Problem Solving and Computer-Assisted Instruction in Science Education: 
Analysis of Research Findings and the Research Process

Jennifer Anastasia Skinner, B.Sc., B.Ed.

Department of Graduate and Undergraduate Studies in Education

Submitted in partial fulfillment of the requirements for the degree of Master of Education

Faculty of Education, Brock University
St. Catharines, Ontario
© June 2001
Abstract

The quantitative component of this study examined the effect of computer-assisted instruction (CAI) on science problem-solving performance, as well as the significance of logical reasoning ability to this relationship. I had the dual role of researcher and teacher, as I conducted the study with 84 grade seven students to whom I simultaneously taught science on a rotary-basis. A two-treatment research design using this sample of convenience allowed for a comparison between the problem-solving performance of a CAI treatment group \((n = 46)\) versus a laboratory-based control group \((n = 38)\). Science problem-solving performance was measured by a pretest and posttest that I developed for this study. The validity of these tests was addressed through critical discussions with faculty members, colleagues, as well as through feedback gained in a pilot study. High reliability was revealed between the pretest and the posttest; in this way, students who tended to score high on the pretest also tended to score high on the posttest. Interrater reliability was found to be high for 30 randomly-selected test responses which were scored independently by two raters (i.e., myself and my faculty advisor). Results indicated that the form of computer-assisted instruction (CAI) used in this study did not significantly improve students’ problem-solving performance. Logical reasoning ability was measured by an abbreviated version of the Group Assessment of Logical Thinking (GALT). Logical reasoning ability was found to be correlated to problem-solving performance in that, students with high logical reasoning ability tended to do better on the problem-solving tests and vice versa. However, no significant difference was observed in problem-solving improvement, in the laboratory-based instruction group versus the CAI group, for students varying in level of logical reasoning ability.
Insignificant trends were noted in results obtained from students of high logical reasoning ability, but require further study. It was acknowledged that conclusions drawn from the quantitative component of this study were limited, as further modifications of the tests were recommended, as well as the use of a larger sample size.

The purpose of the qualitative component of the study was to provide a detailed description of my thesis research process as a Brock University Master of Education student. My research journal notes served as the data base for open coding analysis. This analysis revealed six main themes which best described my research experience: research interests, practical considerations, research design, research analysis, development of the problem-solving tests, and scoring scheme development. These important areas of my thesis research experience were recounted in the form of a personal narrative. It was noted that the research process was a form of problem solving in itself, as I made use of several problem-solving strategies to achieve desired thesis outcomes.
Acknowledgments

Completing this thesis was a rewarding but often difficult task that was only made possible through a network of supportive people. I would now like to express my utmost appreciation toward them.

First and foremost, I would like to acknowledge and give thanks for the extensive time and effort that my faculty advisor, Dr. Joseph F. Engemann, enthusiastically provided over the course of this research. For two years he has been a committed guide, resource, and collaborator. Repeatedly he went above and beyond my expectations of an advisor to ensure the production of a quality piece of work and to facilitate the learning process.

On a personal note I would like to thank my parents, Wayne and Anastasia Skinner, who encouraged me to value education and who have always supported me as I followed my dreams. A special thanks to my fiancé, Ryan Winslow, who was a constant source of understanding and support throughout this two-year endeavor. I would also like to thank close family and friends for their encouragement, especially at times when the process became all-consuming.

Contributions from other faculty members were an integral part of the research process. I would like to express sincere gratitude toward Michelle McGinn, Susan Drake, Carmen Schifellite, and Peter Mernagh for their assistance. Also, a thank you to Larry Bencze from OISE/UT, who was the external examiner at my thesis defense.

I would also like to acknowledge the support I received from the principal, staff, students, and school council members at Glen Brae Middle School in Hamilton. Their willingness to participate in the study helped make my research goals a reality.
Finally, I would like to thank colleagues and classmates who contributed valuable advice and opinions throughout the research process. In particular, I would like to thank Helen Sindall and Kit Davis for their mentorship, as well as Michael, from Education 5P02, for providing me with rubrics which were critical to the development of the scoring scheme.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER ONE: THE PROBLEM</strong></td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background to the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of the Problem Situation</td>
<td>7</td>
</tr>
<tr>
<td>Importance of the Study</td>
<td>8</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>9</td>
</tr>
<tr>
<td>Scope and Delimitations of the Study</td>
<td>10</td>
</tr>
<tr>
<td>Outline of the Remainder of the Thesis</td>
<td>11</td>
</tr>
<tr>
<td><strong>CHAPTER TWO: REVIEW OF RELATED LITERATURE</strong></td>
<td>13</td>
</tr>
<tr>
<td>Introduction</td>
<td>13</td>
</tr>
<tr>
<td>Instructional Methods in Science Education</td>
<td>13</td>
</tr>
<tr>
<td>Computers in Science Education</td>
<td>14</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>20</td>
</tr>
<tr>
<td>Problem Solving and Computer-Assisted Instruction</td>
<td>27</td>
</tr>
<tr>
<td>Logical Reasoning Ability</td>
<td>32</td>
</tr>
<tr>
<td>Analysis of the Research Process</td>
<td>36</td>
</tr>
<tr>
<td>Chapter Summary</td>
<td>39</td>
</tr>
</tbody>
</table>
Appendix A: Ontario Science and Technology Curriculum Expectations .................. 129
Appendix B: Parental Information and Consent Form ............................................. 130
Appendix C: Problem-Solving Pretest ................................................................. 133
Appendix D: Problem-Solving Posttest ............................................................... 137
Appendix E: Problem-Solving Pretest Scoring Rubric ........................................... 141
Appendix F: Problem-Solving Posttest Scoring Rubric ......................................... 143
Appendix G: Problem-Solving Pretest: Possible Solutions ................................... 145
Appendix H: Problem-Solving Posttest: Possible Solutions ................................ 147
Appendix I: Student Exemplars of Problem-Solving Test Responses .................. 149
Appendix J: Problem-Solving Strategies Introductory Lesson .............................. 174
Appendix K: Problem-Solving Strategies Record Sheet ........................................ 175
Appendix L: The Big Mix Up Lab: Lesson Plan .................................................... 176
Appendix M: The Big Mix Up Lab: Student Instructions ....................................... 178
Appendix N: The Big Mix Up Lab: Possible Answers to Problem-Solving Strategies
Record Sheet ........................................................................................................ 179
Appendix O: Science Sleuths: Lesson Plan ........................................................... 181
Appendix P: Science Sleuths: Student Instructions .............................................. 182
Appendix Q: Science Sleuths: Possible Answers to Problem-Solving Strategies
Record Sheet ........................................................................................................ 184
Appendix R: Grade 8 Pilot Test: Problem-Solving Test Feedback ......................... 185
Appendix S: Sample Research Notes Used in Qualitative Analysis ....................... 186
Appendix T: Graphic Representation of Qualitative Findings .............................. 189
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Holistic Scoring Method for Problem-Solving Tests</td>
<td>25</td>
</tr>
<tr>
<td>2. Main Study Procedure</td>
<td>51</td>
</tr>
<tr>
<td>3. Similarities and Differences Between CAI and Traditional Instruction Specific to This Study</td>
<td>56</td>
</tr>
<tr>
<td>4. Demographic Characteristics of Participants</td>
<td>65</td>
</tr>
<tr>
<td>5. Summary of GALT Test Scores</td>
<td>67</td>
</tr>
<tr>
<td>6. Summary of Problem-Solving Pretest Total Scores</td>
<td>68</td>
</tr>
<tr>
<td>7. Summary of Problem-Solving Posttest Total Scores</td>
<td>69</td>
</tr>
<tr>
<td>8. Summary of Problem-Solving Improvement Scores</td>
<td>70</td>
</tr>
<tr>
<td>9. Analysis of Variance or Problem-Solving Improvement Score, GALT score, and Instruction</td>
<td>73</td>
</tr>
<tr>
<td>10. Summary of Problem-Solving Test Scores Given by Rater 1 and Rater 2</td>
<td>74</td>
</tr>
<tr>
<td>11. Kappa Analysis for Pretest Scores Given by Rater 1 and Rater 2</td>
<td>76</td>
</tr>
<tr>
<td>12. Kappa Analysis for Posttest Scores Given by Rater 1 and Rater 2</td>
<td>77</td>
</tr>
<tr>
<td>U1. Open Coding Categories of Qualitative Data Collected During the Independent Study</td>
<td>191</td>
</tr>
<tr>
<td>U2. Open Coding Categories of Qualitative Data Collected During the Research Proposal</td>
<td>195</td>
</tr>
</tbody>
</table>
U3. Open Coding Categories of Qualitative Data Collected During the Thesis.............................................. 198
List of Figures

<table>
<thead>
<tr>
<th>Figure V1: Scatterplot of relationship between pretest total score and GALT</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure V2: Scatterplot of relationship between posttest total score and GALT</td>
<td></td>
</tr>
<tr>
<td>Figure V3: Scatterplot of relationship between improvement score and GALT</td>
<td></td>
</tr>
<tr>
<td>Figure V4: Relationship among problem-solving improvement score, GALT score, and instruction</td>
<td></td>
</tr>
<tr>
<td>Figure V5: Scatterplot of relationship between pretest total score and posttest total score</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER ONE: THE PROBLEM

Introduction

Research on information technology in science education has gone through several phases. Initially, the focus was on comparing the efficiency of learning through a computer-based medium versus learning through a traditional medium (Berger, Lu, Belzer, & Voss, 1994). Alongside the advancement of computer software and hardware has been the advancement of studies surrounding information technology in the classroom. More recently, research has focused on both the content and the methods of teaching and learning through computers, such as learning through constructivism-informed approaches (Berger et al.).

Although a large literature base exists, there is no consensus on the merit of computer-based education (CBE). This research discrepancy is alarming considering the educational funds, time, and instruction invested into the rapid infusion of information technology into our schools. On the one hand, there are researchers who claim that “learning environments which incorporate meaningful and appropriate computer-based instructional activities can stimulate the formation and development of science concepts and problem-solving skills and abilities in learners” (Simmons & Lunetta, 1993, p. 171). On the other hand, there are critics who claim that research does not show a significant advantage of computer-based instruction over traditional instructional methods (Robertson, 1998). The external validity of research in computer-based education has been critiqued for commonly occurring flaws, such as the making of excessively broad generalizations and the use of less than optimal software (Krajcik, Simmons, & Lunetta, 1986).
CBE leads into a great number of possible research avenues, the scope of which is beyond any one study. It encompasses a large range of technological tools and methods for using this technology in the classroom. Within the realm of CBE, computer-assisted instruction (CAI) more narrowly refers to the use of computers in conjunction with regular classroom learning (Berger et al., 1994; Pek, 1996; Tiffin & Rajasingham, 1995). More sophisticated CAI methods are exploratory in nature, meaning they allow students to freely interact with a problem (Pek; Tiffin & Rajasingham). The current study made use of a more sophisticated form of CAI by having students actively engage in learning through an exploratory software program.

The focus of this study was further narrowed down to examine student problem-solving performance within CAI. Problem solving is recognized as an important component of science and mathematics education and has been at the center of many studies (Kilpatrick, 1967; Kinnear, 1986; Krajcik, Simmons, & Lunetta, 1988; Murphy & McCormick, 1997; Polya, 1957; Schoenfeld, 1982; Simmons & Lunetta, 1993; Slack & Stewart, 1990; Soderberg, Stewart, Calley, & Jungck, 1994; Weber, 1993). Moreover, researchers have highlighted the need for further study of problem-solving performance in the use of information technology in science education (Berger et al., 1994; Krajcik et al.).

Background to the Problem

A Need for Problem-Solving Research

Researchers believe that problem solving should be context-specific, active, flexible, inventive, and authentic (Murphy & McCormick, 1997; Slotnick, 1989). To be authentic, problems must be both "personally meaningful and purposeful within a social
null
framework” (Murphy & McCormick, p. 463). In addition, successful problem-solving performance requires explicit teaching, including student-teacher interactions and discussions about decisions and conclusions (Kinnear, 1986; Murphy & McCormick; Weber, 1993).

Traditional problem-solving tasks in science have been textbook- and laboratory-based. Furthermore, due to the demands of laboratory experience (i.e., time, space, and cost) textbook tasks are more often used in the science classroom (Kinnear, 1986). Unfortunately textbook problems solved through the use of the paper-and-pencil method are often narrowed to reinforce only the topic at hand since “problems are usually physically arranged in association with chapters in a textbook or temporally arranged in association with a topic of study” (Kinnear, p. 3). In textbook-based problems the data are immediately available, as compared to the laboratory-based problem in which a student is asked to gather the data before attempting to solve the problem. Although laboratory-based problems are ‘hands-on’ investigations, they often follow prescribed, rather than investigatory or exploratory, methods (Jungek, 1991; Kinnear). These traditional instructional modes have been critiqued for their inability to teach students problem-solving skills which may be transferable to new and/or authentic situations (Pek, 1996; Soderberg et al., 1994). In traditional instruction, students are often passive learners in a teacher-directed learning environment and the emphasis for learning is placed on the “ends” rather than the “means” (Pek).

A review of the literature revealed that certain characteristics of computers are believed to have the potential to better facilitate problem-solving performance (Slotnick, 1989; Friedler, Nachmias, & Songer, 1989; Krajcik et al., 1988; Marsh & Kumar, 1992;
Pek, 1996; Simmons & Lunetta, 1993; Soderberg et al., 1994). For example, computer-based problems can make the problem-solving process more active and authentic, and therefore more purposeful for students (Jungck, 1991; Kinnear, 1986; Slack & Stewart, 1990). Also, students may be required to formulate and evaluate a hypothesis, as well as acquire and evaluate data in a simulated computer problem, thereby making the process more personally meaningful (Kinnear; Slack & Stewart).

Studies have reported a positive effect of CAI on problem-solving performance but under careful scrutiny the methods used to reach these conclusions were questionable. For example, problem-solving behaviours for students engaging in computer-based problems and textbook-based problems were compared (Kinnear, 1986). Computer problem solving was more favourably described than textbook problems, for example students reported that they were easier to solve. However, this traditional form of textbook instruction may not have been an equitable medium with which to compare computer instruction. Perhaps a more active instructional method for the control group, such as a "hands-on" laboratory, would have resulted in greater similarity in the effects of the two forms of instruction. In a 1996 study, Pek concluded that CAI improved problem-solving skills, but did not specify the measures of problem-solving performance. Eylon et al. (1996) found that solving problems through a computer simulation improved some aspects of student learning, but that these effects might have been attributed to enriched concept instruction and not to the simulation alone. In addition, several factors have been found to affect problem-solving performance through CAI (Cypher, 1987; McLean, 1989; Slack & Stewart, 1990). These included such things as the role of the
teacher, student reaction to the simulation, limiting student frustration, interactions in cooperative learning, concept understanding, and logical reasoning ability.

A purpose of this study was to build on previous research in this area. To accomplish this, potential downfalls noted in the methods of previous researchers were purposely avoided and factors thought to affect the relationship between CAI and problem-solving were controlled wherever possible. For example, in the current study, students who received traditional instruction and those who received CAI were both engaged in active learning activities. This was done in hopes of creating a more equitable comparison between problem-solving performance in CAI and traditional instruction, than was reported in previous studies.

Developmental Issues

According to Piagetian theory, people make the transition from the concrete reasoning stage to the formal reasoning stage at approximately twelve years of age. However, some critics of Piagetian theory believe that this transition may occur at a later age. Thus, for most people this transition occurs during or after, the late elementary school years. Since several studies reported a relationship between logical reasoning ability and success in problem solving (Gipson, Abraham, & Renner, 1989; Smith & Sims, 1992; Stewart & Hafner, 1994), a serious educational implication is that there may be a discrepancy between what many students are asked to do in school and what they are developmentally capable of doing. As well, in learning new concepts, some students will have a greater need for concrete hands-on activities than will others (Clewell, Anderson, & Thorpe, 1992). Elementary school teachers should be aware of differing intellectual
abilities among their students and attempt to provide instruction which accommodates these differences.

CAI might be one solution to this instructional dilemma because of its ability to simultaneously offer a wide range of programs to suit the various needs of students. For example, it was suggested that some computer programs adequately present abstract ideas in concrete form, thereby benefiting concrete learners (Papert, 1980; Slotnick, 1980). Other researchers similarly suggested that problem-solving performance through CAI might be connected to logical reasoning ability, but further research was required (Krajcik et al., 1988; Krasnor & Mitterer, 1984).

Development of a Science Problem-Solving Measure

A related purpose of this study was to develop a valid and reliable quantitative measure of science problem-solving performance. It was designed to specifically assess the elements of problem solving utilized by participants in this study. Additionally, a test that could be administered to an entire class simultaneously was needed since I had the dual role of researcher and teacher. Although some methods for assessing problem-solving performance were available in the literature, none adequately suited the specific assessment needs of this study. Most were measures of problem solving in mathematics, rather than science (Kilpatrick, 1967; Norrie, 1988; Schoenfeld, 1982). One was a measure of problem solving for early elementary students, rather than for late elementary students (Kanevsky & Rapagna, 1990). Another took a qualitative approach to assessing problem solving, rather than a quantitative one (Rashid, 1993). Ross and Maynes (1983) discussed the development of experimental problem-solving skills tests. These tests targeted the assessment of several science related skills (such as hypothesizing, collecting
data, etc.); however, only a sample multiple-choice question was provided in the article, not whole tests. The researchers reported that reliability and validity needed to be improved before the tests could be effectively used in research, although it was recommended for use as a supplementary assessment tool in the classroom. Thus, it was decided that a tailor-made measure of problem-solving performance was needed for this study.

**Analysis of the Research Process**

Several authors have reported on the research process itself, thereby providing the reader with detailed information on how the study was conducted (McGinn, 2000; Rashid, 1993). In the current study, the development of a problem-solving measure was considered significant and therefore it was thought that the process should be thoroughly and openly discussed. Other aspects of the research process were similarly considered to be of importance to the reader and were included in this discussion. Thus, another purpose of this study was to examine my thesis research experience at Brock University through a personal narrative approach.

**Statement of the Problem Situation**

**Quantitative Research Questions**

The preparation of this study revealed an absence of adequate research on problem solving through CAI for students of different logical reasoning abilities. This absence formed the basis for the two research questions addressed through quantitative analysis:

1. Does computer-assisted instruction improve students' problem-solving performance?
2. Does logical reasoning ability, as defined by Piagetian reasoning levels, have an
effect on this relationship?

Qualitative Research Question

In order to successfully answer the questions stated above, a customized measure
of problem-solving performance in science was required. Thus, a related purpose of this
thesis was to design, validate, and determine the reliability of a science problem-solving
performance measure for use in this study. To provide the reader with critical insight into
the development of this measure, as well as other significant aspects of the research
process, my entire research experience was qualitatively analyzed; this formed the basis
for the third research focus:

3. What would an analysis of my thesis research experience in the Faculty of Education
at Brock University reveal?

Importance of the Study

There were several reasons why this study was conducted.

1. To make a contribution to the literary debate on the impact of CAI by purposely
   addressing and avoiding previously occurring research flaws.

2. To contribute to a need stated in the literature for further research on the effects of
   logical reasoning ability on problem solving through the use of CAI.

3. To contribute to the understanding of elementary school students' problem-solving
   performance.

4. To develop a valid and reliable pretest, posttest, and scoring scheme to assess
   problem-solving performance in science.
5. To provide insight into the research process of a Brock University M.Ed. thesis for future M.Ed. thesis students and other readers.

6. To provide provincial educators with valuable information on the effectiveness of two teaching methods (i.e., computer- and laboratory-based instruction) for the new Ontario Science and Technology curriculum (MoET, 1998).

Definition of Terms

**Computer-assisted instruction (CAI)**, also called computer-aided instruction, is defined as the use of computers in conjunction with regular classroom instruction under the direction of the teacher (Tiffin & Rajasingham, 1995).

**Computer-based education (CBE)** is defined here as an all-encompassing term referring to the wide range of technological tools available (e.g., word processing programs, spreadsheets, hypermedia, and probes) and the wide range of methods for using these technological tools.

**Logical reasoning ability** is defined by Piagetian reasoning levels (i.e., concrete, transitional, and formal-operational). It has been operationally defined in this study as student score on the Group Assessment of Logical Thinking (GALT).

**Problem solving** is defined by Polya (1980) as:

> finding the unknown means to a distinctly conceived end. To solve a problem is to find a way where no way is known off-hand, to find a way out of a difficulty, to find a way around an obstacle, to attain a desired end, that is not immediately attainable, by appropriate means. (p. 1)

**Problem-solving heuristic** has been defined as a “general suggestion or strategy, independent of any particular topic or subject matter, that helps problem solvers
approach and understand a problem and efficiently marshal their resources to solve it” (Schoenfeld, 1980a, p. 9). Polya’s (1957) problem solving model made use of many heuristics such as, write down the problem, draw figures, relate an unknown problem to a known problem, and check the result. In this study the problem-solving heuristics, or strategies, given to the students in the problem-solving lessons were:

1. Define the problem: What is the root of the problem?
2. Develop a plan: What do I need to know? How am I going to find this information?
3. Gather information: How am I going to record this information?
4. Sort-link-solve: How does this information help solve the problem?
5. Verify the solution: Does the answer make sense? Can I explain my solution? Does my solution agree with the evidence?

Problem-solving performance is the demonstration of the problem-solving process. It has been operationally defined in this study as score on the pretest and score on the posttest, as well as problem-solving improvement score (i.e., the difference between pretest and posttest scores).

Sophisticated CAI is more exploratory in nature than traditional CAI (Pek, 1996; Tiffin & Rajasingham, 1995).

Traditional CAI is usually “drill and practice” in nature; namely a computer-generated question requires a student response which is followed by immediate feedback on the accuracy of the response (Berger et al. 1994; Pek, 1996).

Scope and Delimitations of the Study

This study was based on results obtained from 84 grade seven students attending one elementary school. Thus, results may only be applicable to students with similar
demographics. Additionally, results may have been improved by the use of a larger sample size. Another limitation was the use of only one software program. Thus, results obtained in this study may only be generalizable to software programs with similar strategies and objectives (Krajcik et al., 1986). Similarly, only one laboratory activity was used in the study. Perhaps a different form of traditional instruction would reveal different findings. In addition, it was assumed that four lessons in CAI or traditional instruction was a long enough opportunity to allow for the development of student problem-solving skills. A significant limitation of the results drawn from this study arose from the fact that the problem-solving measure was newly developed. Further experimentation with this measure is recommended before strong conclusions can be drawn. Another potential limitation was my combined role as teacher and researcher; although no conflicts were noted. Qualitative analysis made use of only one data source (my journal notes) and only one participant (myself). Future study might include multiple data sources and/or several participants for a diversified analysis of the thesis research experience at Brock University.

Considering these limitations, the aim of the current study was to provide some understanding into the dynamic relationship of problem solving, CAI, and logical reasoning ability for other educators and researchers. It was also intended to provide some insight into the thesis research experience of a Master of Education student at Brock University.

Outline of the Remainder of the Thesis

Chapter One was an introduction to the thesis; namely, a look at background information on the problem, a statement of the problem situation, the importance of the
study, definition of terms, and the scope and delimitation of the study. Chapter Two is a review of literature related to the quantitative and qualitative questions of the study. Specifically, this is a review of the research on instructional methods in science education; the use of computers in science education; problem solving; the relationship between problem solving and CAI; logical reasoning ability; the relationship among problem solving, CAI, and logical reasoning ability; and analysis of the research process. Chapter Three is a description of the quantitative research analysis. It begins with a description of the methodology including the selection of computer software, selection of participants, instrumentation, procedure, pilot study, research design, data collection, and data analysis methods. The methodological assumptions, limitations, establishment of credibility, and ethical considerations are then presented, followed by the quantitative research findings. Chapter Four focuses on the qualitative analysis. It reports on the qualitative methodology including its limitations and credibility, and then reports the qualitative findings in the form of a personal narrative. Chapter Five is a discussion of the findings of this study, as well as their implications for theory, practice, and future research.
CHAPTER TWO: REVIEW OF RELATED LITERATURE

Introduction

Problem solving is recognized to be an important component of science and mathematics education, as is the use of information technology in the classroom. The previous chapter revealed a need for further research in these areas and an examination of their possible connection to Piagetian reasoning levels. This chapter is an expansion on the related literature reviewed for the current study. It begins with a look at research related to the quantitative component of the study, including instructional methods in science education; the use of computers in science education; problem solving; the relationship between problem solving and CAI; logical reasoning ability; and the relationship among problem solving, CAI, and logical reasoning ability. It is followed by an examination of research related to the qualitative component of the study; namely, the analysis of the research process.

Instructional Methods in Science Education

Traditionally, problem-solving tasks in science have been textbook- and laboratory-based. Furthermore, textbook tasks have been used more often in the science classroom because of the demands of the laboratory experience, such as time, space, and cost (Kinnear, 1986). The downfall is that textbook problems solved through the use of the paper-and-pencil method are often narrowed to reinforce only the topic at hand since "problems are usually physically arranged in association with chapters in a textbook or temporally arranged in association with a topic of study" (Kinnear, p. 3). In textbook-based problems the data are immediately available, as compared to the laboratory-based problem in which a student is asked to gather the data before attempting to solve the
problem. Alternatively, laboratory-based problems, although "hands-on" experiences, are often critiqued for following prescribed rather than investigatory or exploratory methods (Jungck, 1991; Kinnear). Thus, these traditional instructional modes have been criticized for their limitation in teaching students problem-solving skills which are transferable to new and/or authentic situations (Pek, 1996; Soderberg et al., 1994). Additionally, in traditional instruction students are often passive learners in teacher-directed lessons and the emphasis is placed on the "ends" rather than the "means" (Pek).

Current approaches to teaching science have moved away from a teacher-directed approach toward a more student-centered approach. According to instruction based on constructivism-informed ideas, the teacher becomes "someone who facilitates conceptual change by encouraging pupils to engage actively in the personal construction of meaning; and in order to do this there needs to be frequent opportunities for pupils to make their ideas explicit and to communicate them" (Driver & Oldham, 1986, p. 116). Thus, knowledge is actively created by different students within different contexts (Baker & Piburn, 1997). Recent research focusing on CAI, as well as research focusing on problem solving, often make use of such constructivism-informed ideas.

Computers in Science Education

The literature on computers in education uses several descriptive terms interchangeably and usually without providing clear statements of their meanings. For example, computer-based problem, computer exploratorium, computer simulation, computer-based instruction, and instructional software are some of the ambiguous terms found in the literature. For the purposes of this study computer-based education (CBE) will be defined as an all-encompassing term referring to the wide range of technological
tools available (e.g., word processing programs, spreadsheets, hypermedia, and probes) and the wide range of methods for using these technological tools. Within CBE, computer-assisted instruction (CAI) has been more specifically defined as the use of computers in conjunction with regular classroom instruction under the direction of the teacher (Tiffin & Rajasingham, 1995). In CAI, the computer facilitates the interplay among four factors of instruction: teaching, learning, knowledge, and problems (Tiffin & Rajasingham). Over time, CAI has evolved from a basic, traditional form to a more sophisticated one.

In traditional CAI students learned through computer programs which were primarily drill and practice in nature. In other words, the computer presented a question, the student typed in an answer, and then received feedback about the accuracy of his/her response (Pek, 1996; Tiffin & Rajasingham, 1995). If the student was successful then the process was repeated with a new question. If the problem was not solved correctly then more explanation of the knowledge needed and/or problem was given to the student before he/she reattempted to solve the question (Tiffin & Rajasingham).

Currently more sophisticated forms of CAI are available, which better complement constructivism-informed instruction by offering authentic, active, meaningful, and individualized learning (Berger et al. 1994; Eylon, Ronen, & Ganiel, 1996; Jungck, 1991; Krajcik et al., 1988; Marsh & Kumar, 1992; McGrath, 1998; Pek, 1996; Rashid, 1993; Soloway, Pryor, Krajcik, Jackson, Stratford, & Wisnudel, 1997; Weber, 1993). For example, sophisticated CAI is more exploratory in nature than the traditional form, allowing students to personally direct their own learn at their own pace (Pek; Tiffin & Rajasingham, 1995). In addition, students can manipulate audio-
visual icons, guided examples, and hints as many times as they are needed in the learning process (Pek). As well, the use of hypertext allows students to branch from one screen to the next to gather information at their own pace (Pek). Finally, sophisticated CAI allows students to do tests and receive immediate feedback, as well as to allow for teachers to monitor this progress (Pek).

Phases of Research of CBE in Science Education

Three phases of research were noted by Berger et al. (1994) after their exhaustive review of the literature on the use of information technology in science education. These trends were labeled traditional research, current research, and new research in information technology.

The first phase, labeled "traditional research in information technology", consisted primarily of comparison studies. More specifically, researchers sought to determine if more learning was accomplished in the same amount of time, or if less time was required to accomplish the same amount of learning, with the use of technology. This effect was often assessed by the use of box-scores. In other words, the number of positive effects versus negative effects versus no effects found in a study, or between many studies, were counted and compared. Effect sizes were then used to statistically measure the impact of using information technology in the classroom. These traditional studies examined several aspects of CBE. Some examined the effectiveness of drill and practice or tutorial forms of CAI. Many comparison studies were also done on computer-managed instruction (i.e., instruction which provided students with a computer evaluation of their performance). Additionally, microcomputer-based laboratories, which incorporated a computer-based, data-collecting device in the science laboratory, were
traditionally studied. Finally, comparison studies on the effectiveness of computer-simulated experimentation were conducted.

The second phase in information technology research was labeled by Berger et al. (1994) as "current research". Essentially, they noticed a shift from traditional comparative studies to studies on the new uses of information technology. For example, cognitive and motivational aspects of learning, as well as constructivism-informed learning theory, were emphasized more in current research than they were traditionally (Berger et al.; Soloway et al., 1997).

The third phase in research on information technology was labeled by Berger et al. (1994) as "new research". Studies in this phase focused on new and promising research which the authors claimed may be indicative of a paradigm shift in research on information technology. More specifically new research was conducted on both the content and the method of teaching and learning through complex computer instructional systems. This included such systems as hypermedia; telecommunication projects like WaterNet, MIX (McGraw-Hill Information Exchange), and Kids Network; as well as Intelligent tutoring systems (Berger et al.; Marsh & Kumar, 1992; Pek, 1996). Additionally, the use of information technology and its relationship to problem solving in science education has become a focus of new research (Berger et al.; Eylon, et al., 1996; Rashid, 1993; Slotnick, 1989; Soderberg et al., 1994).

Arguments For and Against CBE

An examination of the literature revealed positive and negative opinions on CBE. The following is a summary of the arguments for the use of CBE, as a supplement to teacher instruction (i.e., not as a replacement for it). First, learning has been found to be
more efficient with the use of CBE than without it for learners of all ages (Berger et al., 1994). For example, Soderberg et al. (1994) reported that an extensive study involving many American colleges, universities, community colleges, and high schools revealed that "the majority of faculty who used GCK [Genetics Construction Kit software] felt that student learning was vastly enhanced by problem solving with GCK as compared with previous courses where they had not used GCK" (p. 78). Likewise, Weber (1993) reported that "although little has been done to show exactly why computers help students learn more, retain more or learn the same amount faster, studies have been conducted to show that the gains made by students who have access to computers are very evident" (p. 5). Secondly, it has been reported that CBE has a significant positive effect on students' achievement and understanding as well as their attitude toward the subject, computers, and self (Berger et al.; Eylon et al., 1996; Marsh & Kumar, 1992; Pek, 1996; Soderberg et al.; Soloway et al., 1997; Weber). For instance, in an analysis of interviews with 12 teachers one key theme found by McGrath (1998) was that "technology increases student motivation, and motivated students are more receptive, more engaged, and more likely to learn" (p. 58). A third reported advantage of CBE is its potential to meet constructivism-informed ideas by being authentic, active, meaningful, and by offering individualized learning (Berger et al.; Eylon et al.; Jungck, 1991; Krajcik et al., 1988; Marsh & Kumar; McGrath; Pek; Rashid, 1993; Soloway et al.; Weber). For example, the Genetics Construction Kit (GCK) software program allows the user "to generate problems that are similar to realistic research problems where neither the student nor the teacher knows what the 'answer' is" (Soderberg et al., p. 74). The authors claimed that this feature of GCK promoted "meaningful, rather than rote learning of genetics" (p. 74). Fourthly, CBE
frees the teacher to provide more in-depth help to, and to collaborate with, students (Jungck; Slotnick, 1989; Soderberg et al.; Soloway et al.). Another theme discussed by McGrath (1998) from an analysis of interviews with 12 teachers was that “in classrooms with computers, conversations between teachers and students and among students themselves become deeper and more probing” and that “technology use encourages teacher-as-facilitator approaches” (p. 59). The fifth reported advantage is that computer simulations can effectively enhance the laboratory experience, or provide some laboratory-like experience in situations where the “real thing” is not available due to practical restraints such as time and cost (Eylon et al.; Kinnear, 1986; Krajcik et al. 1988; Marsh & Kumar; Soloway et al.). This is illustrated in a quote from Soderberg et al., who stated that “appropriate use of computer simulations can insure more efficient use of lab time and can provide students with experience in handling large data sets. Or, in the case of many large universities which have been forced to drop their lab components from biology courses, computer simulations can provide students with at least some laboratory-like experience” (p. 70). This is a sampling of the arguments for the use of computers in the classroom.

On the other hand, there are arguments against the effectiveness of CBE. For example, in some situations (such as in the use of Hypermedia and the Internet) it was reported that the inexperienced learner became overwhelmed with information resulting in cognitive overload and “poorer integration of new information than would have been obtained with a traditional method” (Marsh & Kumar, 1992, p. 36; Robertson, 1998). As well, some critics claim that research has not shown a significant increase in achievement or attitudes in CBE over traditional instructional methods. For example, Robertson
reported that "edu-tech's critics" believe that "there is no independent research that meets generally accepted standards for controls, accuracy, transparency, replicability, sample size, and validity. There are claims, inferences, arguments, and assertions. And, of course, there are promises" (p. 136). This quote supports a common argument against the effectiveness of CBE, that is a critique of research design in these studies. For instance, Eylon et al. (1996) found that computer simulation helped some aspects of student learning, but at the same time acknowledged that these effects might have been attributed to enriched concept instruction and not to the computer simulation alone. In another example, the 1985 NARST convention reported that, although there was an increase in the amount of research done on the effectiveness of CAI in science education, the research could be critiqued for having the following problems (Krajcik et al., 1986):

1. Excessively broad generalizations were reported.
2. Inappropriate criterion measures were used.
3. Criterion measures were sometimes not specified.
4. Software assessed was, generally, less than optimal.
5. The research questions lacked substantive theoretical base.

To summarize, the literature presented conflicting views on the effectiveness of CBE and critics have claimed that these conclusions were not always reached by justifiable means.

Problem Solving

Problem solving is recognized as an important component of science and mathematics education, consequently it has been the focus of many studies (Kilpatrick, 1967; Kinnear, 1986; Krajcik et al., 1988; Murphy & McCormick, 1997; Polya, 1957; Schoenfeld, 1982; Simmons & Lunetta, 1993; Slack & Stewart, 1990; Soderberg et al.,
1994; Weber, 1993). Some researchers believe that problem-solving skills are as important as content in preparing students to “recognize, address and shape new scientific and technological changes” (Friedler et al., 1989, p. 58). Similarly, Soderberg et al. stated that problem solving, along with problem posing and peer persuasion, were vital to the teaching of biology, as it was “central to a deep understanding of how scientific knowledge is created, modified, and used” (p. 71).

Researchers have stated that problem solving is context-specific, active, flexible, inventive, and authentic (Murphy & McCormick, 1997; Slotnick, 1989). Furthermore, to be authentic the problems must be both “personally meaningful and purposeful within a social framework” (Murphy & McCormick, p. 463). Some researchers also suggested that successful problem-solving performance required explicit teaching (such as the use of heuristics), and student-teacher discussions about the decisions made and the conclusions drawn (Kinnear, 1986; Murphy & McCormick; Weber, 1993).

Polya (1957) defined the term heuristic “as an adjective, means ‘serving to discover’” (p. 113). The term heuristic has also been defined as the “general suggestion or strategy, independent of any particular topic or subject matter, that helps problem solvers approach and understand a problem and efficiently marshal their resources to solve it” (Schoenfeld, 1980a, p. 9). Kilpatrick (1967) defined it as “any device, technique, rule of thumb, etc. that improves problem-solving performance” (p. 19). Likewise, Slotnick (1989) defined heuristics as “rules of thumb which combine past experience, intuition, and an analysis of the current situation” (p. 198). He went on to say that most problems we face in day-to-day life were solved in this way.
The four main steps of Polya’s foundational problem-solving model were: (1) understand the problem, (2) devise a plan, (3) carry out the plan, and (4) look back. Many other heuristics were embedded within these steps such as, write down the problem, draw figures, relate an unknown problem to a known problem, and check the result. Similarly, Slack and Stewart (1990) suggested the use of strategies to promote success in problem-solving performance such as, “predicting the data, redescribing a problem qualitatively, generating and testing hypotheses, considering alternative hypotheses, and checking results” (p. 66). Researchers have reported that science experts followed these heuristics for a successful problem solving experience (Simmons & Lunetta, 1993):

(a) Analyze the task carefully.
(b) Approach problem solving in a hierarchical fashion.
(c) Describe main components of the problem before considering the details.
(d) Select a general qualitative method before solving the problem.
(e) Plan more completely before solving the problem.
(f) Use procedural knowledge with greater efficiency.
(g) Justify solving with related concepts and/or principles.
(h) Make frequent checks on the consistency of their answers. (p. 154)

In contrast, novice and unsuccessful problem solvers were reported to have done the following (Simmons & Lunetta):

(a) Sort problems on the basis of surface features (such as terminology).
(b) Use more trial and error approaches.
(c) Use a backward working process (a means-ends analysis).
(d) Use goals and subgoals to direct the problem-solving search.

(e) Perform frequent tests to check the sequence of actions. (p. 154)

In summary, many researchers encouraged the use of effective heuristics by teachers and students to promote success in student problem-solving performance (Friedler et al., 1989; Kilpatrick, 1967; Norrie, 1988; Polya, 1957; Schoenfeld, 1980b; Schoenfeld, 1982; Slack & Stewart, 1990; Slotnick, 1989).

**Measures of Problem Solving**

Various methods for measuring problem solving performance were reported in the literature. Some researchers assessed problem solving qualitatively. For example, Rashid (1993) assessed the improvement of problem-solving strategies and attitudes in students with learning difficulties. Changes in problem-solving were noted through qualitative analysis of informal interviews with eight student participants.

Other researchers measured problem solving quantitatively. For instance, Kanevsky and Rapagna (1990) compared learning and problem solving between high and average IQ students (aged 4 to 8). The problem-solving tasks used in this study were two modified versions of the Tower of Hanoi puzzle. One was a computer simulation game called Layer Cake and the second, Monkey Can Puzzle, required the physical manipulation of cans in the solution process. The participants were videotaped when solving each puzzle for subsequent analysis. Three elements of problem solving were then selected for scoring during the analysis: “(a) the number of trials needed to achieve the criterion of success in each phase; (b) the number of unassisted, correct moves executed, and (c) the number of hints used on each experimental trial (not on warm-ups)” (p. 22).
Some researchers measured problem-solving performance specific to mathematics. For example, Norrie (1988) outlined "the development, implementation and evaluation of problem solving objectives within a large public school board" (p. 26). To accomplish the evaluation goal of the project, two mathematical problem-solving tests and two mathematical skill application tests for grade 4, 5, and 6 students were administered. A holistic scoring scheme was then applied to these tests by two independent scorers and checked by a third researcher. This scoring method is summarized in Table 1 (Norrie, p. 27).

Another example of mathematical problem-solving assessment was found in Kilpatrick’s (1967) study. Kilpatrick devised and tested a system for coding think-aloud protocols created by participants as they solved mathematical word problems. The think-aloud method is one which requires participants to talk through the problem-solving process. The first draft of Kilpatrick’s (1967) problem-solving checklist was based on the foundational work of Polya (1957). Several revisions resulted in two final checklists of eight items each, to summarize subjects’ problem-solving approaches. The eight items in the first checklist were: drawing a figure, using successive approximations, questioning the existence or uniqueness of a solution, expressing uncertainty about his/her final solution, saying he/she did not know how to solve the problem, requesting assistance or more information, admitting confusion, and the number of 'careless' or structural errors the subject made. The second assessment checklist found eight patterns with high interrater reliability: “deduction processes, use of equations, trial-and-error processes, reading, checking solutions, structural errors, difficulty in performance, and stops without a solution” (Kilpatrick, p. 99).
Table 1

Holistic Scoring Method for Problem-Solving Tests

<table>
<thead>
<tr>
<th>Problem-Solving Objective</th>
<th>Score</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Understanding the problem</td>
<td>0</td>
<td>-no understanding, complete misinterpretation</td>
<td>-partially misinterpreted</td>
</tr>
<tr>
<td>S-Strategy or planning</td>
<td>1</td>
<td>-no strategy or inappropriate plan</td>
<td>-partially correct procedure -part of data interpreted correctly</td>
</tr>
<tr>
<td>A-Answer to the problem</td>
<td>2</td>
<td>-no answer or a wrong answer from an inappropriate plan</td>
<td>-copying error; computational error; or answer labeled incorrectly</td>
</tr>
</tbody>
</table>
In another study, Schoenfeld (1982) developed and tested three practical measures of mathematical problem solving, to be used in place of protocol analysis, which was stated to be time-consuming and costly. Measure 1 was a quantitative coding of pretest and posttest results on a mathematical paper-and-pencil test via two different scoring methods. The first was ‘Multiple Count Scoring’ which credited students for generating multiple approaches to problems and the second was ‘Best Approach Scoring’ which only assessed a student’s best attempt. Schoenfeld labeled Measure 2 as a qualitative assessment of students’ problem-solving behaviour, although the data was converted from this measure into percentages suitable for basic quantitative comparisons. Finally, Measure 3 assessed student ability to generate and transfer general approaches to new and unfamiliar problems (i.e., heuristics). Schoenfeld concluded that the two scoring versions of Measure 1 were roughly equivalent, both having approximately 90% reliability, and the reliability for Measures 2 and 3 were not given.

Ross and Maynes (1983) reported on their progress in the development of paper-and-pencil instruments used to measure experimental problem-solving skills in grade seven and eight students. The tests were made to assess seven skills used by successful scientists when designing an experiment, “developing a focus (i.e., formulating a hypothesis), developing a framework (i.e., designing an experiment), judging the adequacy of data collected, recording information, observing relationships in data, drawing conclusions, and generalizing” (p. 65). Whole problem-solving tests were not presented in the paper, but the authors stated that a multiple choice format was used and an example question was given. The item pool was field tested in three phases with teachers and classes of students. In one of these phases, they found that the multiple
choice questions did not have a high correlation with similar open-ended questions. Ross and Maynes concluded that "in the multiple-choice instrument, students are required to select/recognize the most appropriate response, whereas in the open-ended items students are required to generate the response. In this sense the open-ended measures have greater similarity to classroom activities than the multiple-choice tests" (p. 72). However, they made no mention of revising the tests to the open-ended format. They went on to conclude that since reliability was less than optimal and validity had some unresolved concerns, the tests should be used to measure experimental problem-solving skills with some reservations, but they should not diminish the value of the tests to classroom teachers. In other words, the authors suggested that the tests might be used in conjunction with other classroom assessment methods to evaluate problem-solving performance. Continued experimentation in school environments was recommended to improve reliability and validity of the tests.

Therefore, a review of the literature revealed that several methods were developed and used by other researchers to specifically assess problem-solving performance according to the needs of each study.

Problem Solving and Computer-Assisted Instruction

Researchers suggest that computers have the potential to facilitate successful problem-solving performance (Eylon et al., 1996; Friedler et al., 1989; Jungck, 1991; Krajcik et al., 1988; Marsh & Kumar, 1992; McLean, 1989; Pek, 1996; Rashid, 1993; Simmons & Lunetta, 1993; Slack & Stewart, 1990; Soderberg et al., 1994). For example, Slotnick (1989) examined educational software such as PsychWare simulations, experiments, and tutorials developed at the New York Institute of Technology, as a
means of enhancing problem solving, critical thinking, and scientific reasoning. In each of these programs there was a mystery which called for a solution to be achieved through logical manipulation of its parts. It was concluded that these programs were effective when used alone, or to supplement lectures and text presentation, in introductory college psychology courses. Based on these findings, Slotnick summarized the advantages of CAI on problem solving as follows:

Educational software can create computer microworlds as an effective medium for helping students to construct and make explicit their own problem-solving solution path. Problems can be represented graphically; students can make numerous solution attempts; they can manipulate the problem environment and try alternate solutions. The program can track student activities; suggest different strategies; and help construct new solution paths. The environment we create with the computer will allow students to focus on the process of solving problems, thereby learning a method, an approach, not just a solution to one type of problem. (p. 194)

Slotnick also suggested that computer software programs can offer real world simulations, thereby making the problem-solving process more authentic.

A research paper presented at the 1986 Annual Meeting of the American Educational Research Association by Kinnear found a positive effect of CAI on problem-solving performance. Kinnear compared the use of computer- and textbook-based problems in an introductory college-level genetics course. All participants (N = 68) were first required to solve dominance-type genetics problems (one from a textbook and one from the CATLAB software program). After each task, students were required to submit
a written record of their work and to note the time taken to complete the problem. Students were also given a brief questionnaire which asked: (i) which, in their view, was the easier problem (i.e., computer- or textbook-based) and why; and (ii) whether or not, the computer-based problem assisted in their understanding of genetics. Secondly, students were asked to solve linkage-type genetics problems (one textbook- and two computer-based) and were asked to rate their confidence in solving the textbook problem and in solving the computer problems.

It was found that on the dominance-type tasks students reported to have spent more time on the CATLAB problem ($M = 42$ minutes), than the textbook problem ($M = 3.8$ minutes). Surprisingly, 73.5% of the students reported that the computer problem was easier than the textbook problem, even though it took considerably longer to solve. Students most commonly expressed visualization (i.e., see, observe, visualize) as the reason for identifying the computer problem as easier. Kinnear (1986) concluded that students perceived a task as easier if it resulted in more meaningful learning. As well, it was concluded that students who could not reach any conclusions, or had obtained incorrect conclusions, did so because of their difficulty with the problem-solving process itself and/or in understanding the genetics concepts. Another significant finding of this study was that some students (31% in the dominance tasks and 20% in the linkage tasks) were successful in solving textbook-based problems, but were unsuccessful in solving computer-based problems. Kinnear (1986) attributed these results to the increased need for conceptual knowledge in CATLAB as compared to the minimal knowledge required for textbook problems. In other words, it was suggested that the computer program extended the boundaries of what traditional learning had prepared them to solve. Kinnear
did not consider the possibility that computer-based problems may not have been as conducive to successful problem solving as textbook-based problems. A comment about this research design is that the comparison between textbook-based problems and computer-based problems may not have been a fair one. It is possible that through the use of a more active instructional technique in the control group, such as a hands-on laboratory activity, a more equitable comparison might have been made.

In another study, Pek (1996) tested the usefulness of CAI software utilized by Diploma in Materials Engineering students \(N = 40\) through two separate experiments. In the first experiment, the control group received regular lecture-tutorial-laboratory teaching and the experimental group received the same instruction supplemented with the software. This comparison may be considered unfair, as the experimental group receive unequal treatment as compared to the control group. Perhaps a more reasonable comparison could have been made if the control group had also received supplemental instruction that was not computer-based. However, from this test Pek suggested that CAI raised student understanding and test performance. In the second experiment, Pek tested the effect of explicit teaching of independent learning skills on student understanding of concepts, through the use of CAI. From this study it was concluded that the teacher’s involvement was essential to the successful implementation of CAI. In both of these experiments no statistical analyses were undertaken and conclusions appeared to be drawn from a graph of the results. As well, Pek concluded from these two experiments that CAI improved problem-solving skills, but the method for assessing problem-solving performance was not specified.
In a study by Eylon et al. (1996) the effect of a computer-simulated optics program was examined with grade ten students ($N = 75$). The authors made use of two control groups and an experimental group. The experimental group solved problems on the computer simulation program and used a concept enrichment workbook. One control used the concept enrichment workbook and solved paper-and-pencil problems, and the other control group solved paper-and-pencil problems only. The pretest and posttest assessed conceptual understanding rather than problem-solving performance. The authors concluded that solving problems with the software program enhanced student understanding of concepts, although they acknowledged that these effects might have been attributed to enriched concept instruction and not to the simulation alone.

Other researchers have similarly reported on factors that affect the relationship between CAI and problem-solving performance. Cypher's (1987) ethnographic study reported that educators played an integral role in determining the effectiveness of computer programs in the enhancement of thinking and problem solving. It was also suggested that a supportive environment was critical to fostering these skills.

McLean's (1989) case study of eight grade six students sought to qualitatively examine the entire relationship of problem solving and computer simulations. This analysis revealed five factors thought to influence student problem-solving performance in a highly interactive computer simulation program. These were students’ persistent use of guessing as a problem-solving technique, student reaction to the simulation, limiting student frustration, interaction in cooperative learning, and the role of the teacher.

Slack and Stewart (1990) also discussed factors influencing the relationship between CAI and problem solving. Specifically, it was concluded that “providing an
environment in which students are presented realistic genetics problems is not sufficient to elicit good problem-solving skills” (p. 65). They recommended that educators place an emphasis on conceptual knowledge and its link to problem solving in order to improve problem-solving skills. Likewise, other studies reported that further research is needed to reveal if problem-solving performance in CAI tasks is related to conceptual knowledge and/or logical reasoning skills (Berger et al., 1994; Krajcik et al., 1988; Simmons & Lunetta, 1993).

Thus, the literature suggested that CAI has the potential to be an effective means for developing student problem-solving skills and conceptual understanding. In quantitative studies finding a positive effect of CAI on problem-solving performance, some flaws in the research design were noted; namely an inequitable control group used for comparison and ambiguous means of assessing problem solving. Additionally, several factors were identified for having an effect on the relationship between CAI and effective problem-solving, such as the role of the teacher and level of conceptual understanding.

Logical Reasoning Ability

Piaget’s theory of intellectual development spans from infancy through adolescence; people are thought to pass through a sequence of four stages in the development of logical reasoning ability: sensorimotor, preoperational, concrete-operational, and formal-operational. The stages that are most relevant to the participants in this study are the concrete and formal stages of development. In the concrete-operational stage people between the ages of seven and eleven “come to understand reversibility, identity, compensation, and numerous other logical concepts in concrete tasks” (Siegler & Richards, 1982, p. 912). Piaget theorized that people in the concrete
stage are unable to think abstractly because “hypothetical reasoning divorced from or contrary to experience” is thought to be beyond their ability (Siegler & Richards, p. 913). On the other hand, Piaget theorized that by the age of 12 or 13 most children have reached the formal-operational stage of development, although it was acknowledged that there are some people who never reach this stage. People at the formal-operational level are able to “apply their new abstract reasoning abilities to all areas of life” and “in short, they possess the reasoning abilities of educated adults” (Siegler & Richards, p. 913).

Additionally, as people pass from one stage of development to another they are said to be in the transitional stage. Several criticisms have been made against Piaget’s theory of intellectual development, including “predictions about the typical developmental sequence, the consistency of children’s performance across tasks and over time, and the possibility of teaching relatively young children to understand complex concepts” (Siegler & Richards, p. 915).

In accordance with the critics, some researchers have reported that many late elementary school students are functioning below the formal-operational level, as might be predicted by Piagetian theory (Bitner-Corvin, 1988; Gipson et al., 1989; Newton, Capie & Tobin, 1981; Padilla, Okey, & Dillashaw, 1983; Smith & Sims, 1992; Staver, 1984; Tobin & Capie, 1981). Accordingly, these students have difficulty completing formal-operational tasks and do poorer on concrete tasks in comparison to those with formal reasoning ability (Newton et al.). This notion should be of concern to educators, since students may not be capable of successfully completing the tasks that are required of them in the classroom because of developmental limitations. It was therefore recommended that teachers take a constructivism-informed approach to instruction. In
other words, they should be aware of differing levels of cognitive development among students and attempt to use various instructional methods to accommodate this variation (Gipson et al.; Linn & Songer, 1991; Newton et al.; Smith & Sims; Staver, Tobin & Capie; Yap & Yeany, 1988). For example, it was suggested that students who have difficulty with formal concepts should be given the opportunity to use hands-on, direct, and concrete learning methods as much as possible (Smith & Sims). It was also suggested that students should generate the empirical data needed (Smith & Sims).

**Logical Reasoning, Problem Solving, and CAI**

A relationship between level of logical reasoning ability and success in problem solving was noted in the literature (Gipson et al., 1989; Smith & Sims, 1992; Stewart & Hafner, 1994). However, formal reasoning ability in itself did not predict problem-solving success because other variables, such as content knowledge and mental capacity, also play a role (Gipson et al.; Krajcik et al., 1988; Linn & Songer, 1991; Simmons & Lunetta, 1993; Smith & Sims; Stewart & Hafner). Similarly, researchers have reported a connection between science process skills and logical reasoning abilities (Padilla et al., 1983; Yap & Yeany, 1988). These process skills included formulating a hypothesis, designing an experiment, and making generalizations after collecting data. These science process skills are similar to those used by successful scientists when designing an experiment to solve a problem, "developing a focus (i.e., formulating a hypothesis), developing a framework (i.e., designing an experiment), judging the adequacy of data collected, recording information, observing relationships in data, drawing conclusions, and generalizing" (Ross and Maynes, 1983, p. 65). Therefore, prior research suggests a
relationship between science problem-solving skills and logical reasoning ability.

Seymour Papert’s (1980) foundational work with LOGO linked the variables of problem solving, logical reasoning, and CAI. The computer program LOGO required users to draw designs by directing a “turtle” around a computer screen (Krasnor & Mitterer, 1984). The purpose of LOGO was to “explicitly facilitate the learning of powerful ideas, skills, and heuristics which transcend the immediate task environment and can be applied in other problem-solving situations” (Krasnor & Mitterer, p. 133). Papert wrote that “the aim of AI (Artificial Intelligence) is to give concrete form to ideas about thinking that previously might have seemed abstract, even metaphysical” (p. 157-158). In other words, he hypothesized that computer simulations had the potential to make abstract ideas more concrete, thereby assisting students functioning at the concrete reasoning level. Slotnick (1980) supported the notion that computer models have the potential to facilitate the problem-solving process in this manner.

In an essay review, Krasnor & Mitterer (1984) took a slightly different viewpoint. They hypothesized that only formal operational students would be able to successfully learn LOGO and then transfer general problem-solving strategies to other situations. They went on to state that other variables affected the success of this transfer of skills. This included explicit teaching of heuristics and social interaction, for example. Thus, the literature suggested that the effect of CAI on problem-solving performance may be connected to a student’s logical reasoning ability, but further research was recommended in this area.
...
Analysis of the Research Process

To address the qualitative question of the current study, literature on two aspects of the research process were reviewed in this section. The first was the process of research as a form of problem solving in itself, and the second was the supervisor's role in graduate student research.

Research as a Form of Problem Solving

Previous researchers have reported on their research experience and have interpreted the process as a form of problem solving (McGinn, 2000; Rashid, 1993). Rashid organized a discussion of the research design process around the steps of Sternberg’s 1985 problem-solving model. These steps were:

1. Understanding the problem.
2. Identifying the task components.
3. Selecting a strategy for solving the problem.
4. Allocation of attentional resources.
5. Solution monitoring and evaluation. (Rashid, p. 47)

Each of these five problem-solving steps were used in the development of the author’s research design. For example, when discussing “Step 3: Selecting a strategy for solving the problem” two strategies that she had used in that step of the research process were identified as, “(1) developing an instructional program and (2) developing a plan for assessing the effectiveness of the program” (Rashid, p. 52).

McGinn (2000) applied qualitative research findings on problem solving to her own research notes and records. She reported that “For me, engaging in this research was a form of situated problem solving that entailed co-emergent construction of problems,
solutions, and solvers' identities; identities through resource-rich practices that make reference to visual representations, socially distributed expertise, exemplar problems and solutions, partial solutions, time, and anticipated audiences” (McGinn, p. 201). These elements of problem solving had emerged from an analysis of case studies and were then applied to the analysis of her own research experience. A specific example from McGinn’s study is as follows. Upon analysis of several case studies it was found that once participants became immersed in solving a predetermined problem, the problem would evolve. In later reflection, she reported that this process was evident in her own research as well, as various circumstances arose (such as research participants moving away during the course of the study) her overarching research question evolved. These studies were two examples that reported on how the experience of conducting research was considered to be a form of problem solving.

**Student-Advisor Relationship**

A qualitative study by Acker, Hill, and Black (1994) examined the supervisory relationship between graduate research students and their advisors. Two models of supervisory style were discussed. The first was the “technical rationality model” of supervision. In this model, the thesis advisor was likened to “a manager or director, keeping the students motivated and on track, providing timetables, and guidelines” (Acker et al., p. 484). Accordingly, the thesis student was seen as “a relatively passive participant” who needed structure (Acker et al., p. 484). The second model discussed was the “negotiated order model” of supervision. In this model “the supervisor’s task becomes one of facilitating rather than directing” and “the student, like the supervisor, participates fully in negotiating and interpreting meanings” (Acker et al., p. 485).
Analysis of these models was based on data collected through semi-structured interviews with students, supervisors, and other ‘key persons’ (such as administrators and secretaries). The focus was on describing three aspects of the supervisory relationship: the tutorial, the level of direction provided by the supervisor, and the interpersonal relationship between student and advisor. The tutorial essentially consisted of “reviewing the student’s work, discussing problems that arose, and planning future work” (Acker et al., 1994, p. 487). It was suggested that this aspect of the relationship was often negotiated between the student and advisor. The degree of direction varied among participant students and advisors. Some viewed the advisor as highly directive and others had a non-directive style. However, the authors concluded that “in line with the negotiated order model, supervisors adjusted their practices to match the perceived student need and stage of the research” (Acker et al., p. 492). The nature of the relationship between the student and the advisor also varied among participants. It was stated that “the appropriate degree of intimacy is negotiated in each relationship” (Acker et al., p. 494). Furthermore, it was recommended that advisors and students work to achieve a relationship that is satisfactory to both people. Thus, Acker et al. concluded that the supervisory relationship was more accurately described by the negotiated order model rather than the technical rationality model. However, the technical model was commended for its ability to fulfill a need for supervisors to “insert some order and control into the process” (Acker et al., p. 496).

Therefore, two important trends were found in the literature on the analysis of the thesis research process. First, that it can be described according to a problem-solving
model, and with respect to graduate student research, the role of the supervisor is often a negotiated one.

Chapter Summary

In this chapter, extensive background information related to the questions posed in the quantitative component of the current study was reviewed first. To recap, it was revealed that instructional methods in science education have moved away from traditional teacher-directed methods towards methods based on constructivism-informed theory, such as context-specific, active, flexible, inventive, and authentic learning. It was also suggested that CAI may be an effective method to promote such individualized learning. However, a review of the use of computers in science education presented conflicting views on its effectiveness; most notably critics have claimed that conclusions on the effectiveness of CBE were not always reached by justifiable means. A constructivism-informed approach to instruction has also been linked to successful problem-solving performance. Thus, researchers discussed the potential for CAI as an effective means to develop student problem-solving skills. However, careful scrutiny of quantitative studies finding a positive effect of CAI on problem-solving performance revealed some flaws in the research design; namely an inequitable control group used for comparison and an ambiguous means of assessing problem solving. Several factors believed to play a role in the relationship between CAI and effective problem-solving, such as social interaction and the teacher, were identified. It was suggested that logical reasoning ability may also be one of these factors, but further research was recommended. Researchers have found that many late elementary students are in the concrete stage of logical reasoning development and hence have difficulty with high-
level reasoning tasks. It was recommended that teachers take a constructivism-informed approach to instruction to accommodate this difference in logical reasoning ability, and that CAI may be one effective method by which to do so.

A review of literature related to the research process provided background information on the qualitative research question. First, it was noted that researchers used a range of methods to assess problem-solving performance in order to meet the specific needs of their research. Secondly, it was revealed that authors have likened the research process to the problem-solving process. A third finding relevant to the qualitative component of the current study was that in graduate student research the role of the supervisor is often a negotiated one.

The next chapter focuses on the quantitative methodology and findings of the current study, with reference to the literature discussed here.
CHAPTER THREE: QUANTITATIVE ANALYSIS

Introduction

Related research reviewed in the previous chapter provided some of the background information needed to establish the methodology and procedures of the current study. This chapter is divided into two main sections. It begins with a discussion of the quantitative methodology used to analyze the relationship among student problem-solving performance, logical reasoning, and CAI in science education. This includes a discussion on the selection of computer software and participants. Next, the instruments used for quantitative assessment are described; one was a standardized test and the other was developed over the course of this thesis. Details of a pilot study conducted prior to the commencement of the main study, as well as the procedures employed in the main study are then highlighted. Methodological assumptions and limitations of the study are openly discussed and credibility is established. Finally, ethical considerations arising from this research are listed. In the second part of this chapter quantitative findings are presented. To recap, these relate to the following research questions:

1. Does computer-assisted instruction improve students' problem-solving performance?
2. Does logical reasoning ability, as defined by Piagetian reasoning levels, have an effect on this relationship?

Quantitative Research Methodology

Selection of Computer Software

The selection and use of a software program that actively engages students in science problem solving was required, since CAI was a focus of this study. After much
consideration, Science Sleuths, a software program developed by a Seattle-based company called Videodiscovery, was selected for the following reasons:

1. The first reason for choosing this software was because the distributor claimed that the program had the ability to improve problem-solving skills. More specifically, the accompanying teacher’s manual claimed that the following problem-solving and critical thinking skills would be built through use of the software: decision-making, reasoning, recalling and relating past experiences, linking ideas, analyzing, synthesizing, recognizing main idea and supporting detail, distinguishing fact from opinion, detecting bias, organizing, justifying conclusions, and evaluation.

2. The second reason for choosing this software was because of its acclaim. For example, Science Sleuths won a prestigious Invision Silver Medal Award in the young adult category. As well, the National Science Foundation awarded a grant to the software publishers to fund further software development based on the success of the Science Sleuths software. Additionally, in 1996 it won Newsweek Editor’s Choice Award for children’s software because of its entertaining, informative, and educational value.

3. The software was also selected on the basis of its content. More specifically, it met several learning expectations of the Ontario Science and Technology curriculum (Appendix A), thus making the software a potentially useful commodity for teachers aiming to cover an extensive curriculum.

4. The software was chosen on the basis of its accessibility. Science Sleuths was licensed by the Ontario Ministry of Education for use by every school in Ontario. Although the school within which the study was conducted did not initially possess
Science Sleuths within its repertoire of networked software programs, it was obtained and installed before the study commenced.

5. Finally, the software was chosen because it was developmentally appropriate for the participants. Videodiscovery personnel recommended that Science Sleuths would be best used with students from grades six to nine.

Science Sleuths presented students with real-life problems to be solved through experimentation in a virtual science lab. Students were required to gather data using 24 interactive science tools and to record this information in an electronic journal. The scientific mystery was solved only when the solution was validated by the steps needed to reach it. In this way, an emphasis was also placed on the process of problem solving and not just on the solution. Although students were required to find the correct answer, they did so in a manner that promoted free exploration of available resources and consideration of different solution paths. In this way, the philosophy behind the software was aligned with constructivism-informed ideas, as it aimed to provide a motivating, real-life experience within which students could apply their science knowledge. However, it should be noted that the problems presented in the software program are somewhat limited in their authenticity as scientific problems (i.e., the students are expected to find a specific pre-determined solution to a pre-determined problem).

Selection of Participants

The selection of participants for this study was another important consideration in the research process. A range in logical reasoning ability among participants was necessary in order to study its effect on the relationship between CAI and problem-solving performance. Grade seven students were purposely selected to participate in this
study because some were expected to be at a level of formal reasoning ability, some were expected to be in the concrete stage of logical reasoning ability, and others were expected to be in transition between the two stages.

The elementary school chosen for the study is located in a large south-eastern Ontario city of approximately 300,000 residents. The selected students were from low to middle income families and ranged greatly in ability, culture, and race. Parental consent was obtained before any student participated in the study (Appendix B). From this one school, 107 grade seven students across four different classes were invited to participate in the study and 84 students chose to do so. The students who did not return their consent forms did not provide an explanation for this decision. Students without consent forms were involved in the same problem-solving lessons and assessment tasks as students who had consented to participate in the study, as the lessons were compatible with required Ontario curriculum. The only difference in programming for these students was that instead of writing the GALT when students with permission wrote it, they were asked to do a reading activity related to problem solving. This was because the GALT was not directly linked to the curriculum, although it can be a useful tool for assessing the learning needs of a class.

Instrumentation

The instruments used to assess the quantitative variables of this study were the GALT, which is a standardized test of logical reasoning ability, and a tailor-made problem-solving pretest and posttest I developed for use in this study.

Group Assessment of Logical Thinking (GALT). Piagetian interview tasks, such as the balance beam, coloured squares and diamonds, and/or electronic switch-box tasks
have traditionally been used to measure reasoning ability (Gipson et al., 1989; Newton et al., 1981; Smith & Sims, 1992; Staver, 1984). However, researchers have reported several drawbacks to these interview tasks: they are impractical for use by classroom teachers, there is the potential for interviewer subjectivity bias, they are not suitable for assessing large numbers of participants, and/or they require substantial interviewer training and expertise (Bitner-Corvin, 1988; Tobin & Capie, 1981). Several objective logical reasoning ability measures, including the GALT, have been developed for use in place of these Piagetian interview tasks.

The GALT is a paper-and-pencil measure of six reasoning modes (conservation, proportional reasoning, controlling variables, probabilistic reasoning, correlational reasoning, and combinatorial reasoning). The full-length version consists of twenty-one questions which are mathematical in nature. In the first eighteen questions, students are required to select a correct response from a set of possible answers, as well as the correct justification for this selection from a set of possible reasons. In the remaining three questions, students are required to show a pattern for the combinations. Students are roughly classified developmentally, in terms of logical reasoning ability, by their scores on the GALT as follows: 0-8 are concrete, 9-15 are transitional, and 16-21 are formal (Bitner-Corvin, 1988). The abbreviated version of the GALT contains twelve items (two items to measure each of the six reasoning modes). More specifically, item numbers 1 and 4 are used to measure conservation, items 8 and 9 measure proportional reasoning, items 11 and 13 measure controlling variables, items 15 and 16 measure probabilistic reasoning, items 17 and 18 measure correlational reasoning, and items 19 and 20 measure combinatorial reasoning. Students are roughly classified in terms of their logical
reasoning ability by the following scores on the abbreviated version of the GALT: 0-4 are concrete, 5-7 are transitional, and 8-12 are formal.

The abbreviated version of the GALT was chosen to quantitatively measure logical reasoning abilities of participants for two reasons. First, the abbreviated GALT better conformed to the time constraints imposed by rotary classes and secondly, research supported the abbreviated GALT as a reliable and valid measure of logical reasoning ability. Bitner-Corvin (1988) used five separate convenience samples and several statistical analyses to test the reliability of both the long and the abbreviated version of the GALT. The author reported high reliability coefficients for the five samples (ranging from 0.76 to 0.86). Roadrangka, Yeany, and Padilla (1983) similarly reported that the GALT was a reliable measure of logical reasoning (total test reliability coefficient was 0.85). Roadrangka et al. found a strong overall validity coefficient for correlation between Piagetian interview tasks and the GALT (0.80).

**Problem-solving pretest and posttest.** I developed a paper-and-pencil pretest and posttest to assess the science problem-solving activities used in this study (Appendices C and D). To recap, the five problem-solving strategies taught and used in the introductory problem-solving lesson and in both CAI and traditional instruction were:

1. **Define the problem:** What is the root of the problem?
2. **Develop a plan:** What do I need to know? How am I going to find this information?
3. **Gather information:** How am I going to record this information?
4. **Sort-link-solve:** How does this information help solve the problem?
5. **Verify the solution:** Does the answer make sense? Can I explain my solution? Does my solution agree with the evidence?
These five strategies used in the lessons then formed the basis for the five questions asked in the pretest and posttest. The pretest and posttest were formatted identically except each presented a different problem scenario to be solved. Both problem-solving tests were administered in two parts. First, students were given Part A consisting of the problem scenario and a response sheet; only after it was completed and submitted to the researcher did participants receive Part B of the test. In Part A students were expected to read the problem scenario and then write a response to the first three problem-solving questions; namely:

1. Write the problem in your own words.
2. What do you need to know?
3. (a) How are you going to gather this information?
   (b) How are you going to record this information (be specific)?
Part B contained the second half of the same problem scenario and the two remaining problem-solving questions; namely:

4. Solve the problem (show your work).
5. (a) Does your answer make sense? Why or why not?
   (b) Does your solution agree with the evidence? Why or why not?

The development of these tests was a lengthy process requiring many revisions. The first draft was based upon an examination of the problem-solving assessment methods used in previous studies (Kanevsky & Rapagna, 1990; Kilpatrick, 1967; Norrie, 1988; Rashid, 1993; Schoenfeld, 1982). Several preliminary drafts were then critically discussed by the researcher, Master of Education classmates, and Brock University faculty members, with an array of experience in teaching and researching problem...
solving and science education. Student feedback on these tests was obtained during the pilot study and also assisted in achieving construct validity.

The science process skills targeted in these problem-solving tests were similar to those assessed by Ross and Maynes (1983) in a measure of experimental problem-solving skills. Specifically, their test made use of seven skills used by successful scientists when designing an experiment, "developing a focus (i.e., formulating a hypothesis), developing a framework (i.e., designing an experiment), judging the adequacy of data collected, recording information, observing relationships in data, drawing conclusions, and generalizing" (p. 65). Parallels can be drawn between the science skills used in Ross and Maynes' tests and the problem-solving tests designed for this study. In Step 1 of my tests, students were asked to develop a focus by examining the scenario to identify the problem, although a hypothesis was not required, as was by the Ross and Maynes' tests. In Steps 2 and 3 of my tests, students were asked to design an experiment, as was asked by the tests of the other researchers. In Step 5 of my tests, as in Ross and Maynes' tests, students were asked to judge the adequacy of the data collected by reflecting upon sources of experimental error. In Step 3 of my tests, students were asked to describe how they would collect and record data, just as students were required to do so in Ross and Maynes' tests. In Step 4 of my tests, students were asked to observe mathematical relationships in the data, draw conclusions, and make generalizations from the data in order to solve the problem, as was required in the other tests too. Thus, the problem-solving tests used in this study were designed to assess how problem-solving skills were used in a science context, just as Ross and Maynes had reported doing. A significant difference between my tests and those developed by Ross and Maynes is that I used open-
ended questions whereas they used a multiple-choice format; this was an acknowledged limitation to their tests as was reported in their study:

In the multiple-choice instrument, students are required to select/recognize the most appropriate response, whereas in the open-ended items students are required to generate the response. In this sense, the open-ended measures have greater similarity to classroom activities than the multiple-choice tests. (p. 72)

There were two important considerations in the presentation of data in the problem scenarios. First, it was decided that extraneous data would be included so that students would be required to filter through the given information to determine which data were relevant to the solving process. Since the sorting of information is an aspect of the problem-solving process, it was therefore incorporated into the tests. Similarly, it was stated by Marsh and Kumar (1992) that in problem solving, “children must not only understand what the question is, but they must also decide what information is relevant” (p. 34). Secondly, it was decided that students would be given the data needed to solve the problem in Part B of the scenario. This was done to avoid introducing a potential confounding variable to the study; since a significant difference between laboratory instruction and CAI is the manipulation of real versus virtual materials, it was thought that, if students had been required to make use of either of these methods to collect data in the pretest or posttest, one treatment group would have the advantage of practice over the other group. An alternative method of data collection, other than through physical or virtual means, was not conceivable so it was decided that students would not be required to gather data in the pretest or the posttest.
After the pretest and posttest data were collected and scored, statistical analysis was used to reveal the reliability between them. Written responses to the problem-solving pretest and posttest questions were scored by the researcher according to a reliable scoring scheme (Appendices E and F). A preliminary draft of the scoring scheme was based on scoring methods found in the literature (Kilpatrick, 1967; Norrie, 1988; Polya, 1957; Schoenfeld, 1982), as well as on possible solutions to the pretest and posttest questions which were determined by the researcher (Appendices G and H). Critical discussions by the researcher, Master of Education classmates, and Brock University faculty members assisted in this developmental process. The final version of the scoring scheme was presented in the form of an assessment rubric and contained four achievement levels for each of the five problem-solving questions. An exemplar illustrating each achievement level of each step in the scoring scheme was identified from student test responses (Appendix I). Interrater reliability was done for 30 randomly-selected test responses which were scored independently by two raters (i.e., myself and my faculty advisor).

Procedure

Data for the current study were collected while the researcher was simultaneously teaching a grade seven 'Pure Substances and Mixtures' unit based on the Science and Technology curriculum (MoET, 1998). The order in which the main study was conducted is presented in Table 2, followed by specific details about this procedure.

**Problem-solving strategies introductory lesson.** Earlier research has suggested that explicit instruction in the use of problem-solving strategies (e.g., heuristics) and student-teacher discussion about decisions made and conclusions drawn during the problem-
[Text content that has been redacted due to confidentiality.]

[Further redacted text]

[Redacted text]

[Continued redacted text]
<table>
<thead>
<tr>
<th></th>
<th>Treatment Group (CAI)</th>
<th>Control Group (Traditional Instruction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Introduction to study and distribution of consent forms</td>
<td>Introduction to study and distribution of consent forms</td>
</tr>
<tr>
<td>2.</td>
<td>Abbreviated GALT (1 lesson/30 min)</td>
<td>Abbreviated GALT (1 lesson/30 min)</td>
</tr>
<tr>
<td>3.</td>
<td>Problem-solving strategies</td>
<td>Problem-solving strategies</td>
</tr>
<tr>
<td></td>
<td>introductory lesson and application (2 lessons/60 min)</td>
<td>introductory lesson and application (2 lessons/60 min)</td>
</tr>
<tr>
<td>4.</td>
<td>Problem-solving pretest (1 lesson/30 min)</td>
<td>Problem-solving pretest (1 lesson/30 min)</td>
</tr>
<tr>
<td>5.</td>
<td>Computer-directed tutorial (0.5 lesson/15 min)</td>
<td>Teacher-directed tutorial (0.5 lesson/15 min)</td>
</tr>
<tr>
<td>6.</td>
<td>Computer-based problem solving (3.5 lessons/1.75 hr)</td>
<td>Laboratory-based problem solving (3.5 lessons/1.75 hr)</td>
</tr>
<tr>
<td>7.</td>
<td>Problem-solving posttest (1 lesson/30 min)</td>
<td>Problem-solving posttest (1 lesson/30 min)</td>
</tr>
</tbody>
</table>
solving process foster successful problem-solving performance (Cypher, 1987; Friedler et al., 1989; Kilpatrick, 1967; Kinnear, 1986; Murphy & McCormick, 1997; Norrie, 1988; Polya, 1957; Schoenfeld, 1980b; Schoenfeld, 1982; Slack & Stewart, 1990; Slotnick, 1989; Weber, 1993). For example, when problem-solving objectives were implemented across a large public school board, teachers first introduced problem-solving strategies and then provided students with an opportunity to apply them in practice (Norrie). Accordingly, problem-solving strategies were explicitly taught to all participants prior to their receiving CAI or traditional instruction. The problem-solving heuristics, or strategies, given to the students were:

1. Define the problem: What is the root of the problem?
2. Develop a plan: What do I need to know? How am I going to find this information?
3. Gather information: How am I going to record this information?
4. Sort-link-solve: How does this information help solve the problem?
5. Verify the solution: Does the answer make sense? Can I explain my solution? Does my solution agree with the evidence?

These problem-solving strategies were adapted from those provided in the Science Sleuths' teacher's manual and are closely related to Polya's (1957) problem-solving strategies. The main difference is that Polya presented four strategic steps and the software manual presented five. More specifically, Polya listed "carry out a plan" as a single strategy, whereas it is broken down into two strategies "gather information" and "sort-link-solve" in the manual and in the present study.

The procedure of the introductory problem-solving lesson was as follows. First, the problem-solving strategies were introduced and discussed in detail through a teacher-
directed mini lesson (Appendix J). Secondly, the class worked together to apply these new skills to solve a practical problem: Find two ways to shorten the time it takes to dissolve sugar in a drink. Students were given the materials needed to experimentally reach a solution (i.e., a beaker, water, sugar, stirring rod, etc.) and a problem-solving record sheet to facilitate the process (Appendix K).

**Traditional instruction.** The control group actively engaged in an exploratory laboratory activity. This activity, called “The Big Mix Up”, was introduced through a teacher-directed tutorial (Appendices L, M, and N). Students were given the problem of separating a heterogeneous mixture into its separate components without affecting the total mass of the mixture. They were required to design an experiment using the problem-solving strategies record sheet to facilitate the process. Students were given an opportunity to experiment with different equipment and substances in an open-ended practice session before starting the separation process. This pre-lab experience as well as occasional assistance from the teacher facilitated the problem-solving process. More equipment and substances than were needed were available for use; hence students had to decide what materials were needed to solve the problem. An enrichment activity which required students to separate a second (unknown) mixture applying principles previously learned to a slightly different context was provided. However, none of the students finished quickly enough to have the opportunity to complete the enrichment task.

**Computer-assisted instruction.** The experimental group received CAI facilitated by the teacher (Appendices O, P, and Q). As students engaged in the use of multimedia science software in the school computer lab, the majority of the teaching, learning, and transfer of necessary knowledge was facilitated by the computer.
Research has shown that the explicit teaching of computer skills leads to a more effective use of the technology (Weber, 1993). Therefore, all students in the treatment group participated in a computer tutorial session prior to engaging in the problem-solving program. The self-directed computer tutorial encouraged students to discuss any difficulties they anticipated with the program. After the tutorial, the Science Sleuths program presented the students with a scientific problem which they could solve using virtual tools in a computer-simulated lab. More specifically, the problem presented to students was to determine if an unknown substance that had washed up on the shore was harmful enough to justify the closing of a beach. Hints and help functions were available in the computer program, as well as assistance from the teacher, as needed. Students worked at their own pace, moving through six difficulty levels, and choosing how and when to manipulate resources in order to achieve a solution. Students attempted to solve the mystery by answering questions generated by the computer and supporting these answers with data collected and recorded in an electronic journal. They were also required to write down their steps on a problem-solving strategies record sheet to facilitate the process.

Several factors have been found to affect problem-solving performance through CAI (Cypher, 1987; McLean, 1989; Slack & Stewart, 1990). These include such things as the role of the teacher, student reaction to the simulation, limiting student frustration, interactions in cooperative learning, and concept understanding. Since the purpose of this study was to examine only the inherent differences between CAI and traditional instruction, these other factors were controlled as much as possible between the two
groups. The similarities and differences between CAI and traditional instruction are summarized in Table 3.

Additional student instructions. The following additional instructions were given to all students during the study:

1. Students were asked to write their names on the back of the GALT, pretest, and posttest response sheets. At a later time a participant ID number was added to the front of the response sheets to maintain anonymity in future reference and to minimize experimenter bias during the coding of the responses.

2. During the problem-solving tasks students were encouraged to collaborate with one another, but each student was responsible for writing out his/her own problem-solving record sheet.

3. It was emphasized that the process of solving the problem was as important as the final solution.

4. Students were encouraged to work through all tasks naturally (since this study aimed to minimize deviation from the regular classroom routine).

Pilot Study

A pilot study was conducted with 19 grade eight science students (selected from the same elementary school as those who participated in the main study). The aims of the pilot study were as follows:

1. To finalize practical considerations necessary for instruction (i.e., time needed to work through the computer tutorial, loading CD-ROM, etc.).
<table>
<thead>
<tr>
<th>Computer-Assisted Instruction (Science Sleuths Software)</th>
<th>Traditional Instruction (Separating Mixtures)</th>
</tr>
</thead>
<tbody>
<tr>
<td>students follow a scientific problem-solving model</td>
<td>students follow a scientific problem-solving model</td>
</tr>
<tr>
<td>problem-solving skills used:</td>
<td>problem-solving skills used:</td>
</tr>
<tr>
<td>• decision-making</td>
<td>• decision-making</td>
</tr>
<tr>
<td>• reasoning</td>
<td>• reasoning</td>
</tr>
<tr>
<td>• recalling and relating past experiences</td>
<td>• recalling and relating past experiences</td>
</tr>
<tr>
<td>• linking ideas</td>
<td>• linking ideas</td>
</tr>
<tr>
<td>• analyzing</td>
<td>• analyzing</td>
</tr>
<tr>
<td>• synthesizing</td>
<td>• synthesizing</td>
</tr>
<tr>
<td>• recognizing main idea and supporting details</td>
<td>• recognizing main idea and supporting details</td>
</tr>
<tr>
<td>• organizing</td>
<td>• organizing</td>
</tr>
<tr>
<td>• justifying conclusions</td>
<td>• justifying conclusions</td>
</tr>
<tr>
<td>virtual, unlimited and immediately accessible resources</td>
<td>real, limited, and sometimes accessible resources</td>
</tr>
<tr>
<td>available resources are:</td>
<td>available resources are:</td>
</tr>
<tr>
<td>• video interviews</td>
<td>• textbook</td>
</tr>
<tr>
<td>• still photos</td>
<td>• science notebook</td>
</tr>
<tr>
<td>• graphs and tables</td>
<td>• materials: table salt, pepper, sand, sawdust, iron fillings, mechanical mixture (salt, sand, iron fillings and sawdust), 4 plastic bags</td>
</tr>
<tr>
<td>• maps and charts</td>
<td>• lab equipment: safety goggles, magnet, hot plate, tongs, evaporating dish, electronic scale, support stand, funnel, filter paper, stirring rod, beakers, magnifier</td>
</tr>
<tr>
<td>• documents</td>
<td>• calculator</td>
</tr>
<tr>
<td>• encyclopedia</td>
<td>• glossary</td>
</tr>
<tr>
<td>• calculator</td>
<td>• search encyclopedia</td>
</tr>
<tr>
<td>• sleuth database</td>
<td>• interactive on-screen tools: Bunsen burner, coefficient, drill, fecal coliform count, gas chromatography, Geiger counter, graduated cylinder, gram stain, hand, indole paper, knife, mass spectrometer, microscope, petri dish, pH, pop-up, scale, sledgehammer, tape measure, thermometer, zoom-in</td>
</tr>
<tr>
<td>• electronic notebook</td>
<td>• applicable to daily life, motivating, authentic, potentially applicable to daily life, authentic, motivating</td>
</tr>
<tr>
<td>• glossary</td>
<td>• simultaneously challenges varying science aptitudes</td>
</tr>
<tr>
<td>• search encyclopedia</td>
<td>• because there are 6 levels of difficulty to choose from</td>
</tr>
<tr>
<td>• applicable to daily life, motivating, authentic,</td>
<td>• role of teacher:</td>
</tr>
<tr>
<td>• potentially applicable to daily life, authentic,</td>
<td>• role of teacher:</td>
</tr>
<tr>
<td>• simultaneously challenges varying science aptitudes</td>
<td>• facilitate computer-directed tutorial</td>
</tr>
<tr>
<td>• an enrichment activity will be provided for students who complete the first activity</td>
<td>• facilitate problem-solving process</td>
</tr>
<tr>
<td>• role of teacher:</td>
<td>• main source for providing help during investigation</td>
</tr>
<tr>
<td>• facilitate computer-directed tutorial</td>
<td>• ask open-ended questions during investigation</td>
</tr>
<tr>
<td>• facilitate problem-solving process</td>
<td>• guide them to resources as needed</td>
</tr>
<tr>
<td>• provide help (secondary to computer)</td>
<td>• provide assessment feedback</td>
</tr>
<tr>
<td>• ask open-ended questions during investigation</td>
<td>• open-ended exploration of problem to find a correct solution</td>
</tr>
<tr>
<td>• guide them to resources as needed</td>
<td>students are active learners</td>
</tr>
<tr>
<td>• supplement feedback from computer</td>
<td>students are active learners</td>
</tr>
<tr>
<td>• open-ended exploration of problem to find a correct solution</td>
<td>students receive immediate assessment feedback (by accessing the &quot;Present Evidence&quot; function)</td>
</tr>
<tr>
<td>students receive immediate assessment feedback (by accessing the &quot;Present Evidence&quot; function)</td>
<td>students receive delayed assessment feedback (after the teacher has marked their work)</td>
</tr>
<tr>
<td>students receive immediate help from computer as needed</td>
<td>students receive immediate help from computer as needed (i.e., by accessing Hints button or Help menu)</td>
</tr>
<tr>
<td>(i.e., by accessing Hints button or Help menu)</td>
<td>students do not always receive immediate help because of teacher to student ratio (approximately 1 to 30)</td>
</tr>
<tr>
<td>students are encouraged to collaborate with each other</td>
<td>students are encouraged to collaborate with each other</td>
</tr>
</tbody>
</table>
2. To gain experience in administering and scoring the GALT.

3. To further validate the problem-solving pretest and posttest with student feedback through field testing.

The procedure of the pilot study was similar to the main study, but was conducted over a shorter time period. The first step in the pilot study was administering the GALT to participants. This was followed by a quick lesson on the five problem-solving strategies prior to students writing the problem-solving pretest. After a single, brief session using Science Sleuths students were then given the problem-solving posttest. Due to time constraints, a limited number of participants, and the use of preliminary versions of the tests, reliable quantitative data was not collected during the pilot study. However, the pilot study did allow for a review of practical considerations and for obtaining valuable feedback from the students on a feedback response sheet written immediately following the completion of the pretest and posttest (Appendix R).

Research Design

The quantitative component of the study was a two-treatment design. The control group received traditional instruction and the experimental group received CAI for four class periods. The sample used in this study was a sample of convenience, since students had been placed into one of three grade seven classes by the school administration prior to the commencement of the study. The fourth grade seven class that participated in the study contained high-achieving students who had chosen to be enrolled in the French-Immersion program. This high achieving French-Immersion class was paired with the lowest achieving regular class to form the first treatment group. The remaining two (middle-achieving) classes were paired to make the second treatment group. It was
assumed that this deliberate pairing of the four classes would result in a similar number of students of concrete, transitional, and formal reasoning ability within the control and experimental group. Level of achievement was based on my assessment of the students in each class according to Ontario curriculum expectations, prior to the commencement of the study. The first pair of classes was then assigned as the experimental (CAI) group and the second pair of classes as the control group (traditional instruction).

Prior to receiving CAI or traditional instruction, all participants wrote the problem-solving pretest and after having received the instruction each wrote the problem-solving posttest. The dependent variable was problem-solving performance and it was operationally defined as score (out of 20) on the pretest, score (out of 20) on the posttest, and problem-solving improvement (determined by subtracting the pretest score from the posttest score). There were two independent variables in this experiment. The first was type of instruction and its two levels were experimental (CAI) group and control (traditional instruction) group. The other independent variable was logical reasoning ability, which was operationally defined by student score (out of 12) on the abbreviated GALT.

The problem statements generated two null hypotheses to be investigated quantitatively:

2. Logical reasoning ability, as defined by Piagetian reasoning levels, has no effect on problem-solving performance.
Researchers have stated that several characteristics of CAI promote successful problem solving and this formed the basis for the first working hypothesis (Friedler et al., 1989; Jungck, 1991; Kinnear, 1986; Krajcik et al., 1988; Marsh & Kumar, 1992; Pek, 1996; Simmons & Lunetta, 1993; Slack & Stewart, 1990; Slotnick, 1989; Soderberg et al., 1994). Quantitative studies which found a positive effect of CAI on problem-solving performance contained some research design limitations; namely an inequitable control group used for comparison and an ambiguous means of assessing problem solving. Additionally, several factors were identified for having an effect on the relationship between CAI and effective problem-solving, such as the role of the teacher and level of conceptual understanding. This study attempted to build on previous findings through careful consideration of methods employed by previous researchers. First, a more fair comparison of problem-solving performance between CAI and traditional instruction was created by engaging students in active learning in both treatment groups. Secondly, the pretest and posttest measures of science problem-solving performance were clearly specified and validity and reliability were carefully addressed. Thirdly, factors which could potentially affect the comparison between CAI and traditional instruction (such as the role of the teacher and the researcher) were controlled wherever possible to minimize confounding factors.

The second working hypothesis was also formed on the basis of previous research. Several studies reported a relationship between logical reasoning ability and success in problem solving (Gipson et al., 1989; Smith & Sims, 1992; Stewart & Hafner, 1994). Other researchers postulated that there may be a relationship among problem-solving performance, logical reasoning ability, and the use of computers, but
recommended further research (Krajcik et al., 1988; Krasnor & Mitterer, 1984; Papert, 1980).

Data Collection

Quantitative data were obtained through the GALT, a problem-solving pretest, and a problem-solving posttest. All responses on the pretest and posttest were scored according to an accompanying scoring scheme. The GALT is a standardized test which was scored according to pre-established criteria. The collected data were entered into SPSS spreadsheet tables for further analysis.

Data Analysis

Demographic characteristics of the sample population were gathered and summarized for each treatment group (i.e., CAI and traditional instruction) and for the total sample. The mean and standard deviation of GALT scores revealed the number of concrete, transitional, and formal reasoning students in each treatment group and in the total sample. The mean and standard deviation of problem-solving pretest scores were determined for both treatment groups and for the total sample. The same was done for posttest scores and for problem-solving performance improvement scores.

Inferential statistical analyses were performed to assist in the interpretation of the research questions. First, independent samples t-tests were done to determine the effect of instruction on problem-solving performance. The effect of logical reasoning ability on problem-solving performance was statistically analyzed through use of scatterplots and bivariate correlation analysis. Furthermore, effects of logical reasoning ability on problem-solving performance in each treatment group was statistically determined through use of a two-way analysis of variance. A scatterplot and bivariate correlation
analysis were conducted to reveal the reliability between the pretest and the posttest. Mean and standard deviation were calculated for scores, independently given by my faculty advisor and myself, on 30 randomly selected student problem-solving test responses. Finally, interrater reliability of the scoring scheme was determined for these scores through bivariate correlation analysis and Kappa analysis.

Methodological Assumptions

Three notable assumptions were made in the quantitative research methodology. First, students in one of the classes had chosen to be in a demanding academic program (i.e., French-Immersion) and this may have had an effect on the level of reasoning ability in the CAI group. However, the effect was assumed to have been counterbalanced by the pairing of this high-achieving class with another (low-achieving) class to create only a minor difference in distribution of reasoning ability between the two treatment groups. Secondly, it was assumed that four lessons in CAI or the traditional laboratory method was long enough to have had an effect on problem-solving performance. Thirdly, it was assumed that reliability and validity were addressed for the tests and scoring scheme of science problem-solving performance developed over the course of this study.

Limitations

Several limitations to the quantitative methodology were noted. First, this experiment was limited to the selection of participants attending one elementary school. Thus, results obtained from this study might only be applied to students with similar demographics to those who participated in the present study. As well, there were 84 participants in this study, although a larger sample size had been anticipated. The 23 students who did not return their consent forms did not give reasons for this decision.
Perhaps, a larger number of participants would have yielded different results. Another limitation was the use of only one software program. Results obtained in this study might only be generalizable to software programs with similar strategies and objectives (Kracjik et al., 1986). Similarly, only one laboratory activity was completed by participants and other forms of traditional instruction may not yield the same results. Another potential limitation was the combined role of teacher as researcher, although no conflict of these roles was perceived. Since the problem-solving pretest, posttest, and scoring scheme were developed over the course of this study, they provided limited results. It is recommended that modifications be made to these tests in further study. A potential limitation to the reliability of the scoring scheme was that it was determined for the scores of only two raters; my advisor and myself. Although student responses were scored independently, both raters were familiar with the scoring scheme.

Establishing Credibility

Methodological problems commonly found in CAI research were carefully considered and addressed by the present study. At the 1985 NARST convention, the external validity of research on the effectiveness of CAI in science education was critiqued for having the following problems (Krajcik et al., 1986):

1. Excessively broad generalizations were reported.
2. Inappropriate criterion measures were used.
3. Criterion measures were sometimes not specified.
4. Software assessed was, generally, less than optimal.
5. The research questions lacked substantive theoretical base.
What follows is a discussion of how these five common flaws were avoided in the current study. First, the results of the study may lend themselves to some generalization, as was addressed in the limitations section but "excessively broad generalizations" will not be attempted. It is the hope of the researcher that educators find the present research useful as a starting point for understanding student problem solving within their own specific context. Second, the reasons for selecting problem-solving performance and logical reasoning ability as criterion measures were their perceived relationship to each other and to CAI, as outlined in the literature review. Third, the criterion measures in this study have been explicitly specified as problem-solving performance and logical reasoning ability and have been operationally defined. Fourth, the software was selected for this study on the basis of its educational aims, widespread acclaim, content suitability, accessibility, and developmental appropriateness. These carefully considered characteristics supported it as an optimal software choice. Finally, the research questions have substantive theoretical base as was developed in the literature review chapter.

Ethical Considerations

Several ethical considerations arose over the course of this study. Most notably there was a need to obtain all necessary permission prior to commencement of the study. This included permission from Brock University, the participating school board, and the principal of the participating school. Since the participants of this study were minors, parental consent was received prior to student participation in either the main or the pilot study. Students who did not return permission forms were involved in the same instructional activities, but their responses were used for classroom assessment purposes only (i.e., not in the research analysis). Another ethical concern was maintaining
participant anonymity. To accomplish this, participants were asked to write their names on the back of the GALT, pretest, and posttest response sheets. A corresponding ID number was later written on the front so that no immediate connection could be made to individual students during analysis. A final ethical concern that arose during this study was the need to provide equal learning opportunities to all students. To meet this need, once the study was complete participants received the opposite form of instruction to that received during the study (i.e., the CAI group received traditional instruction and vice versa).

**Quantitative Findings**

First reported in this section are descriptive statistics, followed by inferential statistical analyses. In total, 84 grade seven students from one elementary school participated in this study. The students were divided into two treatment groups, one received traditional laboratory-based instruction (n = 38) and the other received CAI (n = 46). None of the participants in the CAI group reported having prior experience using the selected software program, Science Sleuths. The gender and age of these participants are summarized in Table 4.

**Logical Reasoning Ability**

Logical reasoning ability was measured by the abbreviated Group Assessment of Logical Thinking (GALT). A total test score, out of 12, was obtained for all 84 participants. Those scoring between 0 and 4 were roughly classified as being in the concrete reasoning ability stage, those scoring between 5 and 7 were roughly classified as being in the transitional stage, and those scoring between 8 and 12 were roughly
### Demographic Characteristics of Participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Age</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Instruction Group (n = 38)</td>
<td>Male</td>
<td>17</td>
<td>12.47</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>CAI Group (n = 46)</td>
<td>Male</td>
<td>19</td>
<td>12.43</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Total (N = 84)</td>
<td>Male</td>
<td>36</td>
<td>12.45</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>
classified as being in the formal stage. Resulting GALT scores are summarized in Table 5.

**Problem-Solving Performance**

Problem-solving performance was operationally defined by score on the problem-solving pretest, problem-solving posttest, and improvement score (determined by subtracting the pretest total score from the posttest total score). A total test score out of 20 was obtained for the pretest, as well as for the posttest through use of a scoring scheme rubric developed for the tests. These three performance scores are summarized in Tables 6, 7, and 8.

**Question 1: Effect of Instruction on Problem-Solving Performance**

Independent samples t-tests were used to determine if any significant difference in problem-solving performance existed between the participants who received traditional instruction as compared to those who received CAI. In each of the three t-test analyses the independent variable was instruction group and it had two conditions, traditional instruction and CAI. In the first t-test, the dependent variable was problem-solving pretest score. No significant difference in problem-solving pretest score was found between the traditional instruction and CAI conditions, \( t(82) = 0.168 \). In the second t-test, the dependent variable was problem-solving posttest score. No significant difference in posttest score was found between the traditional instruction and CAI conditions, \( t(82) = 0.371 \). In the third t-test, the dependent variable was problem-solving improvement score. No significant difference in problem-solving improvement score was found between the traditional instruction and CAI conditions, \( t(82) = 0.210 \).
Table 5

Summary of GALT Test Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Concrete</th>
<th></th>
<th>Transitional</th>
<th></th>
<th>Formal</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Traditional</td>
<td>32</td>
<td>84.2</td>
<td>5</td>
<td>13.2</td>
<td>1</td>
<td>2.6</td>
</tr>
<tr>
<td>Instruction (n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAI (n = 46)</td>
<td>36</td>
<td>78.3</td>
<td>9</td>
<td>19.5</td>
<td>1</td>
<td>2.2</td>
</tr>
<tr>
<td>Total (N = 84)</td>
<td>68</td>
<td>80.9</td>
<td>14</td>
<td>16.7</td>
<td>2</td>
<td>2.4</td>
</tr>
</tbody>
</table>
Table 6

Summary of Problem-Solving Pretest Total Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Instruction (n = 38)</td>
<td>11.34</td>
<td>3.02</td>
</tr>
<tr>
<td>CAI (n = 46)</td>
<td>11.46</td>
<td>3.16</td>
</tr>
<tr>
<td>Total (N = 84)</td>
<td>11.40</td>
<td>3.08</td>
</tr>
</tbody>
</table>
Table 7

Summary of Problem-Solving Posttest Total Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Instruction (n = 38)</td>
<td>11.37</td>
<td>2.48</td>
</tr>
<tr>
<td>CAI (n = 46)</td>
<td>11.61</td>
<td>3.30</td>
</tr>
<tr>
<td>Total (N = 84)</td>
<td>11.50</td>
<td>2.94</td>
</tr>
</tbody>
</table>
Table 8

Summary of Problem-Solving Improvement Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Instruction (n = 38)</td>
<td>0.15</td>
<td>2.83</td>
</tr>
<tr>
<td>CAI (n = 46)</td>
<td>0.0263</td>
<td>2.61</td>
</tr>
<tr>
<td>Total (N = 84)</td>
<td>0.0952</td>
<td>2.72</td>
</tr>
</tbody>
</table>
Question 2: Relationship Among Logical Reasoning Ability, Problem-Solving Performance, and Instruction

Statistical analyses were performed to determine if a relationship existed between GALT score and problem-solving test scores among participants. First, scatterplots were done to reveal possible patterns in the data. A positive relationship between pretest total score and GALT score (Figure V1), and a positive relationship between posttest total score and GALT score were revealed (Figure V2). A positive relationship between problem-solving improvement score and GALT score was also detected (Figure V3). Upon careful examination of these scatterplots, none of the relationships appeared strong enough to warrant a regression analysis. A small number of predictor variables and limited data points also made regression analysis unsuitable. Consequently, bivariate correlation analysis was chosen as the most appropriate statistical analysis for this study and a Pearson product-moment correlation was calculated for each pair of variables shown to be related in the scatterplots.

In the first correlation analysis, the dependent variable was problem-solving pretest score and the independent variable was GALT score. This analysis revealed a significant positive correlation between these two variables, \( r = 0.405, p < 0.01 \). In the second correlation analysis, the dependent variable was problem-solving posttest score and the independent variable was GALT score. This analysis revealed that there was a significant positive correlation between these two variables, \( r = 0.444, p < 0.01 \). In the third correlation analysis, the dependent variable was problem-solving improvement score and the independent variable was GALT score. This analysis revealed that there was no significant correlation between these two variables, \( r = 0.021, p > 0.05 \).
Furthermore, a two-way ANOVA was conducted to determine the effect of logical reasoning ability on problem-solving improvement scores of participants in the traditional group versus those in the CAI group. The dependent variable was operationally defined as problem-solving improvement score. The two independent variables were GALT score and instruction group (traditional and CAI). The ANOVA revealed no significant main effect or interaction effects, as reported in Table 9. However, a graph of these variables suggested a potential trend, in that students with high logical reasoning ability had slightly greater improvement in problem-solving performance through CAI, as compared to those students with high logical reasoning ability in the traditional group (Figure V4).

Relationship Between Problem-Solving Pretest and Posttest

The relationship between the problem-solving pretest and posttest developed for this study was statistically examined. First, a scatterplot revealed a positive correlation between total pretest score and total posttest score (Figure V5). Bivariate correlation analysis then confirmed that this relationship was statistically significant, $r = 0.593$, $p < 0.01$.

Interrater Reliability For Scores on the Pretest and Posttest

A scoring scheme rubric was developed for the pretest and posttest used in this study. The maximum score for each of the five problem-solving test questions was four, giving a maximum total score of 20. Two raters (i.e., myself and my faculty advisor) independently scored 30 randomly selected pretest and posttest responses (Table 10).

Interrater reliability for the pretest total score and for the posttest total score was determined through bivariate correlation analysis. A significant positive relationship
Table 9

Analysis of Variance for Problem-Solving Improvement Score, GALT Score, and Instruction

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>1</td>
<td>1.077</td>
</tr>
<tr>
<td>GALT</td>
<td>9</td>
<td>0.759</td>
</tr>
<tr>
<td>Instruction x GALT</td>
<td>6</td>
<td>1.740</td>
</tr>
<tr>
<td>error</td>
<td>67</td>
<td>(6.604)</td>
</tr>
</tbody>
</table>

Note: The value enclosed in parentheses represents the mean square error.
Table 10

Summary of Problem-Solving Test Scores Given by Rater 1 and Rater 2

<table>
<thead>
<tr>
<th>Section of Problem-Solving Test Scored</th>
<th>Rater 1</th>
<th>Rater 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Pretest Step 1 (n = 30)</td>
<td>2.70</td>
<td>1.06</td>
</tr>
<tr>
<td>Pretest Step 2 (n = 30)</td>
<td>1.70</td>
<td>0.92</td>
</tr>
<tr>
<td>Pretest Step 3 (n = 30)</td>
<td>2.27</td>
<td>0.83</td>
</tr>
<tr>
<td>Pretest Step 4 (n = 30)</td>
<td>2.20</td>
<td>1.16</td>
</tr>
<tr>
<td>Pretest Step 5 (n = 30)</td>
<td>2.03</td>
<td>0.76</td>
</tr>
<tr>
<td>Total Pretest Score (n = 30)</td>
<td>10.90</td>
<td>3.55</td>
</tr>
<tr>
<td>Posttest Step 1 (n = 30)</td>
<td>3.00</td>
<td>1.17</td>
</tr>
<tr>
<td>Posttest Step 2 (n = 30)</td>
<td>1.47</td>
<td>0.73</td>
</tr>
<tr>
<td>Posttest Step 3 (n = 30)</td>
<td>2.30</td>
<td>0.84</td>
</tr>
<tr>
<td>Posttest Step 4 (n = 30)</td>
<td>2.27</td>
<td>0.91</td>
</tr>
<tr>
<td>Posttest Step 5 (n = 30)</td>
<td>1.90</td>
<td>0.66</td>
</tr>
<tr>
<td>Total Posttest Score (n = 30)</td>
<td>10.93</td>
<td>2.98</td>
</tr>
</tbody>
</table>
between pretest total score given by Rater 1 and pretest total score given by Rater 2 was found, $r = 0.972, p < 0.01$. Similarly, a significant positive relationship between posttest total score given by Rater 1 and posttest total score given by Rater 2 was found, $r = 0.952, p < 0.01$.

Kappa was used to confirm interrater reliability on scores given for each step of the randomly selected pretest and posttest responses (Table 11 and 12). This analysis revealed that scores given by Rater 1 and Rater 2 for each step of the problem-solving pretest were in significant agreement at the 0.01 level. Similarly, it was revealed that scores given by Rater 1 and Rater 2 for each step of the problem-solving posttest were in significant agreement at the 0.01 level.

Chapter Summary

To summarize, this chapter presented the quantitative methodology and findings of the current study. First considered was the careful selection of computer software and participants. Next, the instrumentation used in this study were identified as the standardized GALT, as well as a tailor-made problem-solving pretest and posttest. The order and details of the main and pilot study procedures were then presented. Research design, data collection, and data analysis related to the quantitative research methodology was provided. Although great pains were taken to establish credibility in the study, methodological assumptions and limitations were inevitable and were openly discussed. Finally, important ethical considerations were stated.

The second part of this chapter reported quantitative findings. This included the demographics of participants, as well as their mean GALT scores, mean problem-solving test scores, and mean problem-solving improvement scores. From these summaries it was
Table 11
Kappa Analysis for Pretest Scores Given by Rater 1 and Rater 2

<table>
<thead>
<tr>
<th>Pretest Step</th>
<th>Kappa</th>
<th>Significance (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.864</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.834</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.820</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.947</td>
<td>0.000</td>
</tr>
<tr>
<td>Year</td>
<td>Value</td>
<td>Year</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest Step</td>
<td>Kappa</td>
<td>Significance (p)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1</td>
<td>0.852</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.935</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>0.949</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.765</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.752</td>
<td>0.000</td>
</tr>
</tbody>
</table>
noted that many of the participants were roughly identified to be in the concrete stage of logical reasoning development. On average, the group scored in the low-to-average range on the problem-solving tests and no significant improvement in problem-solving performance was noted from the pretest to the posttest. An independent samples t-test analysis revealed no significant effect of instruction on problem-solving performance. Logical reasoning ability was significantly correlated to performance on the problem-solving tests, but not to overall improvement. Finally, logical reasoning ability did not significantly affect the problem-solving performance of students in either treatment group. However, although not statistically significant, a graph of these variables hinted at the possibility of an interaction effect. Specifically, students with high logical reasoning ability showed greater improvement in problem-solving performance after CAI, as compared to those with high logical reasoning ability who received traditional instruction. Further research with students of high logical reasoning ability is recommended to confirm or refute this trend. The correlation between the pretest and posttest was high, as was interrater reliability for the scoring scheme of each test.

In the next chapter, qualitative analysis provides an in-depth look at how this study was conducted. More specifically, a personal narrative is used to answer the third research question of this study: What would an analysis of my thesis research experience in the Faculty of Education at Brock University reveal?
null
CHAPTER FOUR: QUALITATIVE ANALYSIS

Introduction

The previous chapter reported on the quantitative methodology and findings needed to answer the two research questions concerning the relationship among problem-solving, CAI, and logical reasoning ability. This chapter focuses on the qualitative analysis used to address the third research question: What would an analysis of my thesis research experience in the Faculty of Education at Brock University reveal? More specifically, a personal narrative approach was used to organize my research journal notes for insight into this question. Methodological assumptions and limitations of the study are openly discussed and credibility is established. Qualitative results are then presented in the form of a personal narrative.

Qualitative Research Methodology

Data Collection

Over the course of this research I continually collected potential qualitative data: recorded notes (Appendix S), had students provide written feedback about the problem-solving tests in the pilot study, collected student problem-solving record sheets written during CAI and traditional instruction, and collected reflective journals from participants. Since the purpose of the qualitative analysis was to describe my thesis research process at Brock University, it was decided that an analysis of my research journal notes would be the most promising data for this question. These notes were written during a related independent study course, a research proposal course, and time spent working on the thesis. Other researchers have similarly analyzed their own journal notes (McGinn, 2000; Rashid, 1993).
Data Analysis

The qualitative analysis done in this study can be best described as a personal narrative. Creswell (1998) described one form of the narrative as a chronological approach, in which the reader is engaged by the unfolding of events over time (p. 168). Similarly, Coffey and Atkinson (1996) described one form of the narrative as a chronicle in which "we chronicle our lives in terms of a series of events, happenings, influences, and decisions" (p. 68). Creswell then described the variation found in narrative forms, some are "heavily oriented towards theory, whereas others employ little literature and theory" and "some reports rely heavily on description of events, whereas others advance a small number of 'themes' or perspectives" (p. 168). The narrative presented in this study coincides with these descriptions from the literature. First, it can be described as a chronological narrative, in that it reported on the events, influences, and decisions which unfolded through my research process. Secondly, since almost all of the themes emerged from the data, very little reference to previous literature and theory was made. However, related literature did play a small role (i.e., problem solving as a form of research and in discussions on the advisor-student supervisory relationship). Finally, this analysis revealed themes which best described events occurring throughout the research process.

I began the qualitative analysis with a thorough reading of my research journal notes written during a related independent study course, a research proposal course, and the thesis research. To find emergent themes from the data I made use of open coding which, "involves taking data and segmenting them into categories of information" (Creswell, 1998, p. 242). Likewise, Coffey and Atkinson (1996) suggested that in qualitative analysis when "reading through data extracts, one might discover particular
events, key words, processes, or characters that capture the essence of the piece” (p. 31). These small units of information were then naturally combined to create a fewer number of categories. The next step was to logically connect the categories to create a few overarching themes that best described the research process. This method met Creswell’s recommendation that “researchers try to develop a small number of categories, slowly reducing the number from, say 30, to 5 or 6 that become major themes in a study” (p. 242). These final themes were then discussed in a personal narrative.

Limitations

A limitation to the qualitative component of this study is that it describes only one person’s experience of the research process (i.e., my own) through the use of only one data source (i.e., my journal notes). For a more thorough understanding of the experience of conducting M.Ed. thesis research at Brock University, these findings need to be expanded upon through the involvement of more participants and the use of multiple data sources.

Establishing Credibility

The credibility of the qualitative research was also addressed. Constas (1992) stated that “although the use of qualitative methods has increased greatly in popularity, many still question the defensibility of the qualitative orientation” (p. 253). He went on to say that a privatization of the categorization process was partially to blame for this bias. Constas recommended that “qualitative researchers make all aspects of their analysis open to public inspection” in order to establish greater credibility. Following this recommendation, the qualitative process is thoroughly and openly discussed in this study, thereby contributing to its credibility.
Qualitative Findings

Preliminary Themes

Open coding analysis revealed 153 units of information (i.e., the smallest pieces of information that can stand by themselves) in the database (Tables U1, U2, U3 in Appendix U). An example unit of information found in this data was “called local school boards for recommendation on software”. This step of the analysis served to organize my reflections into discrete data units.

As the 153 units of information were entered into a computerized table they were simultaneously organized into 20 categories. To do this, I looked for natural connections between the units of information and some categories reappeared in more than one data set. The resulting categories were: meeting with advisor, literature review, reflection, goal-setting, research question, coursework, ethics, use of model, proposal, methodology, variables, participants, software, curriculum, faculty members, validation committee, pilot study, future modifications, qualitative analysis, and quantitative analysis. This step of the coding process served to reorganize the units of data into a fewer number of related categories.

The last step in the open coding process was to combine the 20 categories to create six overarching themes found across all the data. These themes were selected because they best described my research experience. The themes that emerged from the data were: research interests, practical considerations, research design, research analysis, development of (the problem-solving) tests, and scoring scheme development. Part of these results are graphically displayed in a flow chart (Appendix T).
Previous authors have stated that the process of conducting research is a form of problem solving in itself (McGinn, 2000; Rashid, 1993). Similarly, it was noted that problem-solving strategies were utilized across all six themes describing the research process of this study.

**Question 3: A Description of the Thesis Research Process**

The third research question of this study will now be addressed (i.e., What would an analysis of my thesis research experience in the Faculty of Education at Brock University reveal?). In chronological order, the three main stages in the research process were: related independent study, research proposal, and thesis research. Certain themes were predominant in certain stages of the research process. Specifically, the predominant themes occurring in the independent study stage were research interests, practical considerations, and research design. The significant themes found in the proposal stage were practical considerations, research design, and research analysis. Finally, in the thesis stage the focus was on practical considerations, research design, research analysis, development of tests, scoring scheme development, and practical considerations. What follows is an in-depth discussion of the six main themes to describe my research experience.

**Research interests.** The first step of the research process was translating my broad interests into specific goals for thesis research. From the qualitative analysis, six sub-categories were found to mesh together to provide a picture of how this was accomplished. These categories were: meeting with advisor, literature review, coursework, reflection, goal-setting, and research question.
Interaction with my thesis advisor played an important role throughout the research process. The supervisory relationship followed a “negotiated order model”, as my advisor facilitated the research process, rather than directing it (Acker, et al., 1994). During tutorials we reviewed my work, discussed problems that had arisen, and planned future work (Acker, et al.). Our interpersonal relationship was very affable, thereby creating a “safe” and positive learning environment. My advisor’s implementation of this supervisory model well-suited my needs as a graduate research student.

In this theme, my advisor was instrumental to the refinement of my research goals. Since he had similar research interests he was able to effectively facilitate an exploration of the vast realm of science education literature. An excerpt from my journal illustrates this: “I have been skimming through the Handbook of Research on Science Teaching and Learning for some time now. My advisor suggested that I read it to familiarize myself with current trends in science education research and to help me narrow down my interests in the field.” This was one example of how my thesis advisor helped me to refine my research interests.

The literature review process and completion of M.Ed. coursework went hand-in-hand with the development of my research interests. As my knowledge of current trends in educational research increased, I was better prepared to pinpoint what it was that I wanted to study. For example, I wrote “I have just started the M.Ed. course 5P22: Technology, change and the curriculum’ and would like to explore the possibility of linking knowledge gained in this course specifically to science.” In a later entry I continued with, “I am currently evaluating a software program in 5P22 entitled ‘Science Sleuths’ which claims to develop problem-solving skills. It would be interesting to
investigate the relationship between problem-solving skills, science, and software such as this one.” This is an example of how a course required for the M.Ed. program helped to develop my thesis research interests.

Reflection was a technique which played a small role in the development of my research interests. When considering research goals I knew that I wanted to relate them to some of my own past experiences. An interesting excerpt from my reflective journal illustrates this:

The first time I took an Introductory Genetics course at the University of Waterloo in 1992 we had to do a *Drosophila* fly lab. We actually bred two flies with a certain genetic makeup and then went back to the lab and examined these flies for a particular genetic trait. In 1997 I upgraded at UPEI and took Genetics again. This time around we were required to use an Internet-based virtual fly-breeding program on the computer. Of course there were pros and cons to both types of learning, but for me personally, I think that in this instance I got more out of the lesson the ‘virtual’ way!

The refining of my research interests eventually resulted in the development of specific research questions for my thesis. However, the focus of these inquiries continued to change as I met with my advisor, set goals, reviewed the literature, completed coursework, and reflected on my past experiences.

**Research design.** The first hurdle of this task was choosing a specific research question, as was previously discussed through the development of my research interests. Other necessary considerations of the research design were: methodology, variables, participants, and software. Decisions surrounding these parts of the research design were
accomplished in part through coursework, as well as through meetings with my advisor and other faculty members.

A significant aspect of the research design process was deciding on the methodology. I felt comfortable enough with qualitative and quantitative methods, having studied them through M.Ed. coursework, to want to include both in my thesis. The decision to use one, or both types of methodology continually changed throughout the research process.

For example, at one point I wanted to do a separate qualitative analysis on problem-solving data collected through the use of the think-aloud method. I then decided that student reflective journals would be the most practical way to collect qualitative data on problem solving. At this time, I also wanted to collect data through a custom-made problem-solving measure, for a separate quantitative analysis. Later yet, I decided that two separate analyses were not the best way to achieve my research goals and that it would be better to use qualitative data only to support quantitative findings. I wrote:

I have been very undecided about how to use qualitative analysis in my report. I decided to use direct quotes to supplement the quantitative findings, rather than to do a separate qualitative analysis. This decision fell into place as I was working through my research questions and methodology. I can still change my mind, if I find that there is a need to explore a qualitative analysis of the data.

Once the thesis data were collected this decision did change again. With the help of faculty members, I realized the significance of having developed a measure of problem-solving and a reliable scoring scheme and wanted to include it in my analysis. Since this
process would be best interpreted through qualitative analysis, separate from the quantitative findings, this was the final decision on methodology.

Just as the research question and methodology evolved over the course of the study, so did the variables to be studied. To illustrate this, at one point in the research process I considered using school grade divisions as a variable in the study: "I think it would be interesting to compare problem-solving skills of students in primary, junior, intermediate, and/or senior grades." However, I later wrote, "one problem that I envision is finding software that is similar enough to be compared, but differs enough to be age-specific". This led to the decision to choose one grade division (i.e., late elementary) and to "explore how levels of logical reasoning ability/development relate to problem solving" as a separate variable. This was just one example of how variables in the study were modified as the research design was developed.

Another important aspect of the research design process was choosing a sample population. The first step in this process was to decide which participant characteristics were desired for the study. Late elementary students were eventually chosen because it was my preferred teaching division and because they best suited some other research goals of the study (i.e., expected to demonstrate a range in level of logical reasoning ability, curriculum content, and software compatibility).

The next step was to recruit interested participants. My thesis advisor assisted me in networking possible sample populations. About this same time I began teaching rotary science at an elementary school and decided that my students would be the best candidates for this study. My concern was that I would have the dual role of teacher and researcher, but to the best of my knowledge it did not affect the proceedings of the study.
Once the participants were selected I had to pair classes and assign treatment to each pair. In summary, one aspect of the research design process involved making many important decisions surrounding participants.

Another major aspect in the research design process was the selection of computer software. Various software was critically examined with the help of my advisor and local school board contacts. The software program needed to meet the requirements of my study and also needed to be developmentally suitable for the participants. A sample journal entry illustrates one aspect of this decision process: “My advisor called the BioQUEST company and has ordered a complimentary copy of the GCK (genetics construction kit) [software] program; as soon as it arrives I can review it and decide if I want to use it for my study.” As it turned out, this software was geared towards older students and therefore was not a suitable choice for the late elementary school participants. In conclusion, another important part of the research design process involved the careful selection of suitable computer software.

To recap, research design involved decisions pertaining to the research question, methodology, variables, participants, and software. The strategies used to make these decisions were coursework requirements, as well as meetings with my advisor and other faculty members.

Practical considerations. Many practical considerations arose over the course of my research. From my perspective, as a researcher with limited experience, some of these practical considerations were not anticipated and seemed like major obstacles when they arose. Through qualitative analysis, areas of practical concern were categorized as: software, participants, ethics, curriculum, and proposal. Three strategies I used in
overcoming these difficulties were goal-setting, meeting with my advisor, and conducting a pilot study.

As was previously discussed, the selection of software for this study was very well thought out and was partially influenced by practical considerations. This process was exemplified in the following quote: “I looked into the TVO web site which describes Ministry-licensed software. The advantage of this software is that is easily accessible to Ontario schools. The disadvantage is that the software for the older grades seemed to be more exploration- and reference-based, rather than one using problem-solving skills. If I decide to use this software I might want to have students in late elementary school participate, rather than high school students, so that I can make use of problem-solving software like Science Sleuths.” Since I began teaching late elementary school students shortly after this journal entry was made, I became interested in using Science Sleuths with this group. Thus, it is clear that practical considerations, such as accessibility, played a role in decisions surrounding the selection of software.

As was alluded to in the previous paragraph, decisions surrounding participants were also often affected by practical considerations. For example, when I considered working with secondary school students the timelines for my thesis research were very much dependent upon their school calendar dates, as well as their teacher’s long range plans. I wrote:

I will hopefully be able to collect data starting in February or March. It depends on what is best for the teacher(s) I will be working with (i.e., some considerations are that second terms start the beginning of February and then there is the March Break).
I quickly realized that the research process was not just dependent upon my own plans, but very much influenced by others who would become involved. Although I did not end up working with high school students, I was similarly concerned with the instructional needs and school timelines of the student sample who did participate.

In a similar vein, ethical considerations were critical to the research process. Before commencement of the study, ethics approval was needed from various sources (i.e., Brock University, the participating school board, the participating principal, the students, and their parents). This took considerable time and effort and reinforced the realization that the research process was very much affected by others who would become involved.

Since I chose to conduct my study in the combined role of teacher and researcher, I was very conscientious in meeting curriculum expectations from the Ontario Ministry of Education. These curriculum boundaries partially shaped the instruction delivered and assessments made over the course of study, as well as the selection of software. When I began teaching the grade seven classes I knew that I would be teaching two science units (i.e., 'Heat' and 'Pure Substances and Mixtures'). Simultaneously, I would be teaching one science unit (i.e., 'Fluids') to grade eight classes over two terms. I thoroughly examined the software and the three science units to determine which would have the greatest overlap. The grade seven unit on 'Pure Substances and Mixtures' had the most overlap with Science Sleuths, so I decided to conduct research and teach that unit simultaneously. The practical nature of this decision demonstrates how large a role such considerations play in the research process.
The thesis proposal also required much practical consideration. At the beginning of the proposal stage I had to ask faculty members if they would accept a role on my thesis committee. Another practical consideration noted in my journal was, “I will need to give my thesis committee 2-3 weeks to look at my proposal before I can present it to them in a proposal hearing”. These small, but important planning details were crucial considerations throughout the research process.

I continually set goals for progress on my proposal such as, “I am finished Chapter 2 (literature review) but it needs to be edited, and I have made solid starts on Chapter 1 and 3. My personal deadline is still Friday, but it might not be until late that night!” The setting of, and following through with, such goals were a necessity in order to accomplish tasks in a timely manner.

When faced with practical obstacles I often turned to my advisor for advice. He played a critical role in helping me to consider possible solutions, as well as in the development of some necessary research skills. As a new graduate student at Brock University I was unfamiliar with some practical aspects of the research process such as obtaining ERIC articles and the proper form for writing an annotated bibliography. In my notes, I recorded that I was to “meet with my advisor next Monday to learn how to obtain ERIC articles.” Although such a task may seem small, it was critical to the development of a thorough literature review. Thus, my thesis advisor served as one resource for obtaining the necessary practical information needed to complete my thesis research.

The pilot study was crucial to the research process, as it brought to light many practical considerations that may have been overlooked otherwise. For instance, I “realized that I only had seven MS Windows version CDs and the rest were Macintosh
version CDs”, all for use with MS Windows-based computers. This was only detected after the students in the pilot study spent time on the computers and were unable to access the program. Had this occurred during the main study it would have set back the data collection a considerable amount of time, as more MS Windows version CDs were collected from the school board’s IT team.

Another critical revelation from the pilot study was that I “decided to have a seating plan set up for the students participating in the main study because of the difficulty in managing the students in the computer lab.” This was necessary because I wanted “the focus to be on learning problem-solving skills, not on who is going to sit where, etc.” By pre-establishing the seating arrangement, the students in the main study began using the computer much quicker and with less distraction, than those who participated in the pilot study. These practical problems and all others presented during the pilot study were resolved prior to the commencement of the main study. If not for the pilot study, the results of the main study may have been negatively affected by such obstacles.

In summary, as a relatively new researcher I was not fully conscious of how crucial a role certain practical considerations were to the research process. This included such areas as the selection of software and participants, ethics approval, curriculum requirements, and the research proposal. Goal-setting, meeting with my advisor, and conducting a pilot study were strategies used to overcome practical difficulties. To summarize, elements of this theme significantly impacted my own learning, as well as the research process.
Research analysis. Research analysis was a significant component of the research process. The two main divisions within this theme were: quantitative analysis and qualitative analysis. The strategies employed were: meetings with my advisor, required coursework, and the use of a model.

The research analysis consisted of a quantitative component and a qualitative component. Naturally, different considerations were relevant to each analysis and both were recorded in my journal. As I worked through this aspect of the research process I kept record of what steps I took and the rationale for doing so. For instance, in reflection on the quantitative analysis I wrote: “From these scatterplots there were no relationships strong enough to do a regression analysis (i.e., not enough predictors). As well, there are too few points to do a regression analysis properly”. This careful recording of my steps helped to organize my thoughts throughout the research process.

Meetings with my advisor were invaluable in providing me with direction during the analytical process. At one point I wrote: “I have scored and entered all the data into an SPSS spreadsheet. I will be meeting my thesis advisor next week to discuss statistical analysis of these data.” On another day I wrote: “we discussed the statistical analysis that would be best to use for interrater reliability and decided on Kappa”. I went on to research and record that “Kappa is a measure of interrater agreement that tests if the counts in the diagonal cells (the subjects who receive the same rating) differ from those expected by chance alone, (SPSS Inc., 1998, p. 81)” This illustrates how meeting with my advisor facilitated the completion of the research analyses.

After a phone conversation with my advisor concerning the qualitative component of the research, I wrote that he had told me I was “on the right track” and that he
suggested that I “just keep writing now and then revise it later”. Specifically, I would need to “move some components from Chapter 4: Findings to Chapter 3: Methodology” in the qualitative report. Once again, my thesis advisor played a critical role in the research process; in this discussion he facilitated the formatting of the research analysis.

Required coursework was integral to the completion of my research analyses. I often referred back to textbooks, handouts, required journal readings, and notes from required M.Ed. classes. In particular, I found the introductory research methods course and the qualitative methods course to be useful in this endeavor.

Another strategy I used in the research analysis process was to refer to a model. An example of this was stated very simply in my journal: “Started qualitative analyses. Modeled after McGinn’s dissertation (2000)”. Throughout the research process I found myself referring to previous research for guidance. For example, I recall borrowing M.Ed. theses and projects from Brock’s library to model other students’ formatting styles, when writing my own proposal.

Thus, quantitative and qualitative research analyses were achieved through meetings with my advisor, the completion of required M.Ed. coursework, and the use of models.

Development of problem-solving tests. A significant accomplishment of this study was the development of a valid and reliable problem-solving measure. This qualitative analysis revealed that validity and reliability were achieved through use of the following strategies: coursework, validation committee, proposal, pilot study, and participants. Finally, future modifications were also found to be an important aspect of this theme.
First, I will describe the process of attaining validity for the tests. I began this endeavor by examining the literature on problem-solving performance in mathematics and science. This research provided me with a starting point for developing a valid measure of problem solving (Kilpatrick, 1967; Kinnear, 1986; Krajcik et al., 1988; Murphy & McCormick, 1997; Polya, 1957; Ross and Maynes, 1983; Schoenfeld, 1982; Simmons & Lunetta, 1993; Slack & Stewart, 1990; Soderberg et al., 1994; Weber, 1993).

Recommendations made by classmates in a Master's level science education course were another source for validation. Specifically, I presented a preliminary version of the tests and the scoring scheme, as well as the basic design of my study during one of our in-class sessions. Since my classmates had a wide range of experience and knowledge in education, our discussions provided constructive ideas for improving validity of the measure. For example, they suggested that each test be delivered in two parts, rather than one, so that participants would have the opportunity to demonstrate an understanding of how to collect and then effectively utilize scientific data.

Validity was also addressed by a small committee consisting of myself (the researcher), my thesis advisor, and two other Brock University Faculty of Education members. One of these faculty members had previously conducted problem-solving research and the other was a former science teacher and current secondary school administrator. My thesis advisor had related experience in both areas. The purpose of this committee was to review preliminary versions of the problem-solving pretest and posttest, as well as the scoring scheme for these tests, in order to make suggestions on how to improve validity. In my journal I wrote that "many changes to the pretest and posttest resulted from our discussion". For example, the wording of the test questions...
was changed to a level appropriate for late elementary students. Additionally, the validation committee recommended that I use raw data in the problem-solving scenario, rather than presenting it to participants in table form. The rationale for this recommendation was to set students up to demonstrate another aspect of problem solving in science (i.e., how to organize and interpret data collected in an experiment).

Similarly, my thesis committee, consisting of two faculty members and my thesis advisor, briefly reviewed the problem-solving tests during a hearing for the proposed research. From this hearing I recorded that one committee member “suggested that I underline the phrase ‘most consistently’ in Part B [of the pretest] because it is important for the students to pick it up if I am expecting them to perform average calculation of bubble size”. These external sources provided me with a well-rounded view, thereby helping me to achieve a valid problem-solving pretest and posttest.

Another source for establishing validity was through the implementation of a pilot study. Participants in the pilot were asked to fill out a feedback sheet immediately following the writing of each problem-solving test (Appendix R). The purpose of the feedback sheet was to provide me with greater insight into the appropriateness of the tests for late elementary students, from their perspective. The majority of the students gave positive feedback on the developmental-appropriateness of the test. For example, one question on the feedback sheet was: “Do you think changes should be made to the test? If so, what?”. One student wrote “No because it’s right for a grade 8 level”. Another student wrote “No, because it’s suppose to make you think”. A third student response was, “No everything was okay”. Thus, consideration of the student perspective also contributed to the validity of the problem-solving tests.
A change to the wording of one of the problem-solving questions resulted from the pilot study. Namely, Step 3 was revised from "How am I going to record this information?" to "How am I going to record this information (be specific)?". In my notes I reflected on the rationale for this change: "Why was this change done? Because some students wrote 'chart' and I hoping to see chart headings as well. By adding the phrase 'be specific' [to the question] I hoped to facilitate their doing so". Thus, the field-testing of the problem-solving tests through a pilot study was critical to the development of a valid measure of problem solving.

To summarize, validity was achieved through an examination of previous research on problem solving, suggestions from M.Ed. classmates, discussion with experienced members of a validation committee and a faculty thesis committee, as well as, through information gathered in a pilot study.

Next, I have briefly described the process of attaining reliability for the tests. A dilemma with test reliability was recorded in my notes:

I think that the posttest should be more difficult than the pretest because I do not want improvement in problem-solving to be a result of a learned effect. On the other hand, the tests should be similar enough to one another because they are supposed to measure the same skill.

In other words, a balance was needed to produce two tests which were similar enough to measure the same skill, but different enough to minimize a learned effect.

Through discussion with the validation team, it was decided that the wording of the test questions, as well as the problem-solving steps taught in the introductory lesson, should be identical. Thus, the only difference between them was the problem scenario
and students were expected to apply the same problem-solving strategies each time. The
problem scenario in the posttest was slightly more complex than that of the pretest to
minimize a learned effect. Quantitative analysis revealed a significant correlation in
student achievement on the pretest and posttest. Therefore, reliability between the pretest
and posttest was addressed through meeting with a validation committee and later
affirmed through quantitative analysis. Reliability will be further discussed in relation to
the scoring scheme.

Although many steps were taken to make the problem-solving measures as valid
as possible, a need for further revision was evident throughout the research process.
Accordingly, ideas for future modifications of the tests were often recorded in my
journal. The following quotation illustrated this:

I have learned how important it is to keep revising such an instrument as this one
because there are always more variables than you can possibly consider prior to
field testing it; I think that I could have revised this test over and over again and
still have found some necessary changes after field testing.

I still believe that revisions can be made to improve the reliability and validity of the
problem-solving tests, even after the completion of my thesis. It has been similarly
recommended by other researchers that continued experimentation is necessary in order
to establish an instrument which reliably and validly measures a desired skill
(Frederiksen, 1986; Ross & Maynes, 1983).

Thus, validity and reliability of the problem-solving tests were achieved through
coursework requirements, meetings with a validation committee, the research proposal,
conducting a pilot study, and with help from participants. Future modifications were also an important consideration in this area.

**Scoring scheme development.** Another significant accomplishment of this study was the development of a valid and reliable scoring scheme to accompany the problem-solving tests. From the reflective analysis, five sub-categories described this aspect of the research process: coursework, validation committee, meeting with advisor, future modifications, and quantitative analysis. These were the same as the categories used to describe the development of the tests themselves, as the processes were similar.

The first draft of the scoring scheme was a checklist based on expected solutions to the pretest and posttest problem scenarios, student achievement outcomes from the Ontario Science and Technology curriculum (e.g., the use of science and technology terminology, design of a fair test), as well as some elements of problem-solving measures previously established in the literature (Kilpatrick, 1967; MoET, 1998; Norrie, 1988; Polya, 1957; Schoenfeld, 1982).

After presenting this scoring scheme in an M.Ed. science education course I received some recommended changes from classmates. As well, one classmate gave me copies of problem-solving rubrics (from various unknown sources). It was suggested by this classmate that I revise the scoring scheme into a rubric, from the checklist format, to allow for assessment of a greater variation in student response. This was recorded in my journal: “Developed a rubric for scoring pretest and posttest using three rubrics given to me by a classmate in Education 5P02. References are unknown. My advisor recommends that I refer to them in the acknowledgment section of my report.” This example showed how my scoring scheme was further developed in conjunction with M.Ed. coursework.
I then presented the revised scoring rubric to the validation committee. At this point in time the committee consisted of my thesis advisor, myself, and only one of the faculty members who had previously assisted with the development of the tests. This meeting was recorded in my journal: “I presented a first draft of a scoring scheme rubric to them and we went through each component at each level and discussed wording and meaning for each.” After a lengthy discussion it was decided that I would need to make major revisions to this preliminary draft, with assistance from my advisor. Thus, the validation committee played a critical role in the development of my scoring scheme.

The first step in this lengthy revision was working with my advisor to determine what the “ideal” solution might look like (Appendices G and H). At this time I wrote:

My advisor and I will also each do the pretest and posttest to determine what the “ideal” solution might be for each; this was an excellent suggestion by the committee and in hindsight I would have liked to have ‘piloted’ the tests with the committee in a similar manner prior to having administered them to the students (i.e., have the committee members write the tests themselves first to determine if each question was adequately worded to prompt students to answer in a way that allows insight into the problem-solving process).

By completing the test from the perspective of a participant I realized that the wording of parts of the tests could have been stated more effectively. For instance, I discovered that there was greater overlap between certain questions than was previously thought. Specifically, Question 2 “What do you need to know?” and Question 3 (a) “How are you going to gather this information?” were more compatible than the subsections Question 3 (a), “How are you going to gather this information?” and 3 (b), “How are you
going to record this information?”. Revising the wording of these questions would be one change that I would make if I was to use these tests in future research.

This revision process revealed a need for another modification in the wording of the problem-solving measures. Specifically, Question 5 subsection (a), “Does your answer make sense? Why or why not?” and subsection 5 (b), “Does your solution agree with the evidence? Why or why not?” were redundant. Although I now believe that these two subsections could be combined into one question, this was not apparent until after collection of the main study data.

I have learned from this experience that continual revision is critical to the development of an effective assessment tool. Only after field testing this instrument did I realize that it was not possible to anticipate every variable based on theory alone. From my own experience, I would recommend that a researcher “tries out” a new or newly revised test him- or herself before administering it. After “trying out” the pretest and posttest, I was in a much better position to make changes to improve the rubric and the tests themselves.

The next step in the development process was to make revisions based on a review of student test response exemplars, with the assistance of my advisor. Since I expected higher test scores in the posttest, than in the pretest, this was the starting point for finding Level 4 (the highest rubric level) exemplars. Levels 3, 2, and 1 were then defined for the five steps of the test. A rubric based upon what the participants had demonstrated emerged from discussions with my advisor. This was balanced with what we carefully considered to be important criteria at each level of each step. I wrote: “we
discussed the criteria for Steps 1 and 2 in detail; we applied these criteria to the samples and came up with a level for each [of the randomly selected] student responses”.

Another important revelation occurred at this time. Emphatically, I wrote: “We decided to score answers found anywhere on the sheet, rather than to restrict students to write required information in specified spaces; I believe that this will open up the results to accommodate the natural problem-solving processes of students.” In other words, when I drafted the scoring scheme I expected that each question would be answered step-by-step within each carefully laid out section of the test. However, upon examination of the response exemplars it became evident that the problem solving process was not “cut and dry” and answers appeared in unexpected sections of the response page. Consequently, the rubric needed to accommodate this natural problem-solving process.

As a result, I decided to score answers found anywhere on the sheet rather than to restrict students to write required information in specified spaces.

After “I, the researcher, met with my advisor, the scorer, to explain the [finalized] scoring rubric and to illustrate it with exemplars” we separately scored 30 randomly selected student responses. We then met to compare our scores for every step of the 30 responses and any unjustified difference in score (i.e., an omission) was acknowledged. Differences in interpretation of student work did not result in changes to either researcher’s original score. This comparison process helped me to improve upon the wording of the rubric criterion. Consequently I wrote, “the wording of some of the criteria in the scoring rubric will be changed to be as ‘scorer-friendly’ as possible.”

Quantitative analysis later revealed a high interrater reliability for these 30 randomly selected pretest and posttest responses.
Further scrutiny revealed an imbalance between the pretest and the posttest. This was thoroughly recounted in my journal:

After scoring the posttests, and discussion with my advisor, I came to the realization that there is a serious discrepancy between the design of the pretest and the posttest; I had presented to the [validation] committee the idea of adding extraneous data to determine if students were able to 'sort' through data in the problem-solving process; the extraneous data in the pretest (i.e., the time taken to make bubbles) can be more easily incorporated into the final answer, in a meaningful way, than the extraneous data in the posttest (i.e., the dissolving rate of potassium chloride); without realizing this beforehand, I had originally designed the scoring scheme so that one aspect of the Level 4, Step 4 [criteria] was that the student was required to use the extraneous data in a meaningful way; what happened in scoring was that the students were achieving Level 4 in the pretest but not in the posttest; after discussion with my advisor I realize this limitation to my study and present it as such; Also, I have changed this aspect of the Level 4, Step 4 [criteria] to read 'using the extraneous data in a meaningful way, or not using it at all'; in other words, as long as it is not used incorrectly, the subject would not be penalized for that component of the criteria.

As I had stated in my journal, the purpose of including nonessential data was to determine if students could sort through information and decide what was needed to solve the problem. In a future version of the tests I would still include this aspect, but modify it so that it was better balanced between the pretest and posttest. This example
demonstrated how future modifications of the scoring scheme and the tests were considered throughout this aspect of the research process.

The development of the scoring scheme was a daunting task. It involved the use of M.Ed. coursework requirements, meetings with a validation committee and my faculty advisor, and quantitative analysis. Future modifications for the scoring scheme and related parts of the tests were also discussed throughout this process.

**Problem-solving strategies.** A pattern that was noticed across all six themes was my use of problem-solving strategies to overcome difficulties presented over the course of the research. This deductive finding was based on previous research findings (McGinn, 2000; Rashid, 1993). Both studies presented the notion of research as a form of problem solving, through a rich analysis of the authors' own experiences. In my own study, the use of the following problem-solving strategies was evident across each of the six themes that emerged from the data:

1. Meetings with my advisor and other faculty members (in required courses; validation committee; thesis committee),
2. The use of a model,
3. Goal-setting,
4. Conducting a pilot study,
5. Conducting a literature review,
6. Completing and reviewing required coursework, and
7. Reflection.

Over the course of my thesis I was faced with many challenges and made use of these seven problem-solving strategies to overcome them.
Recapping, these strategies were applied in the research process as follows. In developing my research interests I relied on meetings with my advisor, conducting a literature review, application of M.Ed. coursework, reflection, and goal-setting. During the research design process, the problem-solving strategies I used were meetings with my advisor and other faculty members, and application of M.Ed. coursework. Strategies that helped me to realize and overcome practical considerations were goal-setting, meetings with my advisor, and conducting a pilot study. The research analysis process required meetings with my advisor, application of M.Ed. coursework, and the use of a model. Finally, the strategies used when developing the problem-solving tests and scoring scheme were application of M.Ed. coursework, quantitative analysis, as well as, meetings with the validation committee and my thesis advisor. Therefore, in accordance with previous findings, I too found the research process to be a form of problem solving. More specifically, I made use of seven problem-solving strategies to reach desired thesis outcomes.

Chapter Summary

To summarize, qualitative findings were obtained from analysis of my own research journals. Open coding of these data revealed 153 discrete units of information which were then naturally combined to form 20 categories (meeting with advisor, literature review, reflection, goal-setting, research question, coursework, ethics, use of model, proposal, methodology, variables, participants, software, curriculum, faculty members, validation committee, pilot study, future modifications, qualitative analysis, and quantitative analysis). The final step in the coding process was to combine these 20 categories to create six overriding themes. A personal narrative approach was used to
describe my research experience according to these themes (i.e., research interests, practical considerations, research design, research analysis, development of the problem-solving tests, and scoring scheme development). Finally, in accordance with other researchers’ experiences, I made use of problem-solving strategies throughout my own research experience, as a graduate student in the Faculty of Education at Brock University.
CHAPTER FIVE: DISCUSSION, IMPLICATIONS, AND SUMMARY

Introduction

In this chapter, findings from the current study are discussed with respect to the background literature presented in previous chapters. It begins with a summary of problem-solving test reliability, the demographics of the sample, logical reasoning ability demonstrated by the sample, and the problem-solving performance demonstrated by the sample. This is followed by a discussion on findings related to the first two research questions, namely:

1. Does computer-assisted instruction improve students’ problem-solving performance?
2. Does logical reasoning ability, as defined by Piagetian reasoning levels, have an effect on this relationship?

Finally, implications of the quantitative and qualitative results found in this study are presented in terms of theory, practice, and future research.

Discussion

Before discussing the results that emerged from use of the problem-solving pretest, posttest, and scoring scheme, their reliability is examined. In comparing pretest and posttest responses of participants, a significant positive correlation was found, \( r = 0.593, p < 0.01 \). In other words, students who scored high on the pretest also tended to score high on the posttest. Conversely, students who scored low on the pretest tended to score low on the posttest. These results suggested a consistency in the difficulty level of the two tests, thereby contributing to their overall reliability.

Statistical analyses revealed high interrater reliability for scores given independently by myself and my faculty advisor. Specifically, Kappa analysis showed a
high reliability on the scores given for the five pretest questions. Kappa analysis similarly revealed high reliability for each of the five posttest questions scored by the two raters. In a separate analysis, bivariate correlation revealed high interrater reliability for the total pretest scores of the two raters, as well as for the total posttest scores. Therefore, statistical analysis revealed a high level of interrater reliability for the scoring scheme and high reliability between the problem-solving pretest and posttest. Since my advisor and I were both familiar with the tests these reliability values may be higher than to be expected from raters who are not as familiar with the scoring scheme.

**Description of the Sample**

In review, the mean age of the participants was approximately twelve and a half years. This is the approximate age at which some people are reported to make the transition from the concrete stage to the formal stage of logical reasoning development. The sample consisted of slightly more females than males, but this small difference was not expected to have impacted the results. Also, none of the participants had reported using Science Sleuths prior to engaging in this study so this was not expected to have had an effect on the results of the study. However, it is possible that some students had more prior experience than others with software programs similar to Science Sleuths.

Participants scored much lower in logical reasoning ability than might have been expected for this age group. Within the sample there were only two students identified as having formal reasoning ability and fourteen as having transitional reasoning ability. A surprisingly large percentage of the sample (80.9%) were found to be in the concrete stage. These results are consistent with previous research which states that many late elementary school students are functioning below the formal-operational level (Bitner-
The content of the image is not legible due to the quality of the scan or the image itself.
Corvin, 1988; Gipson et al., 1989; Newton et al., 1981; Padilla et al., 1983; Smith & Sims, 1992; Staver, 1984; Tobin & Capie, 1981). However, it was surprising to find that the majority of participants in this study were functioning as low as the concrete stage of logical reasoning ability (rather than at the transitional stage).

These findings have serious implications for the curriculum that students are expected to learn and the methods we use to teach it. In Ontario, there has been a “pushing down” of curriculum through the grades, so that students are now exposed to certain concepts at a younger age. The developmental appropriateness of this new curriculum may be questionable as students may not be successful in learning some of the more difficult concepts at a younger age. Other researchers have similarly reported that there may be a discrepancy between what students are being asked to do in school and what they are developmentally capable of doing (Gipson et al., 1989; Smith & Sims, 1992).

Educators may not have control over what curriculum is taught in their classroom, but they often decide how it will be taught. It is recommended that educators be sensitive to differing levels of cognitive development among their students and attempt to use instructional methods to accommodate this variation (Gipson et al., 1989; Linn & Songer, 1991; Newton et al., 1981; Smith & Sims, 1992; Staver, 1984; Tobin & Capie, 1981; Yap & Yean, 1988). Since some research has indicated that most late elementary students do not have formal logical reasoning ability, it may be in their best interest to be exposed to concrete hands-on activities as much as possible (Clewell et al., 1992; Smith & Sims).
Problem-Solving Performance

On average, the sample population scored 11.40 on the pretest and 11.50 on the posttest. Since the total possible score for each test was 20, the participants scored in the low-average range. Future use of the test with a higher achieving population and/or older students may reveal higher scores. Further refinement of the problem-solving tests and accompanying scoring schemes, as well as a larger sample size, may also result in different scores.

The mean score for the entire group in problem-solving improvement was low (0.0952). This revealed a very small and insignificant improvement in problem-solving from the introductory lesson through to the end of the four problem-solving lessons (through CAI or traditional instruction). It is possible that an extended opportunity to develop problem-solving skills may have increased performance scores in either one or both groups.

Question 1: Effect of Instruction on Students’ Problem-Solving Performance

The mean pretest score for the traditional instruction group (11.34) and the CAI group (11.46) were not significantly different, t(82) = 0.168. This result was expected since both groups received the same introductory problem-solving lesson and had not received treatment prior to writing the pretest.

The mean posttest score for the traditional instruction group (11.37) and the CAI group (11.61) were not significantly different, t(82) = 0.371. In other words, neither group scored significantly higher on the posttest than the other after having received two different forms of problem-solving instruction. A very small improvement was detected from the pretest to the posttest for the traditional instruction group (0.15), as well as for
the CAI group (0.0263). Although the traditional instruction group had a slightly higher improvement score on average than the CAI group, this difference was not significant, \( t(82) = 0.210 \). Therefore, the first null hypothesis, that this CAI does not improve students’ problem-solving performance, was supported on the basis of the current findings. These findings are contrary to the literature claiming a positive effect of CAI on problem-solving performance (Eylon et al., 1996; Kinnear, 1986; Pek, 1996).

It is possible that, because limitations perceived in previous research were purposely avoided in the current study, discrepant results were found. For instance, Kinnear (1986) suggested that CAI improved problem-solving performance after comparing textbook-based problems and computer-based problems. The current study aimed to create a more equitable comparison by utilizing a more active instructional technique in the control group (i.e., a hands-on laboratory) to compare with CAI. Perhaps this difference in the control group led to a difference in findings between the current study and previous research. Additionally, research suggested that several factors affected the relationship between CAI and problem-solving performance (Cypher, 1987; McLean, 1989; Slack & Stewart, 1990). For instance, it was suggested that the teacher played a role in the effectiveness of problem solving through CAI. To avoid an effect of the teacher’s role in observed problem-solving performance, the current study made the role of the teacher in CAI and in traditional instruction as similar as possible. As an example, in both forms of instruction the teacher facilitated the problem-solving process by providing help to students during investigations as was needed. By keeping as many variables as similar as possible in the two treatments, it was hoped that they had not influenced the problem-solving performance in either group.
Despite these precautions, methodological limitations in the current study arose and must now be addressed. Most significantly, the problem-solving pretest, posttest and accompanying scoring scheme were developed over the course of this thesis, thereby providing only preliminary findings. Before confident conclusions can be drawn on the effect of CAI on problem-solving performance, further development of the tests and the scoring scheme is necessary. As well, the use of a larger sample population is recommended to confirm or refute the results of this study. Although the software and laboratory activities chosen for this study did not significantly improve problem-solving performance, it is possible that other software and laboratory activities might have yielded different results.

To summarize, this study did not provide strong conclusive evidence to support research for the effectiveness of this computer-based education (Slotnick, 1989; Friedler et al., 1989; Krajcik et al., 1988; Marsh & Kumar, 1992; Pek, 1996; Simmons & Lunetta, 1993; Soderberg et al., 1994), nor for research against its effectiveness (Eylon et al., 1996; Marsh & Kumar; Robertson, 1998), and continued study of these variables is recommended.

Question 2: Relationship Among Logical Reasoning Ability, Problem-Solving Performance, and Instruction

A significant positive relationship was found between problem-solving test score and logical reasoning ability. In other words, students who scored high on the GALT tended to score high on the problem-solving pretest, and vice versa, \( r = 0.405, p < 0.01 \). Similarly, students who scored high on the GALT tended to score high on the problem-solving posttest, and vice versa, \( r = 0.444, p < 0.01 \). The results of the current study
correspond to previous research which reported a connection between these two variables 
(Gipson et al., 1989; Smith & Sims, 1992; Stewart & Hafner, 1994).

In further analysis, no significant correlation between problem-solving 
 improvement scores and logical reasoning ability was found for the sample as a whole, \( r = 0.021, p > 0.05 \). Additionally, level of logical reasoning ability did not lead to a 
significant difference in the problem-solving improvement scores of students in the 
traditional instruction group compared to those in the CAI group. These results 
contrasted with research suggesting that computer models may better facilitate the 
problem-solving process for concrete learners, over traditional methods (Papert, 1980; 
Slotnick, 1980).

Furthermore, Krasnor & Mitterer (1984) hypothesized that students with formal 
logical reasoning ability would be most successful in learning through CAI and then in 
transferring general problem-solving strategies developed in CAI to other problem-
solving situations. This trend was observed in a graph of the results, but the interaction 
effect was not statistically significant. Specifically, in Figure V4, students with a high 
level of logical reasoning ability (i.e., those in the transitional and formal stages) showed 
greater improvement in problem-solving performance through CAI, as compared to those 
with high logical reasoning ability who received traditional instruction. Further research 
with students of high logical reasoning ability is needed to confirm or refute this 
observation.

Based on these findings the second null hypothesis, that logical reasoning ability 
as defined by Piagetian reasoning levels, has no effect on problem-solving performance, 
has been partially rejected. Specifically, scores on problem-solving tests were found to be
related to logical reasoning ability, but overall improvement in problem-solving
performance was not, regardless of the instructional method received by participants.

Implications

In this section, implications arising from the results of this study are discussed in
relation to practice, theory, and future research.

Implications for Practice

The results of this study indicated that the majority of participating late
elementary school students were in the concrete stage of logical reasoning development.
This has great implications for what we, as educators, teach our students. Namely, it
creates concern over the discrepancy between student needs and abilities, and the
challenging curriculum recently implemented in schools across Ontario. Also, as
educators, we need to be conscientious of our teaching methods in relation to student
needs and abilities. It is not feasible that every teacher test their students for logical
reasoning ability and to then plan lessons accordingly. However, teachers should be
sensitive to a variation in student abilities and needs and should aim to facilitate learning
by incorporating a variety of methods in their lessons. Furthermore, educators should be
aware of the need for greater “hands-on” and direct learning experiences in the
elementary school classroom.

The results of this study did not indicate that CAI improved student problem-
solving performance over traditional laboratory instruction. However, this study only
looked at one software program versus one particular laboratory activity. The greater the
variety of instructional methods used by educators, the more likely they will be to meet
the wide variety of interests and needs of their students. From this viewpoint, CAI may be seen as one valuable method of instruction.

The problem-solving pretest and posttest developed in this study may have practical usefulness for classroom teachers. Ross and Maynes (1983) suggested that "useful assessment instruments are ones that simultaneously describe existing performance and prescribe instructional acts that promote an upgrading of the performance measured" (p. 63). In this view, the problem-solving tests used in this study may serve to supplement teachers' assessment of their students' problem-solving performance, by providing them with an additional method for measuring the skill. In addition, sharing the selected student exemplars and the scoring rubric with students may assist teachers in improving student problem-solving performance. In other words, a teacher might show students what exemplars for each level look like, as well as to describe the criterion from the rubric for each level in order for students to gain knowledge on how to improve their own responses. The use of exemplars and rubrics for assessment purposes is in line with a recent trend of the Ministry of Education in Ontario, which has provided exemplars of different levels of student achievement in various subjects to Ontario teachers. It should be noted, however, that since the exemplars and rubrics were developed over the course of this study they were not available for instructional purposes by the participants. However, in future use of the problem-solving tests it is recommended that the rubrics and exemplars supplement problem-solving instruction at the discretion of the test administrator (e.g., openly discuss the assessment criterion in the rubric with the students prior to their writing either the pretest, posttest, or both tests).
Finally, it is hoped that the qualitative component of this thesis will serve as an easily accessible resource in Brock University’s Instructional Resource Centre. As an M.Ed. student I was often encouraged to read previous M.Ed. students’ research and found it to be helpful to my own research process. By reflecting on my own research experience in this report, perhaps I will shed some light into the process of conducting research at Brock University for future and current M.Ed. students. Maybe, the sharing of my experiences in a chronological manner will help other students as they think through the “next step” in their own research process.

Implications for Theory

The following is a summary of the theoretical implications of this study. To begin with, CAI was not found to have significantly improved problem-solving performance, over a traditional instructional technique. This finding lends support for those who question whether or not computer-based education is truly a more effective instructional technique over traditional instructional methods (Krajcik et al., 1986; Marsh & Kumar, 1992; Robertson, 1998).

In accordance with prior research, this study found that logical reasoning ability was positively correlated to problem-solving performance (Gipson et al., 1989; Smith & Sims, 1992; Stewart & Hafner, 1994). Additionally, there was some limited support for an effect of instructional technique on problem-solving scores of students with different logical reasoning abilities. This is in line with Krasnor & Mitterer’s (1984) hypothesis that students with formal logical reasoning ability would be most successful in learning through CAI and then in transferring general problem-solving strategies developed in CAI to other problem-solving situations. However, since the problem-solving pretest,
posttest, and accompanying scoring scheme were newly developed, quantitative findings were not strongly conclusive. Further development of the measures is necessary before drawing firm conclusions on the effectiveness of CAI on student problem solving and the role of logical reasoning ability in this relationship.

By its nature, qualitative research is meant to describe only a small part of the big picture, but in great detail. My analysis was based on only one person’s research experience, my own. It was by no means intended to describe the research process experienced by every Master of Education student at Brock University. The purpose was to uncover interesting patterns in my own journal notes to provide readers with some insight into what my thesis research process experience looked like. It was also meant to provide readers with background information on how this study was conducted. The qualitative findings lend support for the notion that the process of conducting research is a form of problem solving in itself, as was suggested by previous researchers (McGinn, 2000; Rashid, 1993). Specifically, I often reflected on the strategies I took to overcome difficulties faced in each stage of the research. For instance, when I faced the difficulty of recruiting participants for the study, the solution process involved contacting networks made in teacher’s college and asking my advisor for possible contacts.

Implications for Further Research

Assessment is often an enigma, whether in a research or a classroom setting. Skeptics may question if the assessment tool truly assesses what it is supposed to. Thus, the development of the problem-solving tests and accompanying scoring rubric was a daunting task. Although the validity and reliability of the tests and scoring scheme were addressed, they can be further developed through continued experimentation. It has been
similarly recommended by other researchers that continued experimentation is necessary in order to establish an instrument which reliably and validly measures a desired skill (Frederiksen, 1986; Ross & Maynes, 1983). Future research might include a comparison between these problem-solving tests and other pre-established tests of problem solving to determine the reliability between them. Another possible research avenue might be a comparison of scores on the same test between groups trained and untrained in problem solving, to see if the results obtained reflect the amount of training received (Frederiksen). I have also learned that it is necessary to answer test questions myself before administering it, in order to gain another perspective into its validity. If I was to become involved in similar research in the future I would be more knowledgeable about such practical considerations as a result of having conducted this study.

An important aspect of the problem-solving tests requiring further study and possible modifications is the use of extraneous data in the two problem scenarios. On the one hand, the extraneous data given in the problem scenario of the pretest was the amount of time taken to make the bubbles. Since participants were asked to decide which detergent should be purchased only on the basis of which would consistently produce larger bubbles, the time taken to blow the bubbles was not required to solve the problem and was therefore considered extraneous. Upon scoring student responses, I noticed that students varied in their use of the extraneous data in the solution process; some did not use the extraneous data at all, some used it incorrectly, and some were able to use it in a meaningful way. For instance, some students used the extraneous data incorrectly by concluding that the detergent which made the bubbles in the shortest amount of time would be the best purchase, rather than the bubbles which were the largest as the
question had asked. Only a few students used the extraneous data in a meaningful way. For example, one student calculated the average time taken to make each bubble as well as the average bubble size. This student reasoned that although the time taken to make the larger bubble was slightly longer it was not significant enough to influence the choice of which detergent to purchase and correctly concluded that the detergent that consistently made the larger bubble would be the better purchase.

In comparison, the extraneous data used in the posttest was the amount of time taken to dissolve potassium chloride. Since participants were asked to determine if unknown chemicals were either potassium iodide or sodium chloride by comparing rates of dissolving, the time taken to dissolve potassium chloride was extraneous. Once again, it was found that some students did not use the extraneous information at all, some used it incorrectly, and some used it in a meaningful way. An example of incorrect usage by a student was naming one of the unknown chemicals as potassium chloride without any rationale for doing so. On the other hand, a student who used the extraneous information in a meaningful way suggested that maybe a mix-up occurred in the ordering of the chemicals, and one of the unknowns was potassium chloride since their dissolving rates were closest to each other. After conducting the study, I realized that the extraneous data in the pretest might have been easier to incorporate in a meaningful way, as compared to the extraneous data in the posttest. Consequently, I designed the scoring rubric to accommodate this potential imbalance between the pretest and the posttest. Students who used the data in a meaningful way and those who did not use it at all were scored equally; in this way, only students who used extraneous data in an incorrect way were penalized
in the scoring scheme. Perhaps future versions developed through further study might better balance the use of extraneous data in the pretest and posttest.

An interesting trend was found in the results from students with high logical reasoning ability. Since the finding was not statistically significant, it is recommended that further research be conducted to determine if students of high logical reasoning ability do indeed have more success in improving problem-solving performance through a computer-based instructional technique, over a traditional instructional method.

Although the software and laboratory activities chosen for this study did not significantly improve problem-solving performance, it is possible that other software and laboratory activities might have yielded significant results. This is another potential area for future study.

The qualitative findings in this study were limited to my own research experience. I suggest that a more thorough analysis be conducted with several data sources and participants in order to expand on the current description of the M.Ed. thesis research process at Brock University. For example, it would be interesting to look for reoccurring themes from a more theoretical, rather than descriptive, perspective across the research experiences of several Brock University Master of Education students.

Summary

This chapter was a discussion of the research findings, followed by implications for theory, practice and future research. I will now conclude this report by briefly summarizing the findings on the three research questions of this study. To answer the first question, data from this study suggested that this form of CAI did not significantly improve students’ problem-solving performance, over this traditional instruction. To
answer the second question, it was found that logical reasoning ability was correlated to problem-solving performance, as students with high logical reasoning ability did better on the problem-solving measures and vice versa. However, there was no significant difference in problem-solving improvement for students in the traditional instruction group versus the CAI group in relation to their logical reasoning ability. Having said this, a trend observed in the data was that students with high logical reasoning ability seemed to show greater improvement in problem-solving performance through CAI than traditional methods of instruction. This finding was not statistically significant but may warrant further study as it has been suggested by other researchers as well (Krasnor & Mitterer, 1984). The third question was addressed through a personal narrative based on my own research journal notes. I found that my thesis research experience as a Brock University Master of Education student was centered around six main areas: research interests, practical considerations, research design, research analysis, development of the problem-solving tests, and scoring scheme development. In accordance with previous authors, I found that my research experience was a form of problem solving in itself. Specifically, I relied on several problem solving strategies, such as meetings with my advisor and the use of models, to overcome difficulties faced throughout the research process.

McGinn (2000) reflected that the research experience helped her to develop her identity on many levels and I have found that to be true for myself as well. This thesis research has strengthened my role as a qualitative and quantitative researcher, as well as an elementary school teacher. I have greater insight into problem solving because of the research that was conducted and because of my reflection on the research process itself.
REFERENCES


...


Appendix A

Ontario Science and Technology Curriculum Expectations (MoET, 1998)

The following specific Ontario Curriculum expectations are targeted by both types of instruction:

- Formulate questions about and identify needs and problems related to the characteristics of mixtures and solutions, and explore possible answers and ways of meeting these needs
- Communicate the procedures and results of investigations for specific purposes and to specific audiences
- Use appropriate vocabulary, including correct science and technology terminology, to communicate ideas, procedures, and results
- Follow safe work procedures and use appropriate tools, materials, and equipment

The following specific Ontario Curriculum expectation is only targeted by the Traditional Instruction:

- Demonstrate different methods of separating the components of mixtures and describe some industrial applications of these methods

The following specific Ontario Curriculum expectations are only targeted by the CAI:

- Describe the concentration of a solution in qualitative terms (e.g., dilute, concentrated) and in quantitative terms (e.g., grams of solute per 100 ml)
- Plan investigations for some of these answers and solutions, identifying variables that need to be held constant to ensure a fair test and identifying criteria for assessing solutions
- Differentiate between raw materials (e.g., wood, coal, natural gas) and processed materials (e.g., plastic, glass, ceramic)
- Identify a variety of manufactured products made from mixtures or solutions and explain their functions (e.g., medicines, cleaning solutions, salad dressings)
- Identify the sources and characteristics of pollutants that result from manufacturing and agricultural systems
- Describe the effects of some solvents on the environment, and identify regulations that are in place to ensure their safe use and disposal
- Identify different types of waste present in the community (e.g., water, sewage, trash, toxic materials) and the environmental considerations related to their disposal
Appendix B

Brock University, Department of Graduate and Undergraduate Studies in Education
LETTER OF INFORMATION (FOR PARENTS/GUARDIANS)

Title of Study: Computer-Assisted Instruction in Science Education: An Inquiry into Problem-Solving Performance and Logical Reasoning Ability

Researchers: J. Skinner and Dr. Joe Engemann (Supervising Professor)

Dear Parent/Guardian,

I would like to begin this letter by introducing myself. My name is Miss J. Skinner and I have been teaching Science at Glen Brae Middle School since mid-March. Prior to that I was a student-teacher, as well as a grade six English teacher at Glen Brae. I am also currently a full-time Master of Education student at Brock University. The purpose of this letter is to ask for your permission for your child to participate in a meaningful research study that I will be conducting in the near future.

The purpose of this research is to document middle school students problem solving performance when taught these skills through the use of a computer. Grade seven and eight students at Glen Brae Middle School will participate in several science lessons designed to improve their problem solving performance. Problem solving is an important component of science and mathematics education and is one of the learning skills assessed in the new Ontario Report Cards.

The main data sources for the research project will be student assignments and tests. Student participation in the research involves the writing of a mathematical test of reasoning ability, as well as participation in lessons which meet the requirements of regular Science and Technology instruction. Agreement to your child’s participation will allow the use of his/her class assignments as evidence of his/her learning in this class.

The following assignments/tests will be completed and submitted by students:
1. A mathematical test of reasoning ability (called the Group Assessment of Logical Reasoning Ability).
2. Participation in a lesson on problem-solving strategies.
4. Problem-solving instruction through the use of a computer.
5. Problem-solving instruction in a science classroom setting.

The reason for including the mathematical reasoning ability test is to better understand the needs of students in middle school science and to then adapt instruction to meet these
needs. Previous research suggests that there is a link between a students’ problem-solving performance and his/her reasoning ability (Gipson, Abraham, & Renner, 1989; Smith & Sims, 1992; Stewart & Hafner, 1994). Researchers have also suggested that these skills are linked to computer-assisted instruction (Krajcik, Simmons, & Lunetta, 1988).

By allowing your child to participate in the study you are giving me permission to consider his/her assignments, tests, and learning experiences in my evaluation of middle-school students’ problem-solving performance. This evaluation may provide guidance for other educators aiming to maximize problem-solving performance of their students. The research also provides valuable information to educators on the merit of computers in science education.

Participation in this study is completely voluntary and your child may withdraw from the study at any time and for any reason without penalty. Your child’s grade in Science and Technology will not be affected by your decision on to allow them to participate or to not participate in the research.

Although students must complete class assignments and tests for grading purposes, they have the option of participating in the research beyond those requirements. If your permission for their participation is not granted, your child’s performance on assignments and tests will not be analyzed for what they have learned or contributed to the project.

There is no obligation for your child to participate in any aspect of this project that he/she considers invasive, offensive, or inappropriate. All personal data will be kept strictly confidential and all information will be coded so that your child’s name will not be associated with his/her work. Only the researchers will have access to this data. This study aims to be nonintrusive to the regular routine of science instruction. There are no known or anticipated risks associated with your child’s participation in this research.

The Research Ethics Board of Brock University has officially approved this study (File # 99-272). Copies of the project’s final report will be available during the month of January 2001 at Glen Brae Middle School.

Thank you very much for taking the time to consider this request.

Sincerely,

J. Skinner
**Brock University, Department of Graduate and Undergraduate Studies in Education**

**INFORMED CONSENT FORM (FOR PARENTS/GUARDIANS)**

<table>
<thead>
<tr>
<th>Title of Study:</th>
<th>Computer-Assisted Instruction in Science Education: An Inquiry into Problem-Solving Performance and Logical Reasoning Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researchers:</td>
<td>J. Skinner and Dr. Joe Engemann (Supervising Professor)</td>
</tr>
<tr>
<td>Name of Student:</td>
<td></td>
</tr>
</tbody>
</table>

I give permission for my child to participate in the research study described in the attached Letter of Information.

I understand that my child’s participation in this research involves the writing of a mathematical test of reasoning ability and that all other activities meet the requirements of regular science instruction. I understand that by allowing my child to participate in the research, I have given permission for use of his/her class assignments and tests as evidence of his/her learning in this class. I understand that my child’s grade in Science and Technology will be unaffected by my decision to allow him/her to participate or not to participate in the research.

I understand that student participation in this study is voluntary and that he/she may withdraw from the study at any time and for any reason without penalty.

I understand that there will be no payment for student participation.

I understand that there is no obligation to the student to answer any question/participate in any aspect of this project that he/she may consider invasive, offensive or inappropriate.

I understand that all personal data will be kept strictly confidential and that all information will be coded so that student name is not associated with answers. I understand that only the researchers named above will have access to the data.

Parent/Guardian Signature_________________________________________ Date

This study has been reviewed and approved by the Brock Research Ethics Board. (File # 99-272)

If you have any questions or concerns about your child’s participation in the study, you may contact J. Skinner at Glen Brae Middle School, (905) 560-6732, or Dr. Joe Engemann (supervising professor) at Brock University, (905) 688-5550 extension 3943.

Feedback about the use of the data collected will be available during the month of January 2001 from Glen Brae Middle School. A written explanation will be provided for you upon request.

Thank you for your help! Please keep the information letter for further reference.

I have fully explained the procedures of this study to the above student.

Researcher Signature_________________________________________ Date ___________________
Appendix C
Problem-Solving Pretest - Part A

The Problem Scenario:

David is a camp counselor at Kindergarten Camp. As he is preparing for his campers to arrive he remembers that making big bubbles was a popular activity last summer. He decides that he will have the campers do it again this year. David will make the bubble solution for his campers by mixing liquid detergent with water.

Making a trip to the grocery store David finds two brands of liquid detergent that are in his price-range (Lemon-Fresh and Clean-All) and one that is too expensive (Spotless). He only wants to use one brand of detergent all summer but he is not sure which one will make the biggest bubbles. He buys one small bottle of Lemon-Fresh Detergent and one small bottle of Clean-All Detergent to try out first. What experiment can David do to solve this problem?
Chapter 5
Effect of Administration

Administration plays a crucial role in the effectiveness of systems. In many cases, the way a system is administered can significantly impact its performance.

For instance, in the case of a distribution system, the administration can affect the flow and distribution of resources. Effective administration can ensure that resources are allocated efficiently, minimizing waste and maximizing efficiency.

Additionally, administration can influence the user experience. If a system is well-administered, users are more likely to find it intuitive and user-friendly, leading to higher satisfaction and productivity.

In conclusion, the importance of administration cannot be overstated. It is essential for ensuring the success of any system, whether it be for resource allocation, user experience, or overall efficiency.

Conclusion


Write your name, class, and the date on the back of this test.

Problem-Solving Pretest - Part A Response Sheet

1. Write the problem in your own words.

2. What do you need to know?

3. (a) How are you going to gather this information?

   (b) How are you going to record this information (be specific)?
Problem-Solving Pretest - Part B

Back at the camp David performed an experiment to find out which liquid detergent most consistently made the biggest bubbles. He used three times as much detergent as water in the bubble solution because it was the ratio that produced the biggest bubbles last summer.

First, he mixed 15 mL of Lemon-Fresh with 5mL of water. He made a bubble with this solution and then measured the diameter of the bubble in cm. He repeated this test nine more times. The size of the bubble diameters were: 24 cm, 23 cm, 22 cm, 25 cm, 23 cm, 27 cm, 27 cm, 28 cm, and 30 cm. The time taken to make each of these bubbles was: 30 s, 29 s, 28 s, 24 s, 25 s, 26 s, 27 s, 25 s, 26 s, and 29 s.

Next David mixed 15 mL of Clean-All with 5 mL of water. He made a bubble with this solution and then measured the diameter of the bubble in cm. He repeated this test nine more times. The size of the bubble diameters were: 24 cm, 28 cm, 28 cm, 25 cm, 28 cm, 27 cm, 27 cm, 29 cm, and 29 cm. The time taken to make each of these bubbles was: 29 s, 29 s, 28 s, 29 s, 25 s, 28 s, 27 s, 25 s, 28 s, and 29 s.

Help David to decide, from the results of this experiment, which brand of liquid detergent (Lemon-Fresh or Clean-All) to buy for his campers this summer.
Problem-Solving Pretest - Part B Response Sheet

4. Solve the problem (show your work).

5. (a) Does your answer make sense? Why or why not?

(b) Does your solution agree with the evidence? Why or why not?
Appendix D

Problem-Solving Posttest - Part A

The Problem Scenario:

Mrs. Chung is a science teacher at Pleasant Valley Middle School. The chemicals that she ordered for student lab work have finally arrived at her school. When she opens the delivery package she finds that the labels have fallen off the container of sodium chloride and the container of potassium iodide. As a result, she does not know which container contains which chemical.

Mrs. Chung looks up these two chemicals in a science book to find out some information about them. She discovers that the two chemicals do not take the same amount of time to dissolve in water.

She decides to investigate which container holds sodium chloride and which container holds potassium iodide using the information from the science book. What experiment can Mrs. Chung do to solve this problem?
null
Write your name, class, and the date on the back of this test.

Problem-Solving Posttest - Part A Response Sheet

1. Write the problem in your own words.

2. What do you need to know?

3. (a) How are you going to gather this information?

(b) How are you going to record this information (be specific)?
Mrs. Chung performed an experiment to find out which unlabeled container holds sodium chloride and which one holds potassium iodide.

First, she timed how long it took for 5 grams of the chemical in Container 1 to dissolve in 100 mL of water. She repeated this test nine more times. The time taken for it to dissolve was: 40 s, 35 s, 35 s, 39 s, 38 s, 42 s, 38 s, 42 s, 41 s, and 37 s.

Then Mrs. Chung timed how long it took for 5 grams of the chemical in Container 2 to dissolve in 100 mL of water. She repeated this test nine more times. The time taken for it to dissolve was: 50 s, 54 s, 54 s, 50 s, 48 s, 47 s, 48 s, 52 s, 41 s, and 47 s.

Mrs. Chung’s science book states that it takes 50 seconds to dissolve 5 grams of sodium chloride in 100 mL of water, 49 seconds to dissolve 5 grams of potassium chloride in 100 mL of water, and 40 seconds to dissolve 5 grams of potassium iodide in 100 mL of water.

Use the results of the experiment and the information from the science book to help Mrs. Chung correctly label the chemical in Container 1 and the chemical in Container 2.
Write your name, class, and the date on the back of this test.

Problem-Solving Posttest - Part B Response Sheet

4. Solve the problem (show your work).

5. (a) Does your answer make sense? Why or why not?

(b) Does your solution agree with the evidence? Why or why not?
<table>
<thead>
<tr>
<th>Level 4</th>
<th>Level 3</th>
<th>Level 2</th>
<th>Level 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write a well-reasoned response.</td>
<td>Include unnecessary details and/or focus on unrelated topic.</td>
<td>omissions: 2 or more</td>
<td>omissions: 3 or more</td>
</tr>
<tr>
<td>Does not give specific examples and/or includes unnecessary details.</td>
<td>Does not give specific examples and/or includes unnecessary details.</td>
<td>omissions: 2 or more</td>
<td>omissions: 3 or more</td>
</tr>
<tr>
<td>Does not refer to the use of chemical information.</td>
<td>Does not refer to the use of chemical information.</td>
<td>omissions: 2 or more</td>
<td>omissions: 3 or more</td>
</tr>
</tbody>
</table>

**Problem-Solving Process - Scoring Rubric**

(All answers are not required to be written in the sections provided below.)

**Problem-Solving Process - Scoring Rubric**

Appendix E
<table>
<thead>
<tr>
<th>Step 5:</th>
<th>Score: Solving Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not reflect any of the possible sources or error(s) or reflects on any of the possible sources or error(s)</td>
<td>APPLY THE SOLUTION</td>
</tr>
<tr>
<td>Reflects on one of the three mistake types as described above in the problem solving strategy.</td>
<td>APPLY THE SOLUTION</td>
</tr>
<tr>
<td>Reflects on two of the three mistake types as described above in the problem solving strategy.</td>
<td>APPLY THE SOLUTION</td>
</tr>
<tr>
<td>Reflects all three mistake types as described above in the problem solving strategy.</td>
<td>APPLY THE SOLUTION</td>
</tr>
</tbody>
</table>

LEVEL 3

LEVEL 2

LEVEL 1

LEVEL 4

(See answers are not restricted to be written anywhere on response sheet part B.

* Rubric Criterion CAN BE WRITTEN ANYWHERE ON RESPONSE SHEET PART B.

Problem Solving Project Part B - Scoring Rubric
**Problem Solving Process Part A - Scoring Rubric**

**Appendix F**

*Rubric criteria can be written anywhere on response sheet part A.*

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Problem-solving strategy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes a well-organized process (includes unnecessary steps, includes headings and subheadings, with an overall topic, date, level)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes a poorly organized process (not included all of the steps that they will use)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does not refer to the use of a table or any other graphic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicates the necessity of the information step</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 1:** Define the problem

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrased clearly, concise (i.e., contains adequate description of the problem to be solved, and all the necessary steps required to solve it)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrased clearly, concise (i.e., contains the problem to be solved, and all the necessary steps required to solve it)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrased clearly, concise (i.e., contains the problem to be solved, and all the necessary steps required to solve it)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguously phrased (i.e., contains the problem to be solved, and all the necessary steps required to solve it)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 2:** Develop a plan

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contains a description of a plan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contains a description of a plan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contains a description of a plan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguously phrased (i.e., contains the problem to be solved, and all the necessary steps required to solve it)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 3:** Execute the plan

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describes the other variable (i.e., contains the variables which are needed and unnecessary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully describes the other variable (i.e., no less than 2 variables, and the variables are fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 4:** Analyze the results

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describes the other variable (i.e., contains the variables which are needed and unnecessary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes the other variable (i.e., no less than 2 variables, and the variables are fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 5:** Interpret the results

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describes the other variable (i.e., contains the variables which are needed and unnecessary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes the other variable (i.e., no less than 2 variables, and the variables are fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 6:** Evaluate the results

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describes the other variable (i.e., contains the variables which are needed and unnecessary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes the other variable (i.e., no less than 2 variables, and the variables are fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 7:** Communicate the results

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describes the other variable (i.e., contains the variables which are needed and unnecessary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describes the other variable (i.e., no less than 2 variables, and the variables are fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially describes the other variable (i.e., only 1 or 2 variables, and the variables are not fully described)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conclusion Was Drawn</td>
<td>Possible Sources of Error</td>
<td>Possible Sources of Error</td>
<td>Possible Sources of Error</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Step 1: Write the solution</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Problem Solving Part B - Scoring Rubric**

- **Level 1**: Incorrect conclusion and incorrect definition of the question.
- **Level 2**: Incorrect conclusion and correct definition of the question.
- **Level 3**: Correct conclusion and correct definition of the question.
- **Level 4**: Correct conclusion and correct definition of the question.

(Rubric criteria can be written anywhere on response sheet Part B)
Appendix G

Problem-Solving Pretest - Part A - Possible Solutions

1. Write the problem in your own words.

Which brand of detergent, Clean-All or Lemon-Fresh, will make bigger bubbles most consistently?

2. What do you need to know?

I will need to design a fair test to determine which brand of detergent most consistently makes larger bubbles (i.e., hold constant the volume of water and detergent and do repeated trials of the same test).

3. (a) How are you going to gather this information?

I will record the size of each bubble’s diameter (in cm).

(b) How are you going to record this information (be specific)?

In a table such as:

| Volume of Water (mL) | Volume of Lemon-Fresh (mL) | Volume of Clean-All (mL) | Bubble Size (cm) |
Problem-Solving Pretest - Part B - Possible Solutions

4. Solve the problem (show your work).

Time taken to make the bubbles was not included in the problem scenario so it is omitted from the solution.

Average Size of Lemon-Fresh Bubbles:
\[ 24 + 23 + 22 + 25 + 23 + 22 + 27 + 27 + 28 + 30 = 251/10 \text{ trials} = 25.1 \text{ cm} \]

Average Size of Clean-All Bubbles:
\[ 24 + 28 + 28 + 25 + 28 + 28 + 27 + 27 + 29 + 29 = 273/10 \text{ trials} = 27.3 \text{ cm} \]

Therefore, on average, Clean-All detergent produced bigger bubbles than Lemon-Fresh detergent. David should buy Clean-All detergent for his campers.

5. (a) Does your answer make sense? Why or why not?

My solution does make sense because a fair test was used to compare bubble sizes produced by the two detergents that he could afford to buy.

(b) Does your solution agree with the evidence? Why or why not?

Yes, because I mathematically calculated the average and found that Clean-All produced bigger bubbles than Lemon-Fresh.
Appendix H

Problem-Solving Posttest - Part A - Possible Solutions

1. Write the problem in your own words.

Which one of the containers holds sodium chloride and which one contains potassium iodide?

2. What do you need to know?

I will need to design a fair test to determine the time it will take the chemical in Container 1 to dissolve and the time it will take the chemical in Container 2 to dissolve (i.e., hold constant the volume of water and the amount of solute in each test; repeat trials for each test). I will then compare these values to those in the textbook to identify them.

3. (a) How are you going to gather this information?

I will record the time taken for Chemical 1 to dissolve and for Chemical 2 to dissolve (in seconds). I will then review the values given in the textbook for comparison.

(b) How are you going to record this information (be specific)?

In a table such as:

<table>
<thead>
<tr>
<th>Volume of Water (mL)</th>
<th>Volume of chemical in Container 1 (mL)</th>
<th>Volume of chemical in Container 2 (mL)</th>
<th>Time taken to dissolve (s)</th>
</tr>
</thead>
</table>

Problem-Solving Posttest - Part B - Possible Solutions

4. Solve the problem (show your work).

**Average Time taken for 5g of the chemical in Container 1 to dissolve in 100mL of water:**

\[
40 + 35 + 39 + 38 + 42 + 38 + 41 + 37 = 387/10 \text{ trials } = 38.7 \text{ cm}
\]

**Average Time taken for 5g of the chemical in Container 2 to dissolve in 100mL of water:**

\[
50 + 54 + 50 + 48 + 47 + 48 + 52 + 41 + 47 = 491/10 \text{ trials } = 49.1 \text{ cm}
\]

From the science book 2 of the 3 values given are relevant to this problem (i.e., the labels fell off sodium chloride and potassium iodide, so the value for potassium iodide is irrelevant).

The chemical in Container 1 dissolved (on average) in 38.7 s which is closest to the value given for potassium iodide to dissolve (40 s). The chemical in Container 2 dissolved (on average) in 49.1 s which is closest to the value given for potassium iodide to dissolve (50 s).

Therefore, Container 1 holds potassium iodide and Container 2 holds sodium chloride.

5. (a) Does your answer make sense? Why or why not?

My solution does make sense because a fair test was used to determine time taken for each unknown chemical to dissolve. A comparison was then made to known values in order to correctly identify the two containers without labels. Potassium chloride was not used in the identification because it was extraneous information to this problem.

(b) Does your solution agree with the evidence? Why or why not?

Yes, because Container 1 had an average dissolving rate closest to potassium iodide and Container 2 had an average dissolving rate closest to NaCl.
Appendix I

Student Exemplars of Problem-Solving Test Responses

STEP 1: DEFINE THE PROBLEM
Level 1 Exemplar
Omits the test criterion and omits the variables to be tested.
Participant #55

1. Write the problem in your own words.
   My problem is he is supposed do smthing on his on because it you do it with the kids first you have to shave them have to.

2. What do you need to know?
   You need to know how to make it and you are supposed to be sure when it this year to both I will do it slowe and safely.
STEP 1: DEFINE THE PROBLEM
Level 1 Exemplar
Ambiguously refers to either the test criterion or the variables to be tested (i.e., chemicals) and omits the other.
Participant #42

1. Write the problem in your own words.

Mrs. Chung doesn't know which container holds which chemical. She also needs to know what kind of experiment she should use to figure out which is which.
STEP 1: DEFINE THE PROBLEM
Level 2 Exemplar
Ambiguously refers to test criterion (i.e., better) and ambiguously refers to variables to be tested (e.g., detergents, products).
Participant #4

1. Write the problem in your own words.

he could try out each of the detergents and see which one is better than the other.
STEP 1: DEFINE THE PROBLEM
Level 2 Exemplar
Specifies either the test criterion or the variables to be tested (i.e., Lemon-Fresh or Clean-All) and omits the other.
Participant #70

1. Write the problem in your own words. David wants to make bubbles for the kids, but he doesn't know which detergent to use: Lemon Fresh or Clean All, so he buys both.
1. Write the problem in your own words.

David needs to find out which product will make bigger bubbles.
154

STEP 1: DEFINE THE PROBLEM
Level 4 Exemplar
Specifies test criterion (i.e., What makes larger bubbles) and specifies variables to be tested (i.e., Clean-All and Lemon-Fresh detergents).
Participant #14

1. Write the problem in your own words.

What makes bigger bubbles Lemon-Fresh or Clean-All?
STEP 2: DEVELOP A PLAN
Level 1 Exemplar
Omits a description of a fair test and omits a description of the basic variables.
Participant #2

2. What do you need to know?

You first have to know how much liquid stuff you wanna make and how many kids there is going to be.
STEP 2: DEVELOP A PLAN
Level 1 Exemplar
Partially describes either a fair test or the variables (i.e., how much time does each chemical take to dissolve in water) and omits the other.
Participant #69

2. What do you need to know?

- How much time does each chemical take to dissolve in water
STEP 2: DEVELOP A PLAN
Level 2 Exemplar
Partially describes a fair test (i.e., only 1 element: you need to do a couple of experiments) and partially describes the basic variable (i.e., only 2: how much detergent; how much water).
Participant #11

2. What do you need to know?

You need to know how much detergent to mix in with the water. And how much water to put in a bowl or container.

3. (a) How are you going to gather this information?

You need to do a couple of experiments to figure out the exact amount of water and detergent to mix to make a good bubble solution.
STEP 2: DEVELOP A PLAN
Level 2 Exemplar
Fully describes either a fair test or the variables (i.e., at least 3: how much detergent; how much water; size) and omits the other.
Participant #10

2. What do you need to know?
- which brand makes bigger bubbles
- how much detergent needs to be used
- how much water needs to be added
- how long the bubble takes to pop
- how many bubbles can be made
- how many big bubbles can be made
- how big bubble is size (small medium)

3. (a) How are you going to gather this information?
- do an experiment
- find out which detergent makes bigger bubbles
- find out how long the bubble takes to pop
- how many bubbles can be made
- chart
- presentation
- diagrams
STEP 2: DEVELOP A PLAN
Level 3 Exemplar
Partially describes either a fair test or the basic variables (i.e., only 1 variable: how long it takes each chemical to dissolve) and fully describes the other (i.e. 2 elements: do the experiment with each chemical several times).
Participant #37

2. What do you need to know?

We need to know how fast each of the chemicals dissolve in water because Mrs. Chung looked them up and found out that they dissolve in water at different speeds.

3. (a) How are you going to gather this information?

We could gather it by doing the experiment with each chemical several times and seeing exactly how long it takes each chemical to dissolve.
STEP 2: DEVELOP A PLAN
Level 4 Exemplar
Fully describes a fair test (i.e., 2 elements: do that 5 times; for each one) and fully
describes the basic variables (i.e., at least 3: 100 mL of water; 10 mL of one chemical;
record the time).
Participant #52

2. What do you need to know?
- Which one takes more time to dissolve.

3. (a) How are you going to gather this information?
- 100 mL of water
- 10 mL of one chemical
- Stir till it dissolves
- Record the time
- Do that 5 times for each one
- Same temp.
- Of the water
STEP 3: GATHER INFORMATION
Level 1 Exemplar
Does not refer to the use of a table or any other effective method for organizing data.
Participant #14

(b) How are you going to record this information (be specific)?

Measure how long the bubble grows by having a friend to measure it with a measuring tape, and repeating the test
STEP 3: GATHER INFORMATION
Level 2 Exemplar
States that they will use a table (i.e., I will record this in a chart), or another effective means to organize data, but does not give specific details.
Participant #5

(b) How are you going to record this information (be specific)?

I will record this in a chart
### STEP 3: GATHER INFORMATION

**Level 2 Exemplar**

Describes a meaningless table (i.e., table headings are substance; difference; results).

Participant #14

---

(b) How are you going to record this information (be specific)?

<table>
<thead>
<tr>
<th>Substance</th>
<th>Difference</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.C.</td>
<td>tastes like salt</td>
<td>tastes like salt</td>
</tr>
<tr>
<td>P.I.</td>
<td>does not taste like salt</td>
<td>does not taste like salt</td>
</tr>
</tbody>
</table>

**Short example**

```
<table>
<thead>
<tr>
<th>Substance</th>
<th>Difference</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.C.</td>
<td>tastes like salt</td>
<td>tastes like salt</td>
</tr>
<tr>
<td>P.I.</td>
<td>does not taste like salt</td>
<td>does not taste like salt</td>
</tr>
</tbody>
</table>
```
STEP 3: GATHER INFORMATION
Level 3 Exemplar
Describes a poorly organized table (i.e., does not include all of the needed headings and/or includes unnecessary headings: potassium iodide and sodium chloride instead of Chemical 1 and Chemical 2).
Participant #78

(b) How are you going to record this information (be specific)?
In a chart like this:

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium Iodide</td>
<td></td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td></td>
</tr>
</tbody>
</table>
STEP 3: GATHER INFORMATION
Level 4 Exemplar
Describes a well-organized table (i.e., headings such as Detergent Type [Lemon-Fresh or Clean-All]; Amount of Detergent; Amount of Water; Size of Bubbles).
Participant #30

(b) How are you going to record this information (be specific)?

In a table:

<table>
<thead>
<tr>
<th>Detergent type (Lemon-Fresh or Clean-All)</th>
<th>Amount of Detergent</th>
<th>Amount of Water</th>
<th>Size of Bubbles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
STEP 4: SORT-LINK-SOLVE
Level 1 Exemplar
Computes an incorrect answer (i.e., Lemon-Fresh = 24.8 cm; Clean-All = 27.1 cm) and reaches an incorrect conclusion (i.e., no conclusion given) and uses extraneous data in a non-meaningful way (i.e., average time to make Lemon-Fresh and Clean-All bubbles is calculated but not used in a conclusion).
Participant #1

4. Solve the problem (show your work).

Lemon-Fresh =
\[
\begin{align*}
24 + 23 + 22 + 25 + 23 + 27 + 28 + 30 &= 168 \div 9 \\
24.8 \text{ cm} \\
30 + 29 + 28 + 24 + 25 + 26 + 27 + 25 + 26 + 29 &= 269 \div 9 \\
29.8 \text{ sec}
\end{align*}
\]
The average size for the bubbles is 24.8 cm.
The average time for a bubble to be that size is 29.8 seconds.

Clean-All
\[
\begin{align*}
24 + 28 + 25 + 28 + 28 + 27 + 27 + 29 &= 244 \div 9 \\
27.1 \text{ cm} \\
29 + 29 + 28 + 29 + 25 + 28 + 25 + 28 + 29 &= 277 \div 9 \\
27.7 \text{ cm}
\end{align*}
\]
The average size for the bubbles is 30.7 cm.
The average time for a bubble to be that size is 30.7 seconds.
STEP 4: SORT-LINK-SOLVE
Level 2 Exemplar
Does any one of the following: computes a correct answer (i.e., Container 1 time [sec] = 38.7; Container 2 time [sec] = 49.1) or reaches a correct conclusion or uses extraneous data in a meaningful way or not at all.
Participant #59

4. Solve the problem (show your work).

<table>
<thead>
<tr>
<th>Container #1</th>
<th>Container #2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time (Sec.)</strong></td>
<td><strong>Time (Sec.)</strong></td>
</tr>
<tr>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>35</td>
<td>55</td>
</tr>
<tr>
<td>39</td>
<td>45</td>
</tr>
<tr>
<td>42</td>
<td>52</td>
</tr>
<tr>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>37</td>
<td>47</td>
</tr>
</tbody>
</table>

$387 \div 10 = 38.7$

$38.7 \div 10 = 3.87$

$3.87 \times 8 = 30.96$

$30.96 \div 80 = 0.387$

Participant #1 must be potassium iodine, and Container #2 must be potassium chloride because it takes 49.1 sec to dissolve in water.
STEP 4: SORT-LINK-SOLVE
Level 3 Exemplar
Does any two of the following: computes a correct answer and/or reaches a correct conclusion (i.e., potassium iodide = container 1; sodium chloride = container 2) and/or uses extraneous data in a meaningful way or not at all (i.e., does not use extraneous data at all).
Participant #76

4. Solve the problem (show your work).

<table>
<thead>
<tr>
<th>Cont. 1</th>
<th>Cont. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
<td>150</td>
</tr>
<tr>
<td>35</td>
<td>54</td>
</tr>
<tr>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>39</td>
<td>50</td>
</tr>
<tr>
<td>38</td>
<td>48</td>
</tr>
<tr>
<td>38</td>
<td>47</td>
</tr>
<tr>
<td>42</td>
<td>47</td>
</tr>
<tr>
<td>41</td>
<td>47</td>
</tr>
<tr>
<td>387</td>
<td>416^1</td>
</tr>
</tbody>
</table>

potassium iodide = (38.7) container 1.
sodium chloride = (46.1) container 2.
STEP 4: SORT-LINK-SOLVE
Level 4 Exemplar
Computes a correct answer (e.g., Lemon-Fresh bubbles = 25.1 or 25 rounded off; Clean-All = 27.3 or 27 rounded off) and reaches a correct conclusion (i.e., If Clean-All makes bigger bubbles) using extraneous data in a meaningful way (i.e., He is still going to stick with Clean-All because I don’t think 1 second makes a difference) or not at all.
Participant #37

4. Solve the problem (show your work).

First we should find the average by adding them all together and then dividing by how many numbers there are.

So Lemon Fresh: \[24 + 23 + 22 + 25 + 23 + 22 + 27 + 27 + 28 + \frac{251}{30} \div 10 = 25.1 = 25 \text{ (rounding off)}\]
Clear All: \[24 + 28 + 28 + 25 + 28 + 27 + 27 + 29 + 29 = 273 \div 10 = 27.3 = 27\]

So now it is obvious that Clearall makes bigger bubbles, but there is always how long it took for:

Lemon Fresh: \[30 + 29 + 28 + 24 + 26 + 27 + 25 + 26 + 29 = 269 \div 10 = 26.9 = 27\]
Clear-All: \[29 + 29 + 28 + 29 + 25 + 28 + 27 + 26 + 28 + 29 = 277 \div 10 = 27.7 = 28\]. He is still going to stick with Clear-All because I don’t think 1 second makes a difference.
STEP 5: VERIFY THE SOLUTION
Level 1 Exemplar
Does not reflect on any of the three possible sources of error.
Participant #84

5. (a) Does your answer make sense? Why or why not?
   Yes

(b) Does your solution agree with the evidence? Why or why not?
   Yes
STEP 5: VERIFY THE SOLUTION

Level 2 Exemplar
Reflects on one of the three possible sources of error (i.e., mathematics used in their answer: I averaged it out).
Participant #39

5. (a) Does your answer make sense? Why or why not?
   Yes, because I averaged it out to see which is which

   (b) Does your solution agree with the evidence? Why or why not?
   Yes, because if you average it out it would be right
There is some additional text here that is not clearly legible.
STEP 5: VERIFY THE SOLUTION
Level 3 Exemplar
Reflects on two of the three possible sources of error (i.e., the experiment: we did it nine times; how their conclusion was drawn: since 53.8 is closest to 49 and 38.4 is closest to 40 the tests make sense).
Participant #44

5. (a) Does your answer make sense? Why or why not?

Yes, because in her text book it said that sodium chloride dissolves in 49.5 and potassium iodide in 40.5 and since 53.8 is closest to 49 and 38.4 is closest to 40 the tests make sense.

(b) Does your solution agree with the evidence? Why or why not?

Yes, just like I said in 5a it agrees with the evidence and is a fair test because we did it nine times.
STEP 5: VERIFY THE SOLUTION
Level 4 Exemplar
Reflects on at least three possible sources of error (i.e., the experiment: the test was performed several times; mathematics used in their answer: an average was done for both containers; how their conclusion was drawn: it was then compared to the textbook information and containers 1 and 2 were labeled).
Participant #30

5. (a) Does your answer make sense? Why or why not?
   My answer does make sense because the test was performed several times and an average was done for both containers. It was then compared to the textbook information and containers 1 and 2 were labeled.

(b) Does your solution agree with the evidence? Why or why not?
   My solution agrees with the evidence because the test on both containers was performed ten times and an average was done.
null
Problem-Solving Strategies Introductory Lesson

STEP 1: DEFINE THE PROBLEM
Ask: What is the root of the problem?
Think about what you are trying to solve.

STEP 2: DEVELOP A PLAN
Ask: What do I need to know? How am I going to find this information?
Real problems are complicated. Making a plan helps us understand the main problem and its smaller parts. Looking at one small question at a time helps us focus and gather all the information needed to solve the problem. Think about different ways to get the information you need to solve the problem (e.g., experimentation, other students, teachers, research, etc.).

STEP 3: GATHER INFORMATION
Ask: How am I going to record this information?
Use different resources to gather information. Think about the best way to record the information (e.g., graph, point-form notes, etc.).

STEP 4: SORT-LINK-SOLVE
Ask: How does this information help solve the problem?
Think about sorting the answers to the smaller questions, linking the information, and finding a solution for the main problem. Remember that some information will be useful and that some will not be needed to solve the problem.

STEP 5: VERIFY THE SOLUTION
Ask: Does the answer make sense? Can I explain my solution? Does my solution agree with the evidence?
Think about the problem and decide whether the solution is the best answer. Some problems will have more than one solution, but each should be supported by facts and logical reasoning. It is important to check over your answer and be willing to start over if necessary.
Appendix K

Problem-Solving Strategies Record Sheet

STEP 1: DEFINE THE PROBLEM
Ask: What is the root of the problem?

STEP 2: DEVELOP A PLAN
Ask: What do I need to know? How am I going to find this information?

STEP 3: GATHER INFORMATION
Ask: How am I going to record this information?

STEP 4: SORT-LINK-SOLVE
Ask: How does this information help solve the problem?

STEP 5: VERIFY THE SOLUTION
Ask: Does the answer make sense? Can I explain my solution? Does my solution agree with the evidence?
### The Big Mix Up Lab: Lesson Plan

**INTRODUCTION:**
- “The Big Mix Up” overhead
- Read through handout: “Big Mix Up” Student Instructions
- Students will be given a mixture (of 4 substances) and will be required to separate them:
  - sand
  - sawdust
  - salt
  - popcorn kernels
- Mass of initial mixture should equal total mass of all 3 separate components
- Students will work in pairs
- Read through problem-solving record sheet

**LAB:**
- Students will practice:
  - Evaporation: the change of state from a liquid to a gas
  - Sifting: the physical separation of different size particles using a sieve
  - Filtration: the separation of a solid from a liquid
  - Distillation: the process of separating a liquid by evaporation and condensation
  - Buoyancy: what floats and can be skimmed from surface?
  - Using electronic scale.
- Students will fill in problem-solving Steps 1, 2, and 3.
- Each student will be responsible for writing on their problem-solving strategies record sheet
- Students will carry out their plan and “solve” the problem.
- Students will fill in problem-solving Steps 4 and 5 before submitting all written steps.

**CONCLUSION:**
- Discuss the problem-solving strategies and the lab results

**ASSESSMENT:**
- Problem-solving steps
- Accuracy of mass conservation during separation

**ENRICHMENT:**
- Provide a second mixture for students to separate (i.e., pepper, metal filings, sand, sugar)
There has been a big mix up in Mr. Mixabut’s science lesson. Someone has mixed together bottles of sand, salt, sawdust, and popcorn kernels to make a real mess.

Now Mr. Mixabut is keeping the whole class in until the mixture has been separated back into the correct bottles.

Your job is to separate the four materials into their individual bottles.
Appendix M

The Big Mix Up Lab: Student Instructions

Your job is to separate the four substances into individual containers.

The mass of the four separated substances must add up to the mass of the mixture you started with.

You will be working in pairs.

You will have 4 days in total to do this lab.

Learn how to do each separation technique and use the electronic scale (go to each station).

Write out problem-solving steps 1, 2, and 3.

Use the separation techniques you just learned to separate the four substances.

Find out the mass of each separate substance. Place the substances into separate petri-dishes. Tape the petri dishes closed and write your name, class, and the mass of the substance on the tape.

Write out problem-solving steps 4 and 5.

Hand in the written problem-solving steps and the petri-dishes to Miss Skinner when you are finished.

HINTS:
WHAT METHOD WILL YOU USE TO SEPARATE EACH SUBSTANCE?
IN WHAT ORDER WILL YOU DO THE SEPARATIONS?
Appendix N

The Big Mix Up Lab:
Possible Answers to Problem-Solving Strategies Record

STEP 1: DEFINE THE PROBLEM
Ask: In your own words, what is the problem?
- How can I separate the mixture into its separate substances, without reducing the total mass?

STEP 2: DEVELOP A PLAN
Ask: What do I need to know?
- PART 1:
  - How to use an electronic scale.
  - Separation techniques:
    - Evaporation: the change of state from a liquid to a gas
    - Sifting: the physical separation of different size particles using a sieve
    - Filtration: the separation of a solid from a liquid
    - Distillation: the process of separating a liquid by evaporation and condensation
    - Buoyancy: what floats and can be skimmed from surface?
- PART 2:
  - Order of separation:
    1. strainer - popcorn kernels (1 and 2 might be reversed)
    2. add water - sawdust/pepper floats and can be skimmed out
    3. filter - sand will stay on filter paper; salt and water will fall through
    4. evaporate - salt from water

STEP 3: GATHER INFORMATION
Ask: How am I going to gather this information?
- PART 1:
  - I am going to have to experiment with the separation techniques.
- PART 2:
  - I am going to have to determine the order of separation for each component of the mixture and measure mass of each substance after it is separated using the electronic scale.

How am I going to record this information?
- PART 1:
  - Use a chart and point form notes to record information.

<table>
<thead>
<tr>
<th>Station</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Electronic Scale</td>
<td></td>
</tr>
<tr>
<td>2. Filtration</td>
<td></td>
</tr>
<tr>
<td>3. Evaporation</td>
<td></td>
</tr>
<tr>
<td>4. Mechanical Separation</td>
<td></td>
</tr>
<tr>
<td>5. Solubility and Buoyancy</td>
<td></td>
</tr>
</tbody>
</table>
• **PART 2:**
  • Use a chart to record mass of each substance after it is separated.

<table>
<thead>
<tr>
<th>Substance</th>
<th>Mass (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. popcorn kernels</td>
<td></td>
</tr>
<tr>
<td>2. sawdust/pepper</td>
<td></td>
</tr>
<tr>
<td>3. sand</td>
<td></td>
</tr>
<tr>
<td>4. salt</td>
<td></td>
</tr>
</tbody>
</table>

**STEP 4: SORT-LINK-SOLVE**

**Ask:** Solve the problem (show your work).

• **PART 1:**
  • We learned how to use an electronic scale.
  • We learned several separation techniques:
    • Evaporation: the change of state from a liquid to a gas
    • Sifting: the physical separation of different size particles using a sieve
    • Filtration: the separation of a solid from a liquid
    • Distillation: the process of separating a liquid by evaporation and condensation
    • Buoyancy: what floats and can be skimmed from surface?

• **PART 2:**
  • We separated the mixtures into their separate substances by the above means.
  • The total mass of the original mixture was \( x \) grams.
  • The total mass of the separate components is \( x + x + x + x = 5x \) grams.
  • Therefore the total mass has not changed/has changed slightly/is very different.

**STEP 5: VERIFY THE SOLUTION**

**Ask:** Does the answer make sense? Why or why not?

• Yes, the answer makes sense because the separate components (almost) were added up to get the total mass.
• The results could have been improved by greater accuracy in each technique. Especially....

**Does my solution agree with the evidence? Why or why not?**

• Yes, the answer agrees with the evidence because the separate components (almost) added up to the total mass.
• No, because the separate components did not add up to the total mass of the original mixture.
Appendix O

Science Sleuths: Lesson Plan
(Grade 7 Pure Substances & Mixtures)

COMPUTER LAB SET-UP:

- Students will be assigned a computer and a partner.
- Go through 2 handouts with all students:
  - Science Sleuths instructions
  - Problem-solving steps
  - Students will receive headphones and CDs

STUDENTS WILL:

- Use Science Sleuths program
- Write out problem-solving steps on record sheet
Appendix P

Science Sleuths: Student Instructions

Your job is to identify the substance on the beach and decide if it is harmful enough to close down the beach.

You will be working in pairs.

You will have 4 days in total to solve this problem.

If you finish solving the problem before that time, you will repeat the problem-solving process at the next difficulty level.

To get started:
Insert disc into CD-ROM drive
Click on START → Programs → CD-Rom Titles → Science Sleuths
BE PATIENT AS PROGRAM IS LOADED!
Click on red light on door
Type one partner’s first and last name → OK
Wait and listen to guide

Do the tutorial session on the computer:
Click on TV screen above filing cabinets; “Tutorial” will light up on screen
Listen and watch tutorial carefully

Learn about the HELP function:
Choose HELP in top tool bar; then CONTENTS
Read about the different topics

Learn about the program:
Click on The Blob mystery and Drawer Level 1
Click on the Sleuth Lab door
Explore how to use the different tools available
Save game: ask Miss Skinner for a disk; click on FILE; SAVE GAME; select A-DRIVE; save as your name + g (for game)
For example, I would save as “skinnerg”

Learn about the notebook
Copy and paste items into notebook
Save notebook: ask Miss Skinner for a disk; click on FILE; SAVE NOTEBOOK TO FILE; select A-DRIVE; save as your name
Teach them to print notebook: click on FILE; PRINT NOTEBOOK
WHEN YOU THINK THAT YOU KNOW SCIENCE SLEUTHS WELL ENOUGH: Write out problem-solving steps 1, 2, and 3.
To start playing the game
Click on Blob mystery and Drawer level 1
Click on the Lab door
Use the tools and notebook to solve the problem
Save game and notebook

After you have solved the problem print out your notebook.

Write out problem-solving steps 4 and 5.

Hand in the written problem-solving steps, the printed notebook and the disk to Miss Skinner when you are finished.

HINTS:
THINK ABOUT WHAT TOOLS YOU HAVE AVAILABLE TO USE!
REMEMBER TO SAVE ALL OF YOUR OBSERVATIONS IN THE NOTEBOOK!
Appendix Q

Science Sleuths:
Possible Answers to Problem-Solving Strategies Record

STEP 1: DEFINE THE PROBLEM
Ask: In your own words, what is the problem?
• What is the unknown blob on the beach? Is it harmful enough to need to close down the beach?

STEP 2: DEVELOP A PLAN
Ask: What do I need to know?
• Refer to teacher's manual

STEP 3: GATHER INFORMATION
Ask: How am I going to gather this information?
• Use the virtual tools to gather information

How am I going to record this information?
• Point-form in electronic notebook

STEP 4: SORT-LINK-SOLVE
Ask: Solve the problem (show your work).
• Refer to teacher's manual

STEP 5: VERIFY THE SOLUTION
Ask: Does the answer make sense? Why or why not?
• Yes it makes sense to keep the beach open because the blob was not toxic.

Does my solution agree with the evidence? Why or why not?
• Yes, the answer was derived from the evidence gathered in the notebook
Appendix R

Grade 8 Pilot Test: Problem-Solving Test Feedback

Name: 

Class: 

Date: 

What was the most difficult part of the test?

What was the easiest part of the test?

Do you think that changes should be made to the test? If so, what?

Any other comments about the test?
Appendix S

Sample Research Notes Used in Qualitative Analysis

- Consent form added to appendix
- Control = 7A, 7D and Experimental = 7B, 7C
- matched for high/low ability
- randomly determined which would receive treatment and which would not
- 5P02 Pre/Posttest Validation (March 29/00)
  - change each to a 2 part test
  - change table in pretest so that one trial is different from other 3
- 5V90 Thesis Proposal Presented and Discussed (April 3/00)
  - not a truly random sample
  - only dependent variable is posttest score; pretest (covariate), GALT, and treatment group are all
  - independent variables
  - change wording of 2nd research question?
- Posttest Validation Meeting (April 6/00)
  - many changes to pre and posttest resulted from this discussion
  - another important contribution I am making through this study is scoring of science problem
  - solving which should be emphasized in my report
- Meeting with Advisor (April 12/00)
  - statistical details added
- Proposal Hearing (April 17/00)
  - Several specific changes to the wording of certain sections of the written proposal were suggested and noted for later revisions of my report
  - Pretest Change:
    - Susan suggested that I underline the phrase “most consistently” in Part B because it is important for the students to pick it up if I am expecting them to perform an average calculation of bubble size.
    - No suggestions for posttest changes were made by the thesis committee.
- Pilot Study (April 18-20/00)
  - Overall comments:
    - Students in pilot study did not receive a full problem-solving lesson which resulted in their having some difficulty with the questions. Students in the main study will receive a full 1 hour lesson on problem-solving including a practical lab component before the pretest.
    - I also had to hint that the scientific method would be a big help in answering these questions. I do not plan to do this with the students in the main study because we will have already done the example investigation as part of the problem-solving strategies lesson, so they should know what I expect for the pre- and posttests without being explicitly told.
    - Only 19 students participated in the pilot (hoped for about 30 participants).
null
Most students gave earnest feedback both verbally and in writing on a feedback sheet I provided.

USE SOME OF THESE QUOTES IN FINAL PAPER

Part A was finished in 20-30 minutes and Part B was finished in 10-30 minutes. I have allotted 1 hour for students to do the pretest in the main study.

GALT:

Some pilot students had difficulty understanding the response sheet. I will address this in the main study by giving explicit instructions on how to use the response sheet.

Two students asked why there were more lines provided for Q20 and Q21 than they needed. I will forewarn the students in the main study that this may be the case for them as well.

First student was finished in 20 minutes and the last student took 35 minutes. I will allot an entire 1 hour period for the GALT in the main study and have a back-up lesson prepared in case they finish early.

Problem-solving strategies:

Step 3 changed from “How am I going to record this information?” to “How am I going to record this information (be specific)?”

Why was this change done? Because some students wrote “chart” and I was hoping to see chart headings as well. By adding the phrase “be specific” I hope to facilitate their doing so.

Problem-solving pretest:

Two forms were administered (raw data and table form). Since students in the pilot study seemed to respond equally to each, it was difficult to decide which to use in the main study. There were two students who stated that they were having difficulty reading the table. I decided to use the raw data form because I was interested to see if any students in the main study would convert the raw data into table form. In the pilot study no tables were made from the raw data.

The second reason I chose to use the raw data form is because of comments made by 5P02 classmates during my validation presentation. They seemed to agree that presenting the data in raw form was a better choice than presenting it in table form. They suggested that otherwise the test might be assessing how well students can “read a table” rather than solve a scientific problem.

The third reason for deciding to use the raw data form was that it was also discussed as a potential option, in place of a table, by the validation team. In fact, this is the reason why it was tested in the pilot study in the first place.

Overall, written feedback from students (4 questions) was positive. No consistent problem or suggested changes were stated in the responses. No further changes were warranted by the pilot study.

One common error in students’ responses was to use the time data in their solution even though it was not stated as a requirement. Should I specify that it is not a determining factor somewhere in the scenario?
- **Problem-solving posttest:**
  - I did not have time to do a problem-solving strategies lesson with these pilot study participants. To compensate, I displayed the problem-solving strategies on the overhead to serve as a “cheat sheet” for the students as they worked through the posttest.
  - The students’ lack of experience with the problem-solving strategies resulted in a few comments on the ambiguity of the problem-solving questions in the posttest. I believe that the grade 7 students will not encounter this same limitation because they will have had more exposure and, time to learn, the problem-solving strategies beforehand.
  - The time taken for Part A ranged from 15-30 minutes and the time taken for Part B ranged from 15-30 minutes as well. Therefore, I am allotting the whole 1 hour class period for students to do the posttest.
  - Only one student asked for a calculator (near the end of the test). This presents the question of whether or not I should provide them, or wait for them to ask because I don’t want to influence their responses. I will talk to my advisor about this and see what he suggests.
  - One student pointed out to me that I forgot to specify the temperature of the water used in Part B. I have made that correction to read 20°C. It was great to have these students so involved to pick up such a critical omission!
  - One student asked me “Can’t Mrs. Chung return the jars for other ones?” That is certainly one way to solve the problem!!
  - The table in Part B was difficult for some students to understand. I am wondering if it is because the temperature of the water was omitted in the raw data section, or if it goes beyond that.
  - I think that the posttest should be more difficult than the pretest because I do not want improvement in problem-solving to be a result of a learned effect. On the other hand, the tests should be similar enough to one another because they are supposed to measure the same skill.
  - Several students mentioned that the chart was too difficult. I decided to replace it with a sentence containing the information needed to solve the problem instead.
  - The scenario in Part A was partially rewritten to be more clear to the students.
- **Science Sleuths:**
  - Realized that I only had 7 Windows CDs and the rest were MAC version. Did not get to actually observe students using the program because there were so many “bugs” to work out (i.e., how students working in pairs would share headphones).
  - Decided to have a “seating plan” set up for the students participating in the main study because of the difficulty in “managing” the students in the computer lab. I want the focus to be on learning problem-solving skills, not on who is going to sit where, etc.
Appendix T

Graphic Representation of Qualitative Findings

153 UNITS OF INFORMATION

- began networking for participant school
- considered using Internet program called "virtual fly lab", but it was too advanced for intermediates
- reviewed science curriculum for intermediate students to find best connections to software programs

decided to teach middle school students and simultaneously conduct research with them

reviewed the GALT and its scoring

20 CATEGORIES

- participants
- ethics
- software
- methodology
- curriculum
- meeting with advisor

- literature review
- reflection
- pilot study
- coursework
- use of a model

- goal-setting
- proposal
- research question
- qualitative analysis

- future modifications
- faculty members
- variables
- quantitative analysis

- validation committee

6 THEMES

- practical considerations
- research interests
- research design
- research analysis
- development of the (problem-solving) tests
- scoring scheme development
Appendix U

Additional Data Tables
Table U1

Open Coding of Qualitative Data Collected During the Independent Study

<table>
<thead>
<tr>
<th>Units of Information</th>
<th>Categories</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-met with advisor to discuss my research interests</td>
<td>meeting with advisor</td>
<td>research interests</td>
</tr>
<tr>
<td>-read through recent research in science education</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-narrowed down my interest in science education to 3 possible areas of interest:</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>use of technology; new Ontario science curriculum; and the transition experienced by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intermediate students from elementary to high school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-continued reading recent research in science education</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-reflected on the use of technology in my own science education background</td>
<td>reflection</td>
<td>research interests</td>
</tr>
<tr>
<td>-continually set “next step” research goals for myself</td>
<td>goal-setting</td>
<td>research interests</td>
</tr>
<tr>
<td>-decided that my topic would be technology in science education</td>
<td>research question</td>
<td>research interests</td>
</tr>
<tr>
<td>-took class on current issues and trends in technology in science education</td>
<td>coursework</td>
<td>research interests</td>
</tr>
<tr>
<td>-reviewed Ministry licensed software called Science Sleuths which claims to promote</td>
<td>coursework</td>
<td>research interests</td>
</tr>
<tr>
<td>problem-solving skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-refined my topic to problem-solving and technology in science education</td>
<td>research question</td>
<td>research interests</td>
</tr>
</tbody>
</table>
-learned how to properly do an annotated bibliography
-learned how to access Brock University's library resources
-considered comparing students at different age levels
-realized difficulty in accessing several software for these different ages
-chose headings to organize articles for annotated bibliography around future literature review chapter
-began literature search on problem-solving
-read past projects and theses from IRC
-considered using both qualitative and quantitative assessments of problem solving
-decided to include logical reasoning ability levels instead of age groups (for further comparison)
-GALT was recommended to measure logical reasoning ability
-decided that participants should be intermediates (Grade 7-10); my teaching preference
-advised to use approximately sample size (30 students in about 4 classes) as a reasonable sample size
-decided to compare treatments (computer vs. other instruction)
<table>
<thead>
<tr>
<th>Units of Information</th>
<th>Categories</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-needed to decide on a software that promoted problem solving</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-called local school boards for recommendation on software</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-called Tangent Scientific company for software information (recommended by school board)</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-looked through STAO (Science Teachers Association of Ontario) publication for software recommendations</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-selected thesis committee members</td>
<td>proposal</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-continually added articles to annotated bibliography</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-read TVO web site for list of Ministry licensed software</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-considered the use of BioQUEST software because of its strong research base</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-considered using Think-Aloud-Protocol to qualitatively assess students’ problem-solving performance</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-advised to design my own quantitative measure of problem-solving to match it with the learning goals of my treatment groups</td>
<td>meeting with advisor</td>
<td>research design</td>
</tr>
<tr>
<td>-set goals for completing ethics form and data collection</td>
<td>goal-setting</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-evaluated pros and cons of each software</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-realized choice of software was limited by accessibility</td>
<td>software</td>
<td>practical considerations</td>
</tr>
<tr>
<td>Units of Information</td>
<td>Categories</td>
<td>Themes</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>-realized that BioQUEST would be most appropriate for high school students</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>(Not for middle school)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-submitted literature review and reflective journal to advisor for Independent study credit</td>
<td>coursework</td>
<td>research interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table U2

Open Coding of Qualitative Data Collected During the Research Proposal

<table>
<thead>
<tr>
<th>Units of Information</th>
<th>Categories</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-began networking for participant school</td>
<td>participants</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-continued literature review</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-considered using Internet program called “virtual fly lab”, but it was too</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>advanced for intermediates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-considered using more than one software application</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-reviewed science curriculum for intermediate students to find best</td>
<td>curriculum</td>
<td>practical considerations</td>
</tr>
<tr>
<td>connections to software programs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-began to finalize details of methodology chapter</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-decided to teach middle school students and simultaneously conduct research with</td>
<td>participants</td>
<td>research design</td>
</tr>
<tr>
<td>them</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-requested ethics approval from school board</td>
<td>ethics</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-continued literature review</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-received permission to use GALT test from its author</td>
<td>ethics</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-reviewed the GALT and its scoring</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-realized limitation of teaching would prevent my use of think-aloud-protocol for</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>qualitative problem-solving assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-decided to have participants do a reflective journal (for qualitative data)</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-devised problem-solving checklists to quantitatively measure problem-solving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Units of Information</td>
<td>Categories</td>
<td>Themes</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>-wrote first draft of the literature review chapter</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-decided to use Science Sleuths because of accessibility, appropriateness, and acclaim</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-requested Science Sleuths CD-ROM's from school board IT team</td>
<td>software</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-read Science Sleuth's website for additional background information</td>
<td>software</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-found many of the software's acclaims were made without research to support them</td>
<td>software</td>
<td>research design</td>
</tr>
<tr>
<td>-continued to review literature</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-considered how to combine qualitative and quantitative components</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-continue to set goals and deadlines for myself</td>
<td>goal-setting</td>
<td>research interests</td>
</tr>
<tr>
<td>-experienced some difficulty in writing literature review because of magnitude of information reviewed</td>
<td>literature review</td>
<td>research interests</td>
</tr>
<tr>
<td>-began thinking about statistical analyses</td>
<td>quantitative analysis</td>
<td>research analysis</td>
</tr>
<tr>
<td>-decided to use qualitative data to supplement quantitative findings, only (i.e., no separate qualitative analysis)</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-submitted first complete draft of proposal to advisor</td>
<td>proposal</td>
<td>practical considerations</td>
</tr>
<tr>
<td>Units of Information</td>
<td>Categories</td>
<td>Themes</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>-decided that Grade 7 “Pure Substances and Mixtures” unit was most compatible with Science Sleuths and would teach that concurrent to the study</td>
<td>curriculum</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-revised original research design with help from my advisor</td>
<td>meeting with advisor</td>
<td>research design</td>
</tr>
<tr>
<td>-revised original research questions to fit revised design</td>
<td>research question</td>
<td>research design</td>
</tr>
<tr>
<td>-advised to use a problem-solving pretest and posttest to improve reliability of results</td>
<td>meeting with advisor</td>
<td>research design</td>
</tr>
<tr>
<td>-realized need to change checklists into more sophisticated problem-solving tests</td>
<td>meeting with advisor</td>
<td>research design</td>
</tr>
<tr>
<td>-revised proposal and resubmitted it to advisor</td>
<td>proposal</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-set proposal hearing date with committee</td>
<td>proposal</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-prepared ethics approval form for Brock University</td>
<td>ethics</td>
<td>practical considerations</td>
</tr>
</tbody>
</table>
Table U3

Open Coding of Qualitative Data Collected During the Thesis

<table>
<thead>
<tr>
<th>Units of Information</th>
<th>Categories</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-finished parent/student consent forms</td>
<td>ethics</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-paired classes together (matched for high/low ability)</td>
<td>participants</td>
<td>research design</td>
</tr>
<tr>
<td>-selected treatment and control groups</td>
<td>methodology</td>
<td>research design</td>
</tr>
<tr>
<td>-advised by 5P02 class to change tests to 2 parts each</td>
<td>coursework</td>
<td>development of tests</td>
</tr>
<tr>
<td>-advised by 5P02 class to give raw data (not in table format) in test scenarios</td>
<td>coursework</td>
<td>development of tests</td>
</tr>
<tr>
<td>-advised by 5P02 class to use a rubric for scoring scheme</td>
<td>coursework</td>
<td>scoring scheme</td>
</tr>
<tr>
<td>-advised by 5V90 class I did not have a true random sample</td>
<td>coursework</td>
<td>research design</td>
</tr>
<tr>
<td>-increased knowledge of qualitative research methods through 5V90</td>
<td>coursework</td>
<td>research analysis</td>
</tr>
<tr>
<td>-advised to change wording of preliminary test scenarios</td>
<td>validation committee</td>
<td>development of tests</td>
</tr>
<tr>
<td>-advised to change wording of preliminary test questions</td>
<td>validation committee</td>
<td>development of tests</td>
</tr>
<tr>
<td>-realized significance of tests and scoring scheme development to my thesis</td>
<td>validation committee</td>
<td>research design</td>
</tr>
<tr>
<td>-discussed use of raw data versus table format in tests</td>
<td>validation committee</td>
<td>development of tests</td>
</tr>
<tr>
<td>-discussed statistical analyses with advisor</td>
<td>meeting with advisor</td>
<td>research analysis</td>
</tr>
<tr>
<td>-advised to review some wording in proposal report</td>
<td>proposal</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-advised to underline part of test question for emphasis</td>
<td>proposal</td>
<td>development of tests</td>
</tr>
<tr>
<td>Units of Information</td>
<td>Categories</td>
<td>Themes</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>-realized students will require a full lesson on problem solving</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-had 19 participants in pilot study (anticipated 30)</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-obtained student feedback on tests</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>-students took 20-30 minutes to complete Part A of the pretest</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-students took 10-30 minutes to complete Part B of the pretest</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-realized explicit instructions were needed to properly administer GALT</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-students took 20-35 minutes to complete GALT</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-realized need to develop secondary lessons for students not participating and for those who finish tasks early</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-restated wording of Step 3 more clearly</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>-administered 2 forms of pretest for comparison (raw data and table format)</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>-chose raw data format for tests</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>-students wrote mainly positive feedback on tests</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>-delivered a limited problem-solving lesson with pilot participants (because of time restraints)</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-students took 15-30 minutes to complete Part A of the posttest</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>-students took 15-30 minutes to complete Part B of the posttest</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>Units of Information</td>
<td>Categories</td>
<td>Themes</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>- made decision to permit the use of calculators during tests</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>- omitted critical data in posttest scenario used for pilot</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>- examined relationship between pretest and posttest</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>- rewrote posttest scenario for greater clarity</td>
<td>pilot study</td>
<td>development of tests</td>
</tr>
<tr>
<td>- discovered that we had been issued Windows and Macintosh CD-ROM’s (for Windows computers)</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>- requested more Windows version CD-ROM’s</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
<tr>
<td>- calculated number of functioning computers versus number of participants</td>
<td>pilot study</td>
<td>practical considerations</td>
</tr>
</tbody>
</table>
Appendix V

Figures
Figure VI. Scatterplot of relationship between pretest total score and GALT score.
Figure V2. Scatterplot of relationship between posttest total score and GALT score.
Figure V3. Scatterplot of relationship between improvement score and GALT score.
Estimated Marginal Means of Problem-Solving Improvement Scores

Figure V4. Relationship between problem-solving improvement, GALT, and instruction.
Figure V5. Scatterplot of relationship between pretest total score and posttest total score