ Students of teachers who are taught science using inquiry-based instructional strategies will demonstrate science process skills with greater frequency while solving science-related problems than will similar students who are not taught science with inquiry-based instructional strategies;

H₂ Students of theachers who are taught science using inquiry-based instructional strategies will demonstrate process skills with greater frequency while solving generic problems than will similar students who are not taught science with inquiry-based instructional strategies;
The omnibus chi-square test to determine whether or not there was a statistically significant difference between experimental (taught with inquiry-based instructional strategies) and control (taught without the use of inquiry-based instructional strategies) groups, reported in Chapter 4, revealed a significant difference between the groups ($X^2(9, \, n=64)=28.53, p<.05$). An examination of Table 4.4 revealed that the science process skills of recording data, measuring, and predicting accounted for the majority of the difference between the experimental and control groups. Table 4.5 shows the skills of recording data, experimenting, observing, and measuring as contributing prominently to the overall chi-square value ($X^2(9, \, n=64)=271.88, p<.05$). Of particular interest in this table is that with the exception of recording data, the top four skills (also including experimenting, observing, and measuring) were used more frequently on the science-related problem solving situations. This may suggest that the problem solving patterns established by the groups for the science-related problems did not persist into the generic problem situations and for this reason, the groups' frequency of use of the skills did not transfer from the science-related to the generic problem situations.

The above conclusions are based upon collapsing the data over all skills observed in the study. That is, the skills were examined
without regard for whether they occurred in a science-related or
generic task. For a greater understanding of the results, it is
necessary to examine the skills individually. Table 4.4 demonstrates
that the skill of recording data was used more frequently by the
experimental group and that the skills of measuring and predicting
were used more frequently by the control groups. Table 4.5 shows
that students' use of the skills of experimenting, observing, and
measuring all differed from the students' use of the skill of recording
data. These three skills were significantly different with regard to
their frequency of use from science to generic, when averaged over
treatment conditions. The skill of recording data was used
significantly more for the generic tasks than the science tasks. This
may be explained by the students arriving at the answers to the
generic problem situations more directly than they did with the
science-related problems. There are no data to support this
supposition, however, if this happened, the students might do less
experimenting, measuring, and observing and might be more likely
to record any data and move on to discussing alternative answers.
By doing this, the students might get into a cycle of repeatedly
looking at what information they are given and analyzing it, thereby
recording data repeatedly without engaging in the experimentation
process.

105
In summary, the data supported the first experimental hypothesis; that the experimental groups would use the science process skills more frequently than the control groups on the science-related tasks. Further, the data supported rejection of the second experimental hypothesis; that there would be a difference in frequency of use of the science process skills when the experimental and control groups are compared.

Discussion of Results Related to Appropriateness-of-Use Data

The research hypotheses related to the students' appropriateness of use of the science process skills were;

H3 Students who were taught science using inquiry-based instructional strategies will demonstrate science process skills with greater appropriateness while solving science-related problems than will similar students who were not taught science with inquiry-based instructional strategies;

H4 Students who were taught science using inquiry-based instructional strategies will demonstrate process skills with greater appropriateness while solving generic problems than will similar students who were not taught science with inquiry-based instructional strategies;
The omnibus test of experimental (inquiry-based instruction) versus comparison (no inquiry-based instruction) groups (see Table 4.3) with regard to the appropriateness of use of the science process skills without considering task types (science/generic) yielded $F=.839$, $p>.05$. This indicated that there was not a significant difference between the two groups. An examination of the group means in Table 4.1 revealed that the experimental groups were, on average, more consistent in the appropriateness of their use of the science process skills over both problem types (science and generic) than were the comparison groups.

Table 4.1 presents the means and standard deviations for the appropriateness-of-use data gathered in the study. A repeated measures ANOVA was performed on the data in order to explore the magnitude of the differences between the means. Table 4.2 reports the results of the ANOVA ($F(1,63)=.839$, $p>.05$). Because both hypotheses related to the appropriateness of use of the skills were based upon a significant difference between the experimental and control groups, both hypotheses were rejected.

A possible explanation of these results is that the John Henry Effect may have occurred. That is, despite the researcher's attempts to keep the goals of the study from the cooperating control teachers, they may have instructed their students in how to solve problems.
It is possible that they may have correctly interpreted the intentions of the study when the researcher described how the room was to be set up and that the students would be video taped doing "activities."

An examination of Table 4.3 shows that the appropriateness of use of the skills of controlling variables, observing, and measuring was greater for the generic tasks than the science tasks. The skills of recording data and applying and generalizing were rated as less appropriately used during the generic tasks in comparison to the science tasks. There are three skills common to both which were used significantly differently on science and generic tasks. The skill of recording data was used both more frequently and more appropriately on the science tasks than it was on the generic tasks. In opposition were the skills of observing and measuring which were used more appropriately and more frequently on the generic problems.

Table 4.3 depicts the relationships of the individual skills by experimental and control conditions. In no case were any of the differences found to be statistically significant. Further, for seven of the ten skills, the mean rating of appropriateness of use is higher for the comparison groups than it was for the experimental groups. In view of the homogeneity of the findings, the researcher is confident of the results of these tests and in the decision not to reject the null.
Discussion of Results Related to Group Cooperation Data

The research hypothesis related to the level of cooperation between students in the groups was;

$H_5$ Students of teachers who are taught science using inquiry-based instructional strategies will demonstrate group cooperation differently than will similar students who are not taught science with inquiry-based instructional strategies.

The results of the test of this hypothesis revealed there was not a statistically significant difference between the experimental and control groups ($F(7,63)=.96$, $p>.05$). The impact of this finding on the study was that the level of cooperation between members of the groups as they attempted to solve the problems was not responsible for the outcome of the tests of the other hypotheses.

In summary, there may be several reasons leading to the findings of this study. The first possibility is that the John Henry effect may have occurred. That is, the control teachers may have correctly predicted the intentions of the researcher and specifically addressed problem solving techniques with their students just prior to the data gathering. The second possible explanation is that the experimental teachers did not regularly use the inquiry-based instructional strategies and approaches in which they had been
trained. There was no way for the researcher to control for this possibility in the experimental design.

The third possible explanation for the results of the study is that a reactivity effect may have occurred. The presence of the researcher (a stranger) in the classroom, the introduction of several pieces of apparatus for recording the data, the manipulation of the classroom environment, and the introduction of problem-solving activities may have combined to affect the results. Although these may have had an effect, the method and apparatus for gathering the data was the same in all classrooms.

The fourth possible explanation of the results of the study may lie in the sensitivity of the rating system/methodology. This seems unlikely because the observation instrument and definitions supporting it were based upon the related literature and interrater reliability was very high. It is possible that use of the observation instrument for recording the actions of individual students rather than groups might add sensitivity to the methodology.

The fifth possible explanation of the findings of the study is that the control students may have had sufficient prior knowledge in the area of problem solving from prior experiences in school to perform well on the problem-solving tasks given to them. This possibility should have been controlled for by the fact that the
principals had randomly assigned the students to classes in the schools.

Implications and Recommendations to Science and Teacher Education

The following are implications drawn from this study. They are offered in an effort to inform and assist those who seek to encourage their students to use science process skills and who would pursue the implementation of inquiry-based instructional strategies in science education.

The implications and recommendations from this study are related to teacher training. The results of the study indicate that students in classes where greater attention was given to inquiry in science, via the teacher/resource person teams, used seven of ten science process skills more frequently in science-related problem-solving tasks than did the students in the control groups. These results may be attributable to two factors. The first is that the groups were trained by their teachers in the use of inquiry-based instruction and the application of science process skills to problem situations. The second is that the instructional methods with which the children were taught were modeled by the teachers in the experimental classrooms in intentional and cooperative ways.
Recommendations and Suggestions for Further Study

In light of the expressed desire of many in education and business to see students become better equipped to solve problems, the first recommendations are:

- Schools should implement programs to train inservice elementary teachers in the nature and use of both inquiry-oriented instructional strategies and science process skills.
- Students' cognitive knowledge of science process skills should be directly measured to determine whether students' lack of use of the skills is based upon lack of cognitive knowledge or some other factor.
- Training in problem-solving skills used in other areas of the curriculum (specifically mathematics and social studies) should be included with training in science process skills for both inservice and preservice education of teachers.

National leaders in the areas of mathematics and social studies (NCTM, 1989; NCSS, 1994) suggest the implementation of problem-solving skills into the broader curriculum. In light of this, the following recommendation is made.

Based upon the results of this study, it is reasonable to expect that if students are exposed, via inquiry-based instructional strategies, to science process skills that facilitate problem solving,
they will use them more frequently when presented with problem-solving tasks or situations (Brotherton & Preece, 1995; Matthews, 1992; Radford & Ramsey, 1996). This would begin to address the concern jointly expressed by many (Aldridge, 1991; NCTM, 1989; NCSS, 1994; Sivertsen, 1993) that students need to see the connections between concepts within disciplines and between separate areas of the curriculum.

Based upon the findings of this study, the following areas of inquiry are recommended for further study:

1. Replication of this study should be conducted with a larger sample of subjects from the full range of elementary grade levels. In addition, the students should be tested several times over a period of at least one year. This would allow time for the habits of mind discussed by Aldridge (1991) to develop and for differences between experimental and control groups to become more pronounced.

2. An experimental study should be conducted to examine which of the science process skills are most often used and in what order students tend to use them. This study should also facilitate the examination of pairs of skills which are closely related to one another for example, measuring and recording data.

3. A qualitative study should be conducted to determine the most effective way to instruct preservice and inservice teachers and
resource professionals in the nature and use of science process skills.

In any future research into the area of science process skills, the instrument and methodology developed for this study would be suitable. One refinement which should be made in order to more accurately collect the data for the above areas of research relates to the time intervals. The time intervals on the instrument represent two-minute blocks. Although this was adequate for this study, for studies involving older students, the instrument should be modified to have one-minute blocks as the skills would likely be used in a more rapid progression.

Summary

The potential of inquiry-based instruction has been discussed earlier. It has been demonstrated within the review of literature that if inquiry-based instruction is carefully and properly done, the return both in terms of student achievement and learner satisfaction can be great. This study provides a small sample view of the effectiveness of inquiry-based instruction with regard to one aspect of elementary science education. Although the scope of this investigation was narrow, it provides a framework, process, and useful instrumentation which may be used to evaluate the effects of inquiry-based instructional strategies on science education.
3-15-95

Dear Parent/Guardian

Next week, Doug Smith, a doctoral student from Ohio State University and an Ohio certified teacher, will be conducting a study in (TEACHER'S NAME) 's class at (SCHOOL) which has been approved by (PRINCIPAL NAME).

Part of the study involves videotaping your child in a group as the group participates in some classroom activities related to problem-solving. The videotapes will only be viewed by people directly involved with the study and neither your child, the teacher, nor the school will ever be referred to in a way that would enable anyone to directly identify your child.

Please indicate below, whether you do or do not grant your permission for your child to be videotaped by placing a mark in the blank to the left of one of the statements.

_____ I DO grant my permission for my child to be videotaped for the study mentioned above.

_____ I DO NOT grant my permission for my child to be videotaped for the study mentioned above.

___________________________________________   ________________
Parent/Guardian signature                        Date

PLEASE RETURN THIS FORM TO THE SCHOOL AS QUICKLY AS POSSIBLE
APPENDIX B

LETTERS TO COOPERATING PRINCIPALS AND COOPERATING TEACHERS
Dear (Principal),

Hello, I am Doug Smith, a doctoral student in Teacher Education at Ohio State University. The purpose of this letter is to request your help in gathering data for my doctoral dissertation. As you are aware, some of your staff members are involved in an environmental science partnership program with the Ohio Department of Natural Resources (ODNR) under the co-direction of personnel from ODNR, the Science and Mathematics Network of Central Ohio, and Dr. Deborah Bainer of the Ohio State University (Mansfield Campus).

Upon your approval, I would like to involve two of your teachers and their classes in to examine some of the ways in which students' thinking habits have been affected by the partnership program. To do this, I plan to first interview the teachers selected, then use the teacher's classrooms as a venue to observe and videotape their students in problem-solving situations (I will provide all the materials), and finally, interview the teachers a second time. The problem-solving exercises will take a total of about 1.5 hours to complete once they have begun and each interview will not exceed 45 minutes.

Each student will be asked to have his/her parents fill out a parental permission form allowing their child to be videotaped. The individual identity of the children on videotapes will be kept strictly confidential and at no time, either now or in the future, will any child, the school, or the district be referred to in any terms that could lead directly to identification of individuals. Further, the tapes will be viewed only by people involved with the study and will be erased at the end of the study.

If you approve of your teacher's and student's involvement, I would like to begin very soon. I have included the name of one teacher in your school with whom I would like to work since she is involved in the partnership process. I would need for you to identify another teacher in the same grade level either in your building or in the nearest building to you who is, in your estimation, as similar as possible to the teacher I have mentioned. (Characteristics to consider in your choice would include teaching style, gender, class size, experience, etc.) This teacher's class will serve as the control group and will be involved in exactly the same way as mentioned above. If you approve of the study, please approach the teachers on my behalf and present the idea and letters to them.

I have included in this packet; a letter to each of the teachers outlining the study and what their involvement would be; a copy of my vita; a tentative time line for their involvement in the study; and a sample parental permission form. Once you have identified the control teacher, would you please allow both teachers to view this packet and to keep a copy of the cooperating teacher letter. I will be calling you shortly to confirm
all of this. Following your confirmation, I plan to call your teachers to set up initial interview dates and times.

If you have any questions or concerns, please do not hesitate to call. My phone numbers are; H- (614) 875-7157, and O- (614) 292-1280. Thank-you in advance for your consideration of these materials and your help in this project.

Sincerely,

Douglas W. Smith
Dear Teacher,

Hello, I am Doug Smith, a doctoral student in Teacher Education at the Ohio State University. The purpose of this letter is to request your help in gathering data for my doctoral dissertation. As you may be aware, some of the staff members in your building are involved in an environmental science partnership program with the Ohio Department of Natural Resources (ODNR) under the co-direction of personnel from ODNR, the Science and Mathematics Network of Central Ohio, and Dr. Deborah Bainer of the Ohio State University (Mansfield Campus).

The purpose of this letter is to ask for your assistance in evaluating some of the effects of the partnership process mentioned above on the way in which students' thinking habits have been affected by it. To do this, I plan to first interview you, then use your classroom as a venue to observe and videotape your students in problem-solving situations (I will provide all the materials), and finally, interview you a second time. The problem-solving exercises will take a total of about 1.5 hours to complete once they have begun and each interview will not exceed 45 minutes.

Each student will be asked to have his/her parents fill out a parental permission form allowing their child to be videotaped. The individual identity of the children on videotapes will be kept strictly confidential and at no time, either now or in the future, will any child, the school, or the district be referred to in any terms that could lead directly to identification of individuals. Further, the tapes will be viewed only by people involved with the study and will be erased following this study.

If you are willing to be involved, I would like to begin very soon. I will be calling you shortly to confirm all of this and to set up initial interview dates and times. If you have any questions or concerns, please do not hesitate to call. My phone numbers are; H- (614) 875-7157, and O- (614) 292-1280. Thank-you in advance for your consideration of these materials and your help in this project.

Sincerely,

Douglas W. Smith
Time Line for Cooperating Teacher Involvement

March 8 - 10  
**Call to Cooperating Teacher** to set up initial interview for the next week.

March 13 - 17  
**Initial Interview.** One meeting of about 45 min. On the same day as the interview, Doug Smith will observe the class, provide parental permission forms for all students, and would appreciate being introduced to the class.

March 17 - 26  
**In-class videotaping of students** doing group activities for about 1.5 hours. This session will be lead by Doug and assisted by the Cooperating Teacher.

March 26 - 31  
**Final Interview.**

* This signifies that the session will be **audio-tape** recorded.

**This signifies that the session will be **video-tape** recorded.
APPENDIX C

FIRST AND SECOND TEACHER INTERVIEW QUESTIONS
First Teacher Interview Questions

1. How many years have you been teaching?

2. What grades and (if applicable) subjects have you taught?

3. How do you usually go about teaching science?
   A. How is your class usually organized? (individuals, groups, any special facilities used, etc.)
   B. How would you describe your role during science lessons?

4. Has anything happened this year that has significantly affected (changed or altered) your approach to teaching?

5. Do you use groups in your class?
   A. Which category best represents the number of times per week that you use groups in your class?
      0  1-3  4-6  7-9  10-12  >12
   B. Please talk about the group process routines you use?

Second Teacher Interview Questions (following data gathering)

1. Are you familiar with the term science process skills?

If teacher was not familiar with them, the researcher gave a brief description of the term and a few examples.

2. What do you know about these skills?

3. How important do you think it is for students to be able to
use science process skills?

If teacher is involved in the Partnership:

4. Would you please describe the partnership team you are a part of and how you have worked out the partnership in your classroom?
Science Process Skills Observation Instrument

Instructions

When filling out the accompanying Observation Instrument, you will be referring to the list of terms and definitions below. Before beginning, please review the terms and definitions. The task codes are: FC, FU, and TR.

As you begin, you will be looking for instances of the use of these skills. Each time you see a skill used, make a mark on the line to the right of the skill you saw. At the end of the session, stop the tape and reflect about the overall quality of each skill. Please note, just because a skill may have been used several times does not necessarily indicate that it was used appropriately! Please use the definitions below of each level of the Likert scale as a guide in your decision-making. Remember, in general, appropriateness is a function of whether the skill was used in a way and/or at a time which facilitated movement of the group toward a solution.

Observing- Use of the senses to gather data about objects and events
Measuring- Reference to units of measure when describing objects or events
Predicting- Forecasting future events based on observations and inferences
Communicating- Use of spoken and written words, graphs, drawings, and diagrams to share information with others.
Formulating Hypotheses- Expression of an educated guess as to the solution of a problem
Experimenting- Carrying out a plan to test a hypothesis.
Controlling Variables- Devising two experimental conditions such that the only difference between the conditions is the variable being tested
Recording Data- Written expression of facts, observations, measurements, etc.
Interpreting Data- Use of analysis and synthesis of data to support or refute a hypothesis
Inferring/Applying and Generalizing- using the findings of an experiment to answer experiment related questions and to address situations that are similar but not directly related to the experiment.
The Likert scale for appropriateness of use of the skills has five points. The definitions of the points are described below and should be used as a guide upon which to rate each skill. For this scale, each skill should be considered for the entire session rather than rating each block within a session.

1  Highly inappropriate- Use of the skill was unjustifiable or misplaced in the problem-solving process and in no way contributed to the group's progress toward solving the problem.
2  Somewhat inappropriate- Use of the skill was justifiable but did not contribute to the groups progress toward solving the problem.
3  Somewhat appropriate- Use of the skill was justifiable and contributed weakly to the groups progress toward solving the problem.
4  Moderately appropriate- Use of the skill was justifiable and contributed moderately to the groups progress toward solving the problem.
5  Highly appropriate- Use of the skill was justifiable and contributed substantially to the groups progress toward solving the problem.

The Likert Scale for the Overall Group Cooperation also has five levels. This scale attempts to quantify the facility with which the students operate as a group in the problem solving settings. For this scale, please consider the entire session.

1  None observed- The students did not work together at all.
2  Poor- A few of the students worked together but no organization of the group was evident.
3  Fair- The students assumed roles but failed to maintain focus on the objective or failed to work cooperatively throughout the problem-solving session.
4  Good- Students assumed roles in the group and worked cooperatively throughout the problem-solving session maintaining focus on the objective.
5  Excellent- Students assumed roles in the group, worked cooperatively throughout the problem-solving session, maintained focus on the objective, and successfully completed the task.
Please use the cells below to record each time you observe the skill being used on the video tape. After you have finished the tape, please rate the overall appropriateness of the subjects use of each skill using the Likert scale to the right of the cells. Please refer to the instruction sheet for descriptions of each level. In the blank at the bottom of the form, please rate the group on how well its members worked together.

<table>
<thead>
<tr>
<th>Observation Time (min. : sec.)</th>
<th>0:00</th>
<th>2:00</th>
<th>4:00</th>
<th>6:00</th>
<th>8:00</th>
<th>Total times</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKILL</td>
<td>1:59</td>
<td>3:59</td>
<td>5:59</td>
<td>7:59</td>
<td>10:00</td>
<td></td>
</tr>
<tr>
<td>Observed Appropriateness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observing</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Measuring</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Predicting</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Communicating</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Formulating Hyp.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Experimenting</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Cont. Variables</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Recording Data</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Interpreting Data</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Apply. and Gen.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Overall Group Cooperation: 5 (excellent) 4 (good) 3 (fair) 2 (poor) 1 (none observed)
APPENDIX E

PROBLEM SITUATIONS EMPLOYED IN THE STUDY
All Wound Up

Greetings inventors! You are a group of transportation specialists. You have been sent to the country of Lower Sloblovia. They have asked you to invent a tractor for them that can travel by itself for at least four feet along a level surface. Since they have very few resources, you will have to use only the items in front of you. Remember, your tractor will have to travel under its own power.

Your group has 10 minutes to solve this problem.

Materials:

4 paper clips 1 Acetate sheet
1 drinking straw 1 Steel washer
1 thread spool 1 pr. scissors
4 rubber bands 1 hole punch
Paper and pencil
THE FOXES AND THE CHICKENS
--(This Could Get Ugly!!)--

Five foxes and five chickens are traveling as a single group through the woods. There are five of each because if the foxes were ever to outnumber the chickens, they would eat the chickens immediately! The group happens upon a river. This is a problem since none of them can swim. Luckily, there is a small canoe on the shore BUT only three of them can be in it at any one time.

Your task is to come up with a way that the whole group can get across the river (a few at a time), using only the canoe for transportation. Remember, the foxes may never outnumber the chickens on either shore or in the canoe. (Yes, the chickens and the foxes can both paddle the canoe.)

Your group has 10 minutes to solve this.
Materials: 5 red chips 5 blue chips
1 "canoe" Paper and pencil

131
Fill Up

Questions: Is there air space between the particles of gravel?
If so, how much?
Could your answers be wrong? Why or why not?

Your group has been challenged to find out how much air space there is between the particles of gravel in the container on the table. Remember, you should be as accurate as possible with your answer. All that you need to figure it out is here.

Your group has 10 minutes to solve this problem.

Materials:
Gravel
Graduated cylinders or measuring cups
Water
Empty container (for gravel)
Paper and pencil
The Desert Problem

You are the only survivors of a plane crash in the Arizona desert. Just as engine trouble developed, you were told by the pilot and co-pilot (both killed in the crash) that the nearest town is 60 miles to the south. The area around you is flat and apparently barren except for some small (and pointy) desert plants. You know that the temperature where you are reaches 110° F in the mid-afternoon and goes down to 35° F at night. Everyone is dressed in light weight clothes and when you checked your pockets, you found handkerchiefs, small change, matches, lipstick, credit cards.

You have all agreed to stay together for survival. Your task is to rank the eleven items on your groups sheet from most important to least important using 1 for the most important and 11 for the least important. Then, describe what you would do to survive using these items.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>_____</td>
<td>A cosmetic mirror</td>
</tr>
<tr>
<td>_____</td>
<td>A pair of sunglasses per person</td>
</tr>
<tr>
<td>_____</td>
<td>Flash light (4 battery size)</td>
</tr>
<tr>
<td>_____</td>
<td>Hunting knife</td>
</tr>
<tr>
<td>_____</td>
<td>Magnetic compass</td>
</tr>
<tr>
<td>_____</td>
<td>One quart of water per person</td>
</tr>
<tr>
<td>_____</td>
<td>One top coat per person</td>
</tr>
<tr>
<td>_____</td>
<td>Parachute (red and white)</td>
</tr>
<tr>
<td>_____</td>
<td>Forty-fivecaliber pistol (loaded)</td>
</tr>
<tr>
<td>_____</td>
<td>Pilots map of the area</td>
</tr>
<tr>
<td>_____</td>
<td>Bottle of Tylenol</td>
</tr>
</tbody>
</table>
APPENDIX F

RATER TRAINING SESSION PLANS
INITIAL RATER TRAINING SESSION PLAN

8:00-8:45 Breakfast and introduction to the day's activities.

8:45-9:15 Introduction to the Study
- Background of the Study
- Discussion of Rater's part in the study
- Introduction of the Science Process Skills

9:15-10:15 Viewing of first tape from the Pilot Study
- The tape will be viewed only for frequency of use of process skills. The tape will be stopped and repeated as necessary to establish agreement among raters and researcher as to what does and does not constitute use of a process skill.

10:15-11:30 Discussion of the high inference scales
(Appropriateness and Cooperation)
- The definitions of each level of the scale for "appropriateness of use" will be discussed.
- The tape used above will be viewed again only for the purpose of recording observations on the "appropriateness of use" scale.
- The definitions of each level of the scale for "overall group cooperation" will be discussed.
- The tape used above will be viewed again only for the purpose of recording observations on the "overall group cooperation" scale.

11:30-12:15 Lunch

12:15-12:45 First Trial
- Raters will evaluate another tape from the pilot study and will discuss their observations and interpretations. During this time, the raters may stop the tape at any time for clarification.

12:45-1:00 Discussion of first trial
- Clarification of any points of disagreement
- Verification that all raters are using the instrument in the same way.
1:00-1:30 Second Trial

- Raters and researcher will simultaneously watch and independently rate a third tape from the Pilot Study. During this trial, the tape will be stopped every two minutes for a period of about 30 seconds to allow the raters to record their observations for each cell (referring to the "frequency of use"). The raters and researcher will also record their observations on the two Likert scale items immediately following the final two minute interval.
- The results will be tabulated on a common form and analyzed for agreement according to the criteria listed in the Methodology section of the proposal.

**NOTE** If suitable reliability is achieved at this point, the raters will be issued their tapes and the training session will be over. If this is not the case, differences between raters will be analyzed and areas of disagreement will be further discussed. Following this, the raters will repeat the steps for the second trial with a different tape.

1:30-2:00 Discussion of differences in the results of the second trial.

2:00-2:30 Third Trial

- Following the same protocol as the second trial.

2:30-3:00 Discussion of differences in the results of the third trial.

**NOTE** If at this point, there is not suitable agreement, the raters will be asked to attend another training session at which the differences in the ratings will be discussed and further trials will be run according to the above protocol.
APPENDIX G

FIELD TEST OF METHODS,

MATERIALS, AND INSTRUMENTS
Field Test

Introduction

Since the problem-solving situations and the rating instrument to be used in the study were either created or modified from their original forms, it was necessary to perform a field test to determine the suitability of these components. Additionally, the issues of time to be allowed for each problem, logistics of coordinating group movement about the room during the data gathering visit, and the suitability of the equipment used in gathering the data needed to be addressed. Thus, a pilot study, designed to address these issues was carried out in a central Ohio, suburban, fifth grade classroom.

The teacher was asked by the researcher to place the students in two of her classes into four groups each and to form the groups as she normally would. Consequently, the students in these classes were heterogeneously divided into four groups per class by the teacher. The researcher visited the class one week prior to the data gathering visit and observed the teacher's classes, spoke informally with students, and was allowed by the cooperating teacher to address the classes involved in the pilot study. During this time, the researcher told the students that they were to be videotaped working on some problems that the
researcher would bring on the next visit. The students were given a parental permission form to take home, have signed, and return. They were told that without this, they would not be allowed to participate. The researcher also discussed with the students what the equipment was that was to be used in the videotaping and asked if there were any students who were unfamiliar with it. There were no students who had never had experience with the equipment. The researcher then asked if there were any questions, there were none.

On the day of the data gathering, the following steps were performed:

1. The problem situations were set up at the stations in the classroom (1-4) as follows:

   Station 1       All Wound Up
   Station 2       Foxes and Chickens
   Station 3       Fill Up
   Station 4       The Desert Problem
   Station 5       Bubble Busters
   Station 6       Afloat

3. The groups of four or five students were randomly assigned to starting positions in the room. Groups traveled from station to station in a random order (determined by the
researcher using a random number table) under the direction of the researcher and the session was finished when all groups had completed all tasks.

4. Ten minutes were used to introduce the procedure and to ensure that all activities and recording equipment were prepared. Since the researcher was the only one to operate the cameras, the starting and stopping times were staggered by about 20 seconds to allow the researcher to turn on/off the cameras.

5. Each group was allowed 10 minutes on each activity. After each of the cameras was activated, each group was instructed to begin. Each group working on each problem was videotaped. There was no researcher/teacher involvement in the problem-solving process, however, the researcher and teacher both circulated about the room to monitor and facilitate the process. For example, clarification of tasks, provision of materials, dealing with classroom management issues, etc. are examples of the facilitative behaviors in which the teacher and researcher engaged which did not impact the problem-solving process in the group.

6. At the end of the ten minute time period, the groups were instructed to stop, clean up, and move to the next activity. A seven minute time was given for the resetting of the activities,
movement of the groups, clarification of any questions, and checking of the cameras/microphones.

7. The activities proceeded in turn as described in steps 5 and 6 until all groups had been video-taped attempting all activities.

8. Upon completion of the above seven steps, groups were gathered as a large group and were asked to write a response to questions regarding how they learned to solve problems and how they were able to use that knowledge in these situations. These questions can be found in Appendix C.

Problem Situations Employed

There were six problem-solving activities used in this study. Four were science related tasks. This means that the tasks required the students to work within the bounds of a science concept which is commonly found in fifth grade science curricula (AIMS, 1987, 1993). The other two problem-solving activities were generic. That is, their solutions were not rooted within the bounds of any particular curriculum content or curriculum-related conceptual framework.

The four science-related activities were obtained from the AIMS Education Foundation materials (1987, 1993) and were
modified slightly for the purpose of facilitating the student's use of science process skills while addressing the tasks without the aid of a teacher. The modifications were necessary for the activities because they were originally intended to be performed as teacher-directed (specifically, guided-discovery), laboratory-type activities. In order to use these in an inquiry-based setting, it was necessary in every case to provide written directions for the students (see Appendix E). These directions presented the problem and any conditions under which the students were to work in solving it, stated the amount of time available to solve the problem, and listed the materials they had been given with which to work.

The observations made in the pilot study revealed that although the groups were allowed 15 minutes to complete the tasks, they finished the tasks in eight to ten minutes in all cases. Based upon these observations, the time allowed for each group to work on each task for this study was ten minutes. The activities used in this study which were generic were of a different nature and are addressed separately. The modified AIMS activities used in this study to assess students ability to use science process skills follow.
All Wound Up (AIMS, 1993). The purpose of this activity was for students to use the materials provided them to construct a "tractor" that would travel at least four feet along a level surface under its own power. The science topic area addressed by this activity is friction. The science process skills listed in the AIMS literature which this activity addressed are; observing, measuring, making and testing hypotheses, data recording, interpreting data, and applying and generalizing results.

Fill Up (AIMS, 1987). The purpose of this activity was for students to discover that space exists between particles of gravel. The science topic area addressed by this activity is space in matter. The science process skills listed in the AIMS literature which this activity addresses are; measuring, estimating, data recording, interpreting data, and applying and generalizing results.

Bubble Busters (AIMS, 1987). The purpose of this activity was for students to address the question, "does the amount of soap in a solution affect the time a bubble will last?" To do this, students will use the materials provided to make soap solutions and blow bubbles to determine how long bubbles of various mixtures will last. The science topic area addressed by this activity is surface tension. The science process skills listed in the
AIMS literature which this activity addresses are; observing, measuring, making and testing hypotheses, data recording, interpreting data, and applying and generalizing results.

Afloat (AIMS, 1993). The purpose of this activity was for students to answer the question, "what happens when a ship goes from fresh water into salt water?" To do this, students were instructed to use the materials provided to simulate the conditions implied in the question. The science topic areas addressed by this activity were density and volume. The science process skills listed in the AIMS literature which this activity addresses are; observing, measuring, making and testing hypotheses, data recording, interpreting data, and applying and generalizing results.

There were two activities used in this study that were not related to any content area. The first problem was modified from a problem presented in the New Games Book (Project Adventure, 1982). The second was taken directly from Beyer (1991). These problem situations were:

The Foxes and the Chickens (Project Adventure, 1982). Five foxes and five chickens are traveling as a single group through the woods. There are five of each because if the foxes were ever to outnumber the chickens, they would eat the chickens immediately! The group happens upon a river. This is a problem
since none of them can swim. Luckily, there is a small canoe on the shore but only three of them can be in it at any one time.

Your task is to come up with a way that the whole group can get across the river (a few at a time), using only the canoe for transportation. Remember, the foxes may never outnumber the chickens on either shore or in the canoe. (Yes, the chickens and the foxes can both paddle the canoe.)

**The Desert Problem** (Beyer, 1991). You are the only survivors of a plane crash in the Arizona desert. Just as engine trouble developed, you were told by the pilot and co-pilot (both killed in the crash) that the nearest town is 60 miles to the south. The area around you is flat and apparently barren except for some small (and pointy) desert plants. You know that the temperature where you are reaches 110 °F in the mid-afternoon and goes down to 35°F at night. Everyone is dressed in lightweight clothes and when you checked your pockets, you found handkerchiefs, small change, matches, lipstick, and credit cards.

You have all agreed to stay together for survival. Your task is to rank the eleven items on your groups sheet from the most important to the least important using 1 for the most important and 11 for the least important. Make sure the whole group agrees
on the rankings before you write them down. Your lives depend on it!

The items listed on the ranking sheet are; a cosmetic mirror, a pair of sunglasses per person, a flashlight (4 battery size), a hunting knife, a magnetic compass, one quart of water per person, one top coat per person, a parachute (red and white), a 45 cal. pistol (loaded), the pilot's map of the area, and a bottle of Tylenol.

These problems required validation by a panel of science educators to verify that the science process skills represented in the science-related problems were applicable to the solutions of the problem situations which are not subject-specific as well. To accomplish this, three science educators (one high school biology teacher, one middle school science teacher, and one elementary teacher) were independently given the definitions of the science process skills in question and the goals of the study. They were then asked to evaluate the activities based upon whether or not they felt students would have opportunities to demonstrate the specified science process skills while attempting to solve the problems. All members of the panel agreed that the problems stated above provided opportunity for students to demonstrate the science process skills of observing, making and testing hypotheses, recording data, interpreting data, and applying and generalizing
results as these skills are defined by Padilla (1986) for the AIMS problems.

Data Collection Procedure

1. Students were given permission forms to be video-taped by the researcher which were returned to the cooperating teacher. Only those students who returned signed forms were taped during the data-gathering sessions.

2. The problem situations were set up at the tables in the classroom as stations 1-6 as follows:

   Station 1  All Wound Up
   Station 2  Fill'er Up
   Station 3  Bubble Busters
   Station 4  Afloat
   Station 5  Foxes and Chickens
   Station 6  Desert Problem

3. Students will traveled from station to station in increasing numerical order. Extant groups of students were randomly assigned to starting positions in the room.

4. Five minutes were used to introduce the procedure and to ensure that all activities are prepared.

5. The first activity was given 15 minutes. After the cameras
were activated, the students were instructed to begin. There was no researcher/teacher involvement in the problem-solving process.

6. A five minute time was given for the resetting of the activities, movement of the groups, and checking of the cameras.

7. The second activity was given 15 minutes and followed the same procedure as in #5 above.

For the pilot study, activities 1, 3, and 5 were taped in the initial class and activities 2, 4, and 6 were videotaped in the second class. Thus, two groups were taped for each of the science-related activities and four groups were taped for the generic activity.
Table 1: Means and standard deviations of frequency-of-use of science process skills by problem type
<table>
<thead>
<tr>
<th>Problem Type</th>
<th>Science</th>
<th>Generic</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23.1500</td>
<td>22.8250</td>
<td>22.9875</td>
</tr>
<tr>
<td></td>
<td>(12.1362)</td>
<td>(10.5305)</td>
<td>(11.3296)</td>
</tr>
<tr>
<td>n=8</td>
<td>n=8</td>
<td>n=16</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Values enclosed in parentheses are standard deviations. Groups of students are treated as subjects. Reported values are omnibus in that they reflect collapsing over all skills used by groups.

Table 2: Means and standard deviations of appropriateness-of-use of science process skills by problem type

The videotapes were rated by the researcher using the Science Process Skills Rating Sheet by watching the tapes all the way through while at the same time, rating what was observed without stopping the tapes.

The data were analyzed using paired t-tests of the means of both the frequency-of-use and the appropriateness-of-use.
ratings. The means and standard deviations for the frequency-of-use of science process skills by problem type are reported in Table 1 and the means and standard deviations for the appropriateness-of-use of science process skills by problem type are reported in Table 2.

There were no significant differences between the performance of the groups on science and generic tasks in terms of either frequency of use of the science process skills or the appropriateness of use of the science process skills. The t-test for differences between the means of the frequency-of-use ratings for the science versus generic tasks revealed $t(7)=.349$, $p>.05$. The same test for differences between the means of the science versus generic tasks for the appropriateness of use scale showed $t(7)=.748$, $p>.05$.

Implications to the Study

Several valuable insights were gained from the conduct of this field test. The first was that the instrument, the Science Process Skills Rating Sheet, performed adequately for the observation of the skills overall. For the larger study, however, trained raters will be used which necessitates a modification of the rating sheet. The modification of the rating sheet should be done
to divide the time period of the overall observation into two
minute segments in order to facilitate a closer evaluation of
interrater reliability. It will be necessary to ensure that the raters
are unified on their interpretations of the skills. If the results of
the raters' observations are the same in each of the two minute
cells on the rating sheet and over the range of skills, reliability will
be claimed.

A second modification to the rating sheet indicated by the
field test was to include a scale for the measure of cooperation
between members of the group. It was noted that the groups did
not interact in uniform ways and that this could be a rival
hypothesis. The ability of the students to work together could be
facilitated and greater frequency of use of the skills could result,
yet not be due to the partnering program.

Regarding the data gathering procedures, two primary
modifications are recommended. First, the groups were all
finished with the tasks in eight to ten minutes. The time allotted
for each of the tasks was 15 minutes. Based upon the students'
actions during the "extra" time, it is advisable that the time limits
for all tasks be shortened to ten minutes. The second
recommendation for modification of the procedure is to not use
the tasks titled "Bubble Busters" and "Afloat." This is
recommended based upon the students' lack of success with the problems and the fact that removing two of the science-related tasks provides an equal number of science and generic tasks.

The materials used for the tasks and the instructions given for the groups to follow were adequate and should remain the same for the study.
APPENDIX H

SUMMER INSTITUTE TRAINING AGENDA

(From the PARTICIPANT MATERIALS Packet)
PARTNERING FOR ELEMENTARY ENVIRONMENTAL SCIENCE
June 27-30, 1994

Monday- June 27

9:30 Registration

10:00 Welcome and grabber activity (15)
   Knots (New Games)
   Pre-test and questions (30)
   Setting the stage for inquiry science (60)
   Invent an Animal (OBIS)

11:45 Lunch and Who's Who introduction
   --Interview in mixed group

12:45 Introduction to ODNR

1:00 Multiple ways to teach: Exposition to inquiry
   --Examples of activities using "soil and water"
   Lecture: groundwater/land use chart
   Discussion: protecting groundwater
   Demonstration: groundwater model (Always a River)
   Guided Discovery: Soil Soakers (AIMS)
   Open Inquiry: Filtering Water (Consider the Earth)

2:15 Job Alike Groups (Culture)

3:00 Break and check-in

3:30 Introduction to Partnering
   --Word of phrase that means partnering to you
   --Debrief activity (10)
   --Overview of continuum (20)
   --Feature partnership from Ike I (30)

4:30 Concurrent sessions -- Examples of partnerships (25)
   --select one of three from Ike I

4:55 Reflection (remain in concurrent group settings)

5:15 On own and check-in

5:45 Dinner

6:45 Repeat concurrent sessions

7:15 Overview of course, journal, project, and notebook
   Overview of ordering books

7:30 Forming partnerships
   --Format to identify needs and resources and develop goals
   --Facilitated group planning/sharing (by teams)
   --Sharing at 8:40 (county clusters)

9:00 Adjourn and homework
Tuesday- June 28

8:00 Breakfast

8:45 Introduction to Science and Math Network

9:00 Overview of EE and Science Education (1 1/2)

10:30 Break

11:00 Child development/age appropriateness (cluster groups)
    --Brainstorming developmental characteristics
    What we should know about primary/intermediate
    Implications for teaching
    --Activity-- Incredible Edible Landfill

12:30 Lunch (ODNR)

1:30 Overview of resources

2:00 Partnering time

3:15 Optional activities

6:00 Dinner

7:00 Reflection

7:30 Process skills (1 1/2)
    --What they are and why we use them (15)
    --Reflection: Process skills used so far (15)
    --Reworking activities to add process skills (30)
    Make a Dichotomous Chart (handout)
    --Activity (30)
    Oil Spill (SEA Grant and Flinn Scientific)

9:00 Adjourn and homework
Wednesday- June 29

8:00  Breakfast

9:00  Management of materials
      --Indoors: watershed management (ODNR non-point)
      --Outdoors: Turtle Huddles (Aquatic Project WILD)
      --Debrief

10:45 Break

11:15 Management of students (45)
       --"Biggest problem with management of students indoors and outdoors"
       Debrief techniques which were modeled
       Share own techniques
       --What concerns from cards have not been addressed?

12:15 Lunch and partner planning time

1:30 Questioning strategies
       --Types of questions
          Convergent/divergent
          Bloom's Taxonomy
       --Wait time I and II
       --Activity: Cookie Mining

2:30 Reflection

2:45 Partner planning time

5:30 Reception

6:15 Dinner

7:00 Partner sharing time (county clusters)

8:00 Partner planning time
Optional activities
Homework--Action Plan
Thursday- June 30

8:00 Breakfast

9:00 Assessment
   --Multiple ways
   --Authentic

10:30 Reflection

10:45 Break/Check-out

11:30 Job Alike Groups (next steps)

12:00 Lunch and final partner planning time

1:30 Sharing of partnership plans
   --Small group sharing (four groups--not county clusters)
   --30 sec. summary to the whole group

2:45 Evaluation and closing activity
   --"What I remember most from this institute is ..."

3:30 Adjourn
LIST OF REFERENCES


166


171