THE IMPACT OF STUDENT THINKING JOURNALS AND GENERIC PROBLEM SOLVING SOFTWARE ON PROBLEM SOLVING PERFORMANCE AND TRANSFER OF PROBLEM SOLVING SKILLS

DISSERTATION

Presented to the Graduate Council of the University of North Texas in Partial Fulfillment of the Requirements For the Degree of

DOCTOR OF EDUCATION

By

Gary E. Sullivan, B.S., M.S.
Denton, Texas
August, 1993
THE IMPACT OF STUDENT THINKING JOURNALS AND GENERIC PROBLEM SOLVING SOFTWARE ON PROBLEM SOLVING PERFORMANCE AND TRANSFER OF PROBLEM SOLVING SKILLS

DISSEPTION

Presented to the Graduate Council of the University of North Texas in Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF EDUCATION

By

Gary E. Sullivan, B.S., M.S.
Denton, Texas
August, 1993

This study examined the effects of specially designed thinking journal activities that have been attributed with encouraging reflective thinking, on instruction using generic, or content-free problem solving software. Sixty-three fourth grade students participated in four instructional sessions using a software package called Moptown Hotel. Students completed separate posttests that measured (1) performance on problems of the same kind as those used in instruction, and (2) transfer of skills to other kinds of problems. Scores of students who wrote thinking journals prior to testing were compared with scores of students who did not.

Results indicate that students who wrote thinking journals performed the same as students who did not when tested on problems similar to those practiced in class. Tests in which students transferred their skills to word problems, however, produced significant differences. There was no significant difference between scores when averaged over all four weekly occasions. However, for the final
session alone, students who wrote thinking journals scored higher on tests of problem solving transfer than students who did not (p < .01).

The study also examined the relationship between the degree of metacognitive thought displayed in students' journal entries, and their measured problem solving ability. Results indicate that students who had higher average reflectivity scores also had higher average problem solving performance and transfer scores (p < .05). It was also noted that the significant relationship between reflectivity and scores of problem solving ability was only observed in male students.

It was concluded that under the right conditions, and for the right kinds of problems, thinking journal writing can help students understand their own thinking processes, resulting in improved problem solving behavior. The study also raises the question of whether there are differences between the ways that male and female students apply metacognitive awareness gained through journal writing experiences.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>vi</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>Purposes of the Study</td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Significance of the Study</td>
<td></td>
</tr>
<tr>
<td>Definition of Terms</td>
<td></td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Assumptions</td>
<td></td>
</tr>
<tr>
<td>II. SYNTHESIS OF RELATED LITERATURE</td>
<td>12</td>
</tr>
<tr>
<td>Problem Solving and Problem Solving Software</td>
<td></td>
</tr>
<tr>
<td>Transfer of Problem Solving Skills</td>
<td></td>
</tr>
<tr>
<td>Metacognition, Reflection, and Problem Solving</td>
<td></td>
</tr>
<tr>
<td>Thinking Journals and Problem Solving</td>
<td></td>
</tr>
<tr>
<td>Summary and Conclusions</td>
<td></td>
</tr>
<tr>
<td>III. DESIGN AND METHODOLOGY</td>
<td>36</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Research Design</td>
<td></td>
</tr>
<tr>
<td>Procedures for Collection of Data</td>
<td></td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
</tr>
<tr>
<td>Procedures for Analysis of Data</td>
<td></td>
</tr>
<tr>
<td>IV. RESULTS AND DISCUSSION</td>
<td>69</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
</tr>
</tbody>
</table>
Descriptive Statistics for Hypotheses 1 and 2
Testing Hypothesis 1
Testing Hypothesis 2
Testing Hypothesis 3
Testing Hypothesis 4
Reflectivity and Journal Writing
Overview of Significant Findings
Discussion

V. SUMMARY AND CONCLUSIONS................................. 119

Conclusions
Implications for Problem Solving Instruction
Suggestions for Further Research

APPENDIX ............................................................. 134

BIBLIOGRAPHY ....................................................... 174
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Strategies taught with Moptown Hotel</td>
<td>49</td>
</tr>
<tr>
<td>2. Results of Split-Half Analyses of Internal Consistency</td>
<td>57</td>
</tr>
<tr>
<td>3. Chi-squared Goodness-of-fit Test for Random Assignment</td>
<td>63</td>
</tr>
<tr>
<td>4. Cell Sizes and Variances for MANOVA Analyses</td>
<td>72</td>
</tr>
<tr>
<td>5. Performance and Grand Mean Performance Scores</td>
<td>74</td>
</tr>
<tr>
<td>6. Results of MANOVA for Mean Problem Solving Performance</td>
<td>75</td>
</tr>
<tr>
<td>7. Results of MANOVA for Performance, Weeks 1 - 5</td>
<td>77</td>
</tr>
<tr>
<td>8. Transfer and Grand Mean Transfer Scores</td>
<td>80</td>
</tr>
<tr>
<td>9. Results of MANOVA for Mean Problem Solving Transfer</td>
<td>81</td>
</tr>
<tr>
<td>10. Results of MANOVA for Transfer, Weeks 1 - 5</td>
<td>82</td>
</tr>
<tr>
<td>11. Transfer Score for Individual Classes, Week Five</td>
<td>85</td>
</tr>
<tr>
<td>12. Results of Multiple Linear Regression Analysis, Reflectivity with Performance</td>
<td>86</td>
</tr>
<tr>
<td>13. Reflectivity Scores, Weeks 1 - 5</td>
<td>89</td>
</tr>
<tr>
<td>14. Correlations, Reflectivity with Performance</td>
<td>90</td>
</tr>
<tr>
<td>15. Results of Multiple Linear Regression Analysis, Reflectivity with Transfer</td>
<td>91</td>
</tr>
<tr>
<td>16. Results of Multiple Linear Regression Analysis, Reflectivity with Transfer, Week 5</td>
<td>95</td>
</tr>
<tr>
<td>17. Correlations, Reflectivity with Transfer</td>
<td>97</td>
</tr>
<tr>
<td>18. Changes in the Length of Students' Journal Responses</td>
<td>98</td>
</tr>
<tr>
<td>19. Reflectivity Scores, Weeks 1 - 5</td>
<td>99</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Simple Randomized Subjects Design</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>Individual Treatments and Observations</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>The Problem for Moptown Hotel Lesson Two</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>Sample thinking journal questions</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>Example Student Journal Response</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>Problem Solving Performance for Females</td>
<td>74</td>
</tr>
<tr>
<td>7</td>
<td>Problem Solving Performance for Males</td>
<td>74</td>
</tr>
<tr>
<td>8</td>
<td>Interaction Plot, Week Five, Performance</td>
<td>78</td>
</tr>
<tr>
<td>9</td>
<td>Problem Solving Transfer for Females</td>
<td>80</td>
</tr>
<tr>
<td>10</td>
<td>Problem Solving Transfer for Males</td>
<td>80</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Computer software to teach problem solving has been popular almost as long as microcomputers have been used in the classroom. Many popular titles, like Gertrude's Secrets, Moptown Hotel, and Gears were developed in the early 1980's, and are still widely used today. These programs are examples of generic, or content-free problem solving software, meaning that they provide practice in general problem solving skills, rather than those specific to any particular subject. It is assumed that the skills learned with such software are useful in many kinds of problems encountered by students. However, researchers seldom report that skills learned with problem solving software help students perform better on other problems, or on tests of general problem solving ability. Applying skills to a kind of problem other than the one in which the skills were learned is called transfer. The skills learned in a computer simulation or game are worthless if they cannot be transferred to other academic or real-world problems.

These programs, however, have many redeeming qualities; they are highly motivating, and they provide opportunities
for students to practice "lower-level" thinking skills, like comparing, classifying, and discriminating. Such skills, however, are not problem solving skills.

Problem solving is a process by which a learner discovers the rules that can be used to solve a puzzling situation that the learner has not previously encountered (Gagné, 1970, p. 214; Tennyson, 1989). If students are not aware of the strategies and rules used for problem solving, however, like "breaking down a complex problem into smaller problems", they are likely to wander aimlessly, or to solve problems without knowing how they were solved.

There are two kinds of problem-solving strategies. One kind of strategy involves specific, multiple-step procedures that guide learners in planning and organizing problem-solving tasks, such as the well-known four-step process suggested by Polya. The other type of strategy is referred to as a heuristic by most authors. Heuristics are simple, "unordered" strategies (Marzano, et. al., 1988, p.46-47), usually consisting of general rules-of-thumb intended to help students focus on the problem and save time (Sherman, 1988). Teaching heuristic strategies is considered a very effective method for improving problem solving performance; its mention in the objectives often distinguishes problem-solving software from software for other kinds of thinking skills. Several authors have identified specific heuristic
strategies (Andre, 1986; Gagné et. al., 1988; Frederiksen, 1984; Cyert, 1980), such as:

1. Working backwards from the goal statement
2. Breaking down a complex problem into one or more simple problems
3. Hill climbing - choosing an action which will result in getting closer to the goal
4. Means-ends analysis - monitoring the difference between the goal and the current state, seeking actions by which the difference can be lessened
5. Withholding judgement
6. Seeing the total picture - avoiding getting lost in detail
7. Discussing the problem with others

Problem solving software packages do, in fact, require students to solve problems and to use such strategies. Why, then, do studies of problem solving software seldom report improvement in general problem solving ability?

One of the many answers to this question has been that students cannot effectively learn problem solving skills, and therefore cannot hope to apply them to other problems, unless they openly think about, communicate, and in the process, become aware of their own problem solving behavior. Students should reflect on their own problem solving, asking questions like "how well did I do?", "how well did my
strategies work?", and "what should I do differently next time" (Barell, 1991, p. 153). Students should also be asked to identify ways in which skills just learned could be applied to other familiar situations.

There are many ways to encourage students to articulate their thinking, including teacher questions, classroom dialogue, and peer collaboration. Many software packages suggest such activities. Recently, however, teachers of even very young children have engaged them in writing activities intended to improve self-awareness, self-esteem, and communication skills. This kind of writing is called journal writing, and special purpose thinking journals have been designed to support cognitive processes like problem solving. Perhaps, then, journal writing could make instruction with problem solving software more productive and more likely to result in the successful transfer of skills.

Statement of the Problem

The problem of this study is one of identifying teaching and learning strategies that contribute to development and transfer of problem solving abilities.

Purposes of the Study

This investigation will compare problem solving performance, and transfer of problem solving skills, of students who receive software instruction supported by
student thinking journal activities, with that of students who receive the same instruction, but who perform non-reflective writing activities. The study will also examine (1) the relationship between the degree of metacognitive and reflective thought displayed in students' journal entries, and their problem solving performance; and (2) the relationship between the degree of metacognitive and reflective thought displayed, and the transfer of learned problem solving skills to a related, but different problem solving task.

Rationale

Use of generic problem solving software motivates students and can improve students' problem solving performance, although such software is ineffective unless supported by other teaching or learning activities, like teacher modeling, classroom dialogue, or teacher-guided discovery learning. Also, children are more likely to learn problem solving skills, and more likely to transfer those skills, when they verbalize about their activities and their thinking, or reflect on their own problem solving behavior. Journal writing provides students with the opportunity to verbalize, and specially designed thinking journal activities can encourage reflective thinking. Therefore, it is possible that instruction using thinking journal activities with problem solving software experiences will
lead to better problem solving performance, and greater transfer of problem solving skills, compared with problem solving software used alone.

Hypotheses

1. Students who write thinking journal entries after each instructional session will score significantly higher on tests of problem solving performance than students who do not write thinking journals.

2. Students who write thinking journal entries after each instructional session will score significantly higher on tests of transfer of problem solving skills than students who do not write thinking journals.

3. There will be a significant positive relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving performance.

4. There will be a significant positive relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving transfer.

Significance of the Study

There are too few experimental studies of problem solving software. It is also important that studies investigate transfer as well as problem solving performance. If teachers are to use problem solving software, it is
important that research identify strategies for enhancing its effectiveness. This study is significant in that it will:

1. Determine whether journal writing is an effective strategy for improving instruction with problem solving software.

2. Investigate the relationships between student reflective thought and both problem solving performance and transfer of problem solving skills.

Definition of Terms

The following terms will have restricted meaning, and are thus defined for this study:

1. Problem solving performance

Problem solving performance will be defined as the students ability to solve problems of the same kind and content as those used for instruction and practice. It is operationally defined as student performance on written worksheets consisting of problems of the same kind and content as those present in the software.

2. Transfer of problem solving skills

Transfer of problem solving skills will be defined as the student's ability to solve problems of the same kind, but different content than those used for instruction and practice. The transfer problems used in this study are
considered examples of near transfer. *Transfer of problem solving skills* is operationally defined as student performance on written worksheets consisting of problems that are *isomorphs* of those present in the software. Problem isomorphs are sets of problems that share the same abstract structure, but different content; they are often used to establish the extent that learned skills transfer (Newman, Griffin, & Cole, 1989, p. 21).

3. Degree of metacognitive and reflective thought

Metacognition refers to the "higher order" control processes an individual uses to regulate thinking. However, some have described degrees of metacognitive thinking along a continuum, ranging from complete lack of awareness of decision making processes, through various degrees of awareness, to the point that individuals can reflect upon their own thinking, evaluating and comparing specific thinking processes (Barell, 1990, p.211). *Degree of metacognitive and reflective thought* is defined as the degree to which the problem solver has progressed along this continuum. *Degree of metacognitive and reflective thought* is operationally defined as a numeric rating from zero to three, based on an analysis of student writing in thinking journals. The levels are: 0 - Unaware; 1 - Aware / Mentioning; 2 - Descriptive; 3 - Reflective. This metric is described in Chapter 3, Design and Methodology.
4. Reflective comments

Following methods developed by Higgins, Flower, & Petraglia (1991), and extended for this study, reflective comments are defined as ones that include at least one of these features:

1. explicit evaluation of problem solving activities
2. explicit comparison or consideration of alternatives
3. explicit reasoning or justification for strategies
4. recognition of the strategic value or effectiveness of specific reflective thinking processes, e.g., evaluating, comparing, or justifying one's problem solving strategies.

Limitations

1. This study is intended as an exploratory study. A limitation of this study is that the sample was selected from the population of a local public school. Selection of participating classes was at the discretion of the school administration. There was no attempt to randomly select classes from a larger population, and hence, no claim of generalizability to a larger population is made.

2. The researcher for this study has no access to individual student records. Therefore, information regarding students' grades, standardized achievement scores, intelligence, or other data that would ordinarily be helpful for blocking assignment to treatment levels, or in analyzing results, was not available.
3. The Apple II computers used for the experiment were unreliable. At times, computers stopped working and students were either disrupted or forced to work in groups of threes, rather than in pairs. With three students at a computer, it was difficult for the researcher to monitor whether all students were participating. This condition was improved after the second week, so there were few disruptions during the last two weekly sessions.

4. The current study was limited by the fact that the total treatment was confined to four weekly sessions, ranging from one hour and twenty minutes to two hours each. Research indicates that teaching problem solving takes time, and that one should not expect students to demonstrate gains in problem solving skills after a few hours of instruction and practice. However, allocating long class periods or lengthy sequences of lessons for content-free problem solving instruction is problematic at the elementary level, since time must be subtracted from usual subject-area schedules, and content-free instruction cannot be easily assigned to any particular subject. Leah Melnik, who conducted a quasi-experimental study of content-free problem solving software, also encountered this difficulty, and states that teaching problem solving at the elementary level, although important, is considered secondary to teaching basic skills (1982, p. 37).
Assumptions

1. Since the software includes no provision for on-line measurement of problem solving performance, it is assumed that written worksheets measure the student's ability to solve equivalent software problems.

2. It is assumed that, as stated in the software's documentation, worksheets of related problems measure identical problem solving skills in a different context, and therefore constitute an assessment of transfer of problem solving skills. This assumption was supported by analyses of the problem sets prior to their use.
CHAPTER II

SYNTHESIS OF RELATED LITERATURE

Problem Solving and Problem Solving Software

1. Generic problem solving software

Software for teaching problem solving skills is classified as "Type II" instructional software because it emphasizes higher-order thinking skills and provides students with more control than drill-and-practice or tutorial software (Mandinich & Linn, 1986; Duffield, 1991; Craig, 1985). There are three kinds of software for teaching problem solving, (1) generic software designed to teach general problem solving skills, (2) software designed to teach specific problem solving skills within subject areas, and (3) software not designed to teach problem solving, but used to do so (Gore, 1988). Generic problem solving software is characterized by problems in the form of puzzles or simulations. It is also called content-free; it is not bound to any particular subject, and can therefore be applied to problem solving in more than one subject (Shaw, 1986). Generic problem solving software, the subject of this study, attempts to teach general problem solving skills that can be transferred to other kinds of problems.
2. Effectiveness of problem solving software

The literature on the use of computer software to teach general problem solving skills is sparse and inconclusive. It is often informal; there have been few experimental studies of computers and problem solving (Burton and Magliero, 1988; Yates and Moursund, 1988; Roblyer, 1989). Support for use of the Logo programming language for teaching problem solving is just as inconclusive (Hofmeister, 1989), despite the fact that there are more studies than for generic problem solving software.

Several studies have reported that problem-solving software instruction increases problem solving ability (Stein & Linn, 1985; Lieberman & Linn, 1991; Norton & Resta, 1986), although they often find no difference in gains between students receiving computer treatments and those receiving non-computer treatments.

An exemplary study of this nature, a quasi-experimental study by Melnik (1986), extended prior research that had observed significant transfer of skills from an instructional strategy game to problem solving ability measured by a standardized problem solving test. Melnik's study supported the assumption that generic problem solving software practice improves a child's general problem solving skills. It involved six classes of 5th and 6th grade students; each intact class was assigned to one of six treatments, in a two by three factorial design. The study
was limited, however, by lack of random assignment to treatments, although this was controlled by applying a non-equivalent control groups design (Campbell and Stanley design number 10), in which all students were pretested and post tested, and mean gains scores compared among groups.

Melnik assigned two independent variables, Method (problem solving worksheets, microcomputer problems solving software, and a control group); and Model ("Strategy Model" and "No Strategy Model"). The Method control group received unspecified "traditional" classroom instruction (p. 47). The Factory and Code Quest were used for the computer software treatment; these programs teach general problem solving and thinking skills such as working backward, determining sequence, inductive thinking, discrimination, and pattern identification (p. 49-50). Polya's four-step problem solving process was used as the Strategy Model; it involved direct instruction in problem solving strategies. Problem solving performance was measured using the Mathematics Application Subtest (intermediate 2 level) of the Stanford Achievement Test. The subtest involves general mathematics word problems; it was selected, in part, because the researcher expected that the kind of strategies taught using the Strategy Model would be useful for mathematics problems.

Melnik reported that students using problem-solving worksheets and students using computer software made
significantly greater gains in measured problem solving performance than students in the control group (p < .05). However, there was no significant difference between gains of students using problem-solving worksheets and students using computer software. Also, there was no significant difference between groups receiving Strategy Training and groups that did not. This study supports the use of generic problem solving software for teaching transferable problem solving skills. However, it does not support the teaching of general problem solving strategy models.

Researchers often attribute failure to find significant gains in problem solving skills to reasons other than ineffective computer software, such as studies not grounded in problem solving theory, subjects who are too young for the instruction provided, and inadequate quality, length, or intensity of treatment (Palumbo and Reed, 1991). For instance, there was no gain in problem solving ability in a study of Solutions Unlimited; researchers attributed it to the brief six-hour treatment (Langholz and Smaldino, 1989, p. 274). Sensitivity of measurement can also be problematic; students might acquire an intuitive understanding of concepts, but be unable to explain them or respond correctly in "dichotomous" testing of problem solving skills (Burton & Maglierio, 1988). The authors also state that computers themselves are often considered a treatment, with the selection of software and instructional
methods left to chance or not fully described. For example, Gilman & Brantley (1988) found no significant difference in achievement or problem solving skills between an experimental group exposed to computer software for a year, and a control group that did not use computers. The treatment involved the students' self-selection of a variety of unspecified software.

3. General theory and classroom research

Those who study problem solving software often ground their studies in the large body of general theory and research on problem solving. While early studies investigated the characteristics of good problem solvers, more recent research examines practical, school related situations (Lockhead, 1981, p. 69). Such classroom research often indicates positive results for teaching specific kinds of problem solving, particularly mathematics, but research on teaching general problem-solving strategies is a largely unexplored area of research (Picus, et al., 1983, p. 8, 10; Eylon & Linn, 1988, p. 280). Some report that problem solving strategies can be taught, but that there is no single best method for doing so (Picus, et al., 1983, p. 6; Suydam, 1982, p. 56; Silver & Thompson, 1984, p. 537). Even young children can understand general problem solving strategies (Barell, 1991, p. 155-156), however, students of different abilities might not benefit equally from
instruction in problem solving (Eylon & Linn, 1988, p. 283; Krutetskii, cited by Silver & Thompson, 1984, p. 536). At a minimum, we know enough to teach well-defined problems in relatively narrow problem domains (Frederiksen, 1984).

4. The importance of instructional support

Researchers often report that problem solving software is ineffective unless supported by classroom activities or instruction. According to Lieberman and Linn (1991), generic problem solving software does not teach students how to solve problems; it relies on the classroom teacher to provide the guidance and instruction needed. For example, students who received no problem solving instruction in a study of The King's Rule and Safari Search (Duffield, 1990), and students who were self-directed in the use of Rocky's Boots (Burbules and Reese, 1984), made little or no gains in the use of problem-solving strategies. Stein & Linn (1985) reported that even gifted students did not master Rocky's Boots during free exploration, but a brief instructional intervention had dramatic results. In particular, the literature repeatedly calls for explicit teaching of problem-solving concepts and skills, particularly the general applicability concepts that students often fail to see on their own (Picus, et al., 1983, p. 118; Simon, 1980). Even when general concepts are taught, however, students need detailed instruction in specific problem-
solving procedures in order to "fill the gap between the more general idea and the detailed procedure that needs to be constructed" (Eylon & Linn, 1988, p. 283).

Both non-computer problem solving literature and studies of problem solving software identify many factors that can improve success in learning problem solving skills, such as keeping the locus of control with the student rather than the software (Lieberman and Linn, 1991), and having students work at a level slightly higher than their current ability (Stonewater & Stonewater, 1984). In particular, researchers have repeatedly emphasized the critical importance of high motivation, self-confidence, persistence, and engagement (Picus, et al., 1983; Stonewater & Stonewater, 1984; Silver & Thompson, 1984; Barell, 1991), although the "role of affect" has been largely ignored by cognitive psychology (Meichenbaum, 1985).

5. Summary

The use of generic problem solving software is supported by a substantial base of general problem solving theory and research, by evidence that generic problem solving software can be effective, and recently, by classroom research that identifies the factors that influence learning. Although the literature is not conclusive, there is substantial evidence that when generic problem solving software is used with adequate instructional support, and when children are
motivated and self-confident, even elementary age children can gain general problem solving ability.

Transfer of Problem Solving Skills

1. Studies of problem solving transfer

Problem solving skills are transferred when they are applied in a problem domain other than the one in which the skills were learned. Transfer is difficult to attain, difficult to measure, and is seldom reported in the literature (Frederiksen, 1984; Yates & Moursund, 1988; Palumbo & Reed, 1991). If the new problem is unlike the original, the situation is called far transfer, and such transfer is more difficult to achieve (Silver & Thompson, 1984; Gagné et. al., 1988). Even near transfer is problematic, however. Students in a study by Malojkovich (1983, cited by Lieberman & Linn, 1991) learned to operate the logic gates in Rocky's Boots and to solve problems, but they did not transfer skills to similar logic software. Pea, Kurland, and Hawkins (1985) also failed to show transfer from a computer-based microworld that modeled a problem similar to a classroom planning task. However, Barell (1991) cites evidence from a student's thinking journal indicating that students can apply strategies learned in the classroom to other kinds of problems (p. 148, 230); Swan (1989) reported that fourth through sixth grade students transferred skills from a computer environment to
paper-and-pencil exercises; and the Melnik study already cited reported skills transferred from generic problem solving software instruction to a standardized test of mathematics problem solving skills (Melnik, 1986).

Some studies of problem solving software ignore the issue of transfer, focusing only on student performance within the original task. Others have used standardized tests to measure skills transfer. The Ross Test of Higher Cognitive Processes and the Watson-Glaser Critical Thinking Appraisal (Palumbo & Reed, 1991), the Purdue Elementary Problem Solving Inventory (Langholz & Smaldino, 1989; Norton & Resta, 1986), the Scott-Foresman Developing Cognitive Abilities Test, and the Mathematics Applications Subtest of the Stanford Achievement Test (Melnik, 1986) have been used in studies of problem solving software.

Other researchers have measured transfer by testing performance on a problem solving task similar to the one used in instruction. Problem isomorphs are sets of problems that share abstract structure, but different content. At a structural level, two isomorphic problems are essentially the same, in that the specific set of strategies and rules used to solve one problem can also be used in solving the other. Problem isomorphs are often used to establish the extent that learned skills transfer (Newman, Griffin, & Cole, 1989, p. 21). For instance, Duffield (1990) had students learn problem solving skills with one software
Students failed to demonstrate transfer, but the researcher had not established that the two programs were isomorphic. Researchers used a problem isomorph successfully in another study of a computerized Tower-of-Hanoi problem (Lehrer & Randle, 1987). Using isomorphs to test for transfer, however, requires careful analysis of both tasks, and this is seldom done (Belmont & Butterfield, 1977).

2. Methods for improving transfer of skills

The literature identifies many methods that encourage transfer of problem solving skills, such as direct instruction (rather than learning by discovery); practice in a wide range of examples and contexts (Greeno, 1980; Martin & Hearne, 1990); guided discovery; and explicitly pointing out the similarities among different contexts (Duffield, 1991, p. 54-55). Other "tools that foster transfer" focus on student awareness; they include (1) a process of setting goals, implementing strategies, and reflecting on progress, (2) posing questions, and (3) having students keep thinking journals (Barell, 1992, p. 2). These principles are evident in the teacher's guides for computer games like Gertrude's Puzzles and Moptown Hotel (Perl, 1985A, 1985B), which suggest guided discovery, followed by independent practice and related non-computer examples to allow strategies to transfer.
3. Summary

Transfer of problem solving skills is an important, yet problematic dependent variable. The literature suggests several methods for measuring transfer, but it does not document greater success for any particular method. Problem isomorphs might be more convenient, however, since some software packages include sets of non-computer problems suitable for use as post-tests.

A study of problem solving software should employ teaching strategies supported by recent research. Teaching for metacognitive awareness, discussed in the next section, is a strategy considered to improve problem solving ability and increase the probability of general transfer.

Metacognition, Reflection, and Problem Solving
1. Metacognition

Cognition refers to the thinking processes associated with some goal; metacognition refers to the "higher order" control processes an individual uses to regulate thinking. Metacognition is believed to play an important role in many cognitive tasks, including problem solving (Flavell, 1981, p. 37-38; Flavell, 1979, p. 906). Metacognitive strategies include relating new information to former, selecting thinking strategies, and planning, monitoring, and evaluating thinking processes (Blakey & Spence, 1990; Sternberg, 1985). Metacognition also involves "knowledge
of, awareness of, and control of the feelings that accompany certain situations", helping students recognize, understand, and control frustration, panic, or negative attitudes when solving problems (Barell, 1991, p. 207-208).

The literature suggests that teaching metacognitive skills is associated with increased learning and more effective problem solving (Lockhead, 1981, p. 68; Sherman, 1988; Blakey & Spence, 1990; Lieberman & Linn, 1991, p. 375). However, metacognitive experiences can only occur when the cognitive load of the task permits the additional working memory load required of metacognitive activity; overt attention to a metacognitive skill becomes a second learning task for the student (Flavell, 1981, p. 49; Beyer, 1988, p.230).

Schoenfeld (1989) reported that the problem solving activities of college students taught metacognition and self-regulation more closely resembled those of experts than the same students before instruction (p. 94-99). While students worked on geometry problems for which they did not possess a ready solution, and which were presented out of context, they were asked such questions as "What exactly are you doing?", "Why are you doing it?", and "How does it help you?". At the beginning of the study, students would read a problem, choose an approach, and work with it for the entire 20 minute duration, regardless of its success. By the end of the study, students' problem solving activities more
closely resembled those of experts, who tend to generate a large number of trial solutions, discarding inappropriate ones quickly.

Some state that teaching for metacognition is essential if thinking skills are expected to transfer. Belmont & Butterfield reviewed 114 studies on the use of cognitive instruction, none of which involved training in "executive decision making" (1977, p. 467). While the authors found no reports of generalizable results, a subsequent review by Belmont, Butterfield, and Ferretti reported six studies that produced transfer, all of which included instruction in metacognitive skills (Meichenbaum, 1985, p. 413).

2. Reflection

Metacognition can also include reflecting on one's thinking, before, during, and after problem solving. Reflection is a particular kind of metacognitive activity, involving evaluative thinking about problem solving or other intellectual processes. It is considered by some to be the highest form of metacognitive thought (Higgins, et al., 1991, p. 1; Barell, 1991, p. 209-212).

Student reflection is an effective strategy for improving thinking skills, metacognitive skills, and problem solving ability. "Reflecting on progress" was previously mentioned as a tool that fosters transfer, and Callahan & Garofalo list "assignments that require reflection and
analysis" among methods that help develop metacognitive awareness (1987, p. 22-23). Higgins, et al. (1991) state that there is consensus that "reflection plays an integral part in independent problem-solving and self-regulated learning. However, reflection is not appropriate in all problem situations; creative activities can be impeded by excessive or premature critical scrutiny (p. 1-2).

Reflective activities have also improved learning with instructional problem solving software. Students using a computer game called Zapworld were unable to describe their plans, and when they did solve a problem, they had no clear memory of the solution path due to frequent incorrect or backward moves. Zapworld was changed to require that the subjects record how they solved each problem. Researchers concluded that reflection of problem solving resulted in clearer development of problem-solving procedures. (Levin, et al., 1986). It has been suggested that problem solving software might be more effective if it were modified to include reflective activities (Yates and Moursund, 1988; Levin, et al., 1986).

There is disagreement regarding young children's ability to engage in reflective activities. The authors of Dimension of Thinking (Marzano, et al., 1988, p.131) state that objectives involving reflection need not be reserved for intermediate students, and Barell (1991) cites evidence that some youngsters, even first graders, can reflect on
their own thinking. Beyer disagrees, stating that until children move into the stage of formal abstract reasoning, they are not ready for direct instruction in metacognition (1988, p. 230). Research suggests that young children are quite limited in their metacognitive knowledge, do relatively little monitoring of their own cognition, and often fail to reflect on their own problem solving strategies (Flavell, 1979, p. 906; Eylon & Linn, 1988, p. 281).

3. Summary

Children are more likely to learn and transfer problem solving skills when they are taught metacognitive strategies and encouraged to reflect on their own problem solving behavior. Lack of metacognitive awareness is a possible explanation for experiments that do not report positive effects of instruction, especially when strategies for encouraging such awareness are not part of the treatment. While there is concern that metacognition is difficult for very young children, it is possible that activities that require reflection of problem solving processes will improve the effectiveness of problem solving software, at least for older students.
Thinking Journals and Problem Solving

1. Language and Thinking

Can language assist in the construction of one's knowledge? Staton (1984) believes that it does, and describes two different models of the relationship between language and knowledge. Staton rejects a "picture metaphor" of language that holds that thought must always precede utterance, such that language cannot reveal or generate any information not already known to the speaker. However, an "enlarged definition" of language states that knowledge is not merely a picture of reality, but is actively constructed through interaction of the mind with the external world. Thus, language serves to mediate a constructive cognitive process. This is supported by the theories of Piaget, who believed that at higher, more complex levels of thinking, language is a necessary condition "for the construction of logical operations" (Staton, 1984, p. 56-60).

Several studies have reported that children demonstrate higher levels of thinking and problem solving when they verbalize their activities and their thinking, either orally or in writing. Analysis of the language of children in a peer cooperative learning environment for three communications games revealed greater frequency of use of complex linguistic forms and questions than occurred in other classroom activities (Dickson, 1982, p. 145). In a
study of the effects of children's verbalization of spatial relations on their ability to solve a spatial task, children who verbalized spatial relations outperformed children who simply made picture choices (Olson & Ives, 1983, cited in Dickson, 1985). Also, mathematics students who explained solutions in class were better problem solvers than those who did not (Metwali, 1979, cited in Silver & Thompson, 1984).

Literature on classroom dialogue and social learning also indicates the importance of verbalization. Asking questions was more likely than direct instruction to foster cognitive change in several studies of problem solving and problem solving software (McClane, Marth, & Hartman, 1991; Norton & Resta, 1986; Shalaway, 1990; Pogrow, 1990B). Peer communication and collaboration is also widely associated with effective problem solving behavior (Silver & Thompson, 1984; Duffield, 1991; Levin, et al. 1986; Tennyson, 1989).

2. Writing and Thought

It is often assumed that when ideas are written down, they take on greater clarity, but until recently, there has been little empirical evidence to support this assumption (Lockhead, 1981; Staton, 1984, p. 67). According to Staton, writing creates opportunities for reflection and "articulation of feelings and thoughts in symbolic form" (p. 489). The author states that writing is generative - the
student must decide what to say, and how much, and one written thought can lead to another, extending and modifying ideas. Writing also creates a permanent, visual object, that can be reviewed by the writer. Writing allows us to manipulate and change our own cognitive processes.

Vygotsky believed that written language is closer to thought than it is to oral language because written language must be more detailed and specific. Oral language is often effective, even when incomplete or less specific. Luria, a colleague of Vygotsky, believed that writing leads to repeated analysis, which in turn leads to increases in metacognitive awareness and reflectivity (Staton, 1984, 68-69).

3. Journal Writing

While there are many kinds of student journals, the literature on three types - dialogue journals, learning logs, and thinking journals - is of interest in the study of thinking and problem solving skills. A dialogue journal is a continual, written conversation between two persons, regarding various matters of individual or mutual interest (Bode, 1989, p. 568-569). Learning logs are more focused than dialogue journals, directing the student's attention to specific academic content. Learning logs have also been used to encourage reflective thought, or "learning how to learn" (Kuhrt & Farris, 1990, p. 437). The thinking
journal, the subject of this study, can be thought of as a specialized kind of learning log, one directed specifically toward thinking and problem-solving skills, rather than academic content.

For each of these types of journals, journal writing is intended to be private, self-expressive, and reflective. It helps the student recognize self-opportunity, and allows the student to complain and question, always without fear of punitive results (Craig, 1983). Journals also allow for dynamic exchange of information, rather than simple transmission of static knowledge (Bode, 1989, p. 570). However, journal writing does not guarantee active participation; it does, however, make it more difficult for students to be passive (Kuhrt & Farris, 1990, p. 437).

A major purpose for dialogue journals is to help students learn what Jana Staton calls "practical reasoning" skills through interaction with adults, who provide scaffolding for the child's thinking. Practical reasoning refers to "a reflective process of considering one's actions and their causes or outcomes" (Staton, 1984, p. 30), an essential element of thinking journals as well. Staton's study identified several ways in which journal writing addresses the needs of students: (1) they provide students with an opportunity to discuss problems with someone (although not necessarily to solve them); (2) they help students understand that they have the power to change
outcomes; (3) they encourage students to accept personal responsibility; and (4) they help students to understand that they have choices. The study found that upper elementary students will engage in a personal dialogue, and will bring up significant issues (p. 517).

Journal writing empowers both students and teachers. Student journals allow teachers to assess student understanding (Kuhrt & Farris, p. 437), and they help them make subtle adjustments in their teaching (Sanders, 1985). Journals can be used successfully with even very young children; Kintish (1986) outlined a developmental series of journaling techniques for pre-kindergarten through fourth grade. The author reported that students in a four-year study exhibited a sense of enthusiasm and accomplishment for writing, displaying richer written vocabularies and experiencing an easier transition into middle school writing. However, journal writing is neither a quick fix, nor an easily adapted gimmick (Pradl, 1985, p. 8). Journal writing is very time consuming; it requires emotional commitment, and it requires careful design and consideration of the many ways in which student writing can be directed. - Research indicates that active, rather than passive learning is an important factor in student achievement.
4. Thinking Journals

Journal writing provides students with opportunities to verbalize, and specially designed thinking journals can encourage metacognitive and reflective thinking. In the past, journal writing has been associated with improved academic performance in content areas, improved thinking skills, increased writing and composition skills, and improved reading comprehension (Robinson-Armstrong, 1991; Edwards, 1992, p. 313). Barell (1991), however, describes a special kind of journal that encourages metacognitive awareness and reflection of problem-solving activities. Students keep thinking journals, in which they identify problems and record how they solved them. They also record what they learn about their own thinking abilities and attitudes. Barell describes several varieties of thinking journals, including a problem solving journal for awareness and transfer (p. 225). Problem solving journals were used to foster goal setting for sixth grade students and in-depth understanding of second-grade math problems, as well as for discussing general problem-solving strategies, and for visualizing problems (p. 72-76, 145, 225, 228).

In a study of journal writing and high school students' geometry and thinking skills, most students scored higher marks while engaging in journal writing than before, although students in the lowest ability group did not
improve (Linn, 1987). Qualitative analysis of journal entries indicated that students' metacognitive ability was enhanced through journal writing activities (p. 66). However, Robinson-Armstrong (1991) reports that in-class journal writing diminishes class time for other activities, and students sometimes are pressured to write when they are unprepared or not in the correct frame of mind.

5. Summary

Children often demonstrate higher levels of thinking and problem solving when they have opportunities to verbalize their activities and their thinking, since language is an important, perhaps essential element in the construction of knowledge. Written language is perhaps more powerful than spoken language, helping clarify and modify one's thoughts, creating opportunities for critical reflection, and supporting thinking that is too complex to be thought of otherwise. Student journal writing is a method through which teachers can support and develop students' reasoning skills, help students become more self-expressive and reflective, and provide a welcome opportunity to discuss with an adult matters important to the child. Used with skill, journal writing can be empowering for both student and teacher alike. Of central importance to this study, researchers and teachers have developed specially focused thinking journals that encourage the metacognitive and
reflective thinking considered an important factor in developing effective problem solving skills.

Summary and Conclusions

Problem solving software requires supportive activities; thinking journals provide the kinds of support associated with effective problem solving instruction. The literature suggests guided discovery in which teachers react to student problem solving activities, and dialogue journals are an efficient method for such individualized communication. Journaling is also associated with the high motivation and positive student attitudes required for effective problem solving behavior. Most importantly, student reflection improves problem solving ability, and thinking journals have been designed to encourage reflection of problem solving activity.

Problem solving software is more effective if used in an environment in which metacognition is encouraged, and might be more effective if activities designed to encourage and support reflectivity are added. Journal writing provides such an opportunity, and the computer software itself can encourage such metacognitive and reflective thought (Yates and Moursund, 1988).

The literature suggests, then, that instruction combining thinking journals and problem solving software might lead to better problem solving performance, and
greater transfer of problem solving skills, compared with problem solving software used alone.
CHAPTER III

DESIGN AND METHODOLOGY

Subjects

Sixty-three (63) fourth-grade students, from three fourth-grade classes, served as subjects in this experiment. Fourth-grade classes were requested because of the instructional level of the computer software selected for study. Classes were selected by the principal of a public elementary school, from a population of five fourth-grade classes, based on schedule constraints (the school follows two academic calendars, a "traditional" calendar, and a "year round" calendar; the three classes selected follow the traditional calendar). The selected classes included two classes of twenty-two, and one class of twenty-four. Sixty-three students participated in the study, such that class sizes for instruction were twenty, nineteen, and twenty-four, although three students withdrew from the study after the first week. Each class was scheduled for one weekly instructional session, during the three-week period of March 22 through April 9, 1993. Each class participated in one final session, scheduled the week of April 19. One class participated in a session the week of April 12, although data were not collected for this occasion. All classes met
at the same time of day, from 8:00 A.M., to approximately 9:45 A.M.

The subject matter for this study was new to the students involved. None of the subjects had been exposed to the problem solving software selected for this study, and problem solving software instruction is not a part of their usual curriculum. Classroom teachers reported that the students have substantial experience in response journal writing; the school's curriculum includes journal writing activities that begin in first grade. However, students had no prior experience with directed thinking journal activities.

Classroom instruction was provided by the researcher. Each post-test was administered by a classroom teacher for the first half of the experiment. After analysis of the first two weekly sets of data, there were concerns regarding reliability of measurement. To minimize the possibility of threats to reliability of measurement due to inconsistent administration among classes, subsequent post-testing was administered by the researcher, with assistance from the classroom teacher. This also allowed descriptive data to be collected regarding student behavior during testing. Measured reliability of instruments improved, although there were many factors that contributed to this; details are reported in a subsequent section.
Research Design

1. Hypotheses 1 and 2

This experiment examined the relationship between two levels of an independent variable - Method ($X_A$ and $X_B$), and two dependent variables, problem solving performance ($\text{Performance, } Y_1$), and transfer of problem solving skills ($\text{Transfer, } Y_2$). Students assigned to Method A ($X_A$, Journals) participated in instruction and computer practice, followed by a thinking journals exercise, while those assigned to Method B ($X_B$, Control) participated in the same instruction and computer practice, but were given an alternate, non-reflective writing exercise.

The purpose is similar to that of a study by Linn (1987), who studied mathematical problem solving supported by journal writing. However, Linn studied only one group of students, comparing second-semester performance, supported by journal writing, with first-semester performance without journaling. Linn's design does not control for maturation, history, or confounding variables; according to Kerlinger (1986, p. 317), this kind of design should have been limited to qualitative analysis. The present study instead utilized a "Simple Randomized Subjects" design (Kerlinger, 1986, p. 307, 319), a generalization of the post-test only true-experimental design described by Campbell and Stanley (1963,
Within each class, students were randomly assigned to two treatment groups, group A (Journals) and group B (Control). Both groups participate in four sessions, each consisting of instruction, computer practice, and writing activities (the treatment) followed by written post-tests for both independent variables (see Figure 1). Group A students participated in reflective thinking journal activities ($X_A$), while group B students performed alternate writing activities designed to elicit declarative responses ($X_B$).

Figure 1. Simple Randomized Subjects Design

Since each class contained members of both treatment groups, and since the two groups within each class received simultaneous instruction, the design controlled for inadvertent differences or biases in instruction among classes. An advantage of the design is that it is tolerant of minor adjustments to teaching procedures for subsequent sessions of the same lesson. Since each instructional session included an approximately equal number of students
from each treatment level, all variants of teaching procedures were present in equal amounts for both groups.

To test the hypotheses that thinking journal activity would improve problem solving performance and transfer (hypotheses 1 and 2), separate comparisons of group means were performed for each independent variable, Y₁ (Performance) and Y₂ (Transfer). At the end of the experiment, the mean of all observations (0) of Y₁ (Performance) associated with Method A (Journal) treatments, X₁A, X₂A, X₃A, and X₄A, were compared with the mean of all observations of Y₁ (Performance) associated with Method B (Control) treatments, X₁B, X₂B, X₃B, and X₄B (Figure 2). The same procedure was followed for measures of the other independent variable, Y₂ (Transfer).

Figure 2. Individual Treatments and Observations

<table>
<thead>
<tr>
<th>Group A</th>
<th>X₁A</th>
<th>X₂A</th>
<th>X₃A</th>
<th>X₄A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B</td>
<td>X₁B</td>
<td>X₂B</td>
<td>X₃B</td>
<td>X₄B</td>
</tr>
</tbody>
</table>

Because of the fact that the students were unfamiliar with both thinking journal writing and problem solving software, it was expected that the experimental treatment would have less impact on learning in earlier lessons. It was possible that significant differences between means might occur in latter testing occasions, while overall mean difference could remain insignificant. Therefore, separate
comparison of means for each test occasion were performed, to determine whether a trend would occur, and to test hypotheses 1 and 2 independently for each testing occasion.

2. Hypotheses 3 and 4

The study also examined a second, mediating variable, degree of metacognitive and reflective thought (Reflectivity). The original research proposal had specified that a measure of the number of reflective statements would be made, in which only comments that met the criteria defined by Higgins, Flower, & Petraglia (1991) would be counted. These conservative criteria require that only explicit evaluation, reasoning, or justification of problem solving processes or rules, or explicit comparison or consideration of alternative strategies could be considered reflective; other, less explicit or merely descriptive comments are considered to be at lower levels of metacognitive awareness. However, the subjects of this study initially responded at very low levels of awareness, with responses indicating lack of awareness, and difficulty recalling any explicit details of problem solving behavior. Since it was expected that there would be very few reflective comments recorded if these conservative criteria were applied, yielding no data to analyze, it was necessary to redesign the measure to include a greater range of
metacognitive thought, while adhering to the original intent to measure the student's degree of reflective thinking.

There were two possibilities for adjusting the measure. First, it would be possible to relax the conservative criteria to include even *inexplicit* evaluation, reasoning, justification, or comparison. This might be a valid approach for a study of undirected dialogue journal writing or unstructured student conversation, since it is of interest whether students, of their own inclination, offer any evaluation, reasoning, justification, or comparison. However, the approach is inappropriate under the conditions of this experiment, since the structured journal format *directs* students to offer such observations; it is only of interest if the student can offer *explicit* discussion. For instance, to the journal question "What other strategies could you use next time? Why would this be better?" two students answered as follows:

Student A: Go through all the Moppets slowly, and not fast. It would be better because sometimes you don't see whether it is short or tall, or thin or fat, or Bibbit or Gribbit!

Student B: I could try to spot the Moppet just by looking, and if I couldn't tell, I look closer. I don't think this would be better because you suggested the first one [strategy] to me.
Both students considered alternative strategies, because the journal question directed them to do so. Each is saying that they would work or look more carefully, although student A has specifically focused on the pace of the work as a source of problems. The important factor that discriminates these answers, however, is that student A explained why working too quickly is a bad strategy (an explicit rationale) - it is inefficient because the student sometimes failed to discriminate Moppet features, leading to careless mistakes. Student B's response lacks a definition of "looking closer", and it lacks a reason for the alternate strategy; in fact, the student subsequently recanted, deferring to an original strategy suggested by the teacher.

The second possibility was to design an instrument with which journal responses could be rated in terms of progressive degrees of metacognitive awareness, with the conservative definition of reflective thinking as the highest possible rating. Degree of metacognitive and reflective thought (Reflectivity), then, was measured by analyzing students' thinking journal entries, and assigning a numeric rating (0, 1, 2, or 3) to each individual response, according to the following scale: 0 = unaware, 1 = aware / mentioning; 2 = descriptive; 3 = reflective. Journals for all thirty Method A students were analyzed and ratings recorded. Each student received a mean reflectivity rating for each of the four journal writing sessions, and a
grand mean reflectivity rating was calculated for each student by averaging the four individual scores. Both the means for individual occasions and the grand mean were tested for association with problem solving performance, and with transfer of problem solving skills.

Procedures for Collection of Data

1. Computer software

Students practiced problem solving strategies and thinking skills with an interactive computer game titled Moptown Hotel. Moptown Hotel is an example of generic, or "content-free" problem solving software, meaning that it contains no subject-matter content. However, it teaches several thinking skills often used in mathematics and science. Moptown Hotel provides a progressive series of puzzles that help students develop the following kinds of skills (Perl, 1984B).

1. discriminating properties
2. using analogies
3. processing negative clues
4. controlling variables
5. making inferences
6. developing organized problem solving strategies

A learning task analysis, of a style similar to that described by Gagné (1988), was performed on the first five units of the Moptown Hotel sequence, so that (1) subskills
not identified by Perl could be listed, (2) component skills and strategies taught in one lesson and prerequisite to another could be identified, and (3) computer problems, "performance" worksheets, and "transfer" worksheets could be analyzed to ensure that the same component skills and strategies were required. The learning task analysis is included in Appendix A.

The software consists of a sequence of seven computer activities; four activities were selected for study. Based on the learning task analysis and the requirement that there be equivalent written problems for assessment, the following sequence of Moptown Hotel lessons was selected: (1) lesson one, Who's Next Door, an activity involving analogical relationships among graphic figures called "Moppets"; (2) lesson two, Change Me, in which students must control variables, changing characteristics of Moppets so they progress toward a provided "goal" Moppet (Figure 3); (3) lesson three, Spot Me, in which students identify a figure with the "greatest difference", defined as the Moppet with a unique trait; and (4) lesson five, Secret Pal, in which students must again control variables, changing traits and testing each change to see if the guess is closer or further from a secret "goal" Moppet.
Lesson four was not selected because its problem is very interactive, and does not lend itself to a written evaluation. It was possible to omit lesson four, despite the fact that Moptown Hotel lessons are sequential. The learning task analysis revealed that lesson five is most heavily dependent on skills developed in lessons one and two, while requiring only a single subskill introduced in lesson four (using negative information, i.e., "the Moppet isn't blue, therefore it is red").

However, an unsuccessful attempt to teach lesson five to the first class resulted in withdrawing it from the study. The lack of success with this lesson was indicated by difficulty in group discussions and in computer practice that lengthened the session to two hours, and by post-test results that indicated widespread random guessing.
Descriptive data collected during classroom discussions and computer practice, as well as post-test scores, were analyzed to determine the cause of failure. Analysis of lesson five data suggests that students understood the problem definition, and could successfully negotiate the mechanics of the worksheet, but that the difficulty was with the subordinate skills that had been taught in lesson two; specifically, they were unable to control variables (to vary only a single trait in order to test it), or to use a heuristic strategy of taking one step at a time, stopping to decide whether the new state is closer or further from the goal state. A review of descriptive observations collected during week two, and low lesson two post-test scores supported this conclusion. A reasonable alternate hypothesis is that lesson four, had it been taught, would have provided additional practice prior to lesson five. However, neither controlling variables nor "hill climbing" are required or practiced in lesson four. It is not possible, however, to rule out the alternate hypothesis.

Rather than continue with the risk that the final week of instruction would be unsuccessful, the study was suspended for one week. All three classes concluded the instructional sequence the following week with a second, modified treatment of lesson two. The final sequence of lessons was (1) lesson one *Who's Next Door*, (2) lesson two, *Change Me*, (3) lesson three, *Spot Me*, and (4) a second
treatment of lesson two. Instructional materials were changed for the second treatment, and the focus of instruction was adjusted to address the difficulties that had been observed. New, refined post-tests were also developed, to minimize the reactive effects of previous testing, and to improve reliability.

2. Instruction and guided practice

Each class was taught a sequence of four lessons, designated $X_1$, $X_2$, $X_3$, and $X_4$. Within each class, students were randomly assigned to two treatment groups, group A and group B. Within each class, all students in both treatment groups remained together, receiving the same instruction and computer practice. Minor adjustments to lesson plans were made between classes to correct unclear or confusing directions, to improve the allocation of time, or to change emphasis.

Instruction was adapted from the written lesson plans for teaching *Moptown Hotel*, included with the software. Each computer activity was preceded by an introductory lesson, including both direct instruction in strategies and rules, and a whole-group walk-through of the activity. The introductory lesson was followed by guided practice with the software, as directed by the *Moptown Hotel* lesson plans. A brief follow-up discussion of strategies and rules preceded journal writing activities. Each lesson and software
activity required approximately forty-five minutes to one hour; students typically practiced computer problems for at least twenty-five minutes. Although there were sufficient number of computers to allow students to engage in individual on-line activities, hardware problems with some of the computers made it necessary to have all students work in pairs.

The focus of instruction was twofold. First, lessons included direct instruction and "group discovery" of specific rules that could be applied to the problems at hand. Secondly, it sought to familiarize children with general "unordered" heuristic problem solved strategies (rather than ordered strategies or plans). Some of the heuristics were identified by the Moptown Hotel teacher's guide, and some were identified through the learning task analysis. Table 1 summarizes the heuristic strategies applicable in each Moptown Hotel lesson.

Table 1. Strategies Taught with Moptown Hotel

<table>
<thead>
<tr>
<th>Heuristic Strategy</th>
<th>Moptown Hotel Lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hill climbing</td>
<td>✔️ ✔️</td>
</tr>
<tr>
<td>2. Means-ends analysis</td>
<td>✔️ ✔️</td>
</tr>
<tr>
<td>3. &quot;Looking Ahead&quot; toward the goal</td>
<td>✔️</td>
</tr>
<tr>
<td>4. Discussing the problem with others</td>
<td>✔️ ✔️</td>
</tr>
<tr>
<td>5. Focus on one feature at a time</td>
<td>✔️ ✔️ ✔️</td>
</tr>
<tr>
<td>6. Controlling variables</td>
<td>✔️ ✔️</td>
</tr>
<tr>
<td>7. Being organized</td>
<td>✔️ ✔️</td>
</tr>
</tbody>
</table>
All students in treatment group A wrote responses to questions in a thinking journal, immediately following computer problem solving. Each journal activity required approximately twenty minutes. The journal was composed of four sets of questions, one for each Moptown Hotel lesson taught. All group A students answered identical questions, and each student was directed to answer all questions. The researcher collected and analyzed all student entries.

Figure 4. Sample Thinking Journal Questions

<table>
<thead>
<tr>
<th>Journal 3 - Spot Me</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What was the problem you had to solve?</td>
</tr>
<tr>
<td>2. What is the rule for spotting the Moppet with the greatest difference? Explain.</td>
</tr>
<tr>
<td>3. How did you discover this rule - in other words, what was your strategy? Describe your thinking.</td>
</tr>
<tr>
<td>4. What other strategies could you use next time? Why would this be better?</td>
</tr>
<tr>
<td>5. How did you feel working through the problem? Why?</td>
</tr>
</tbody>
</table>

Thinking journal questions for this study focus on procedural knowledge, and encourage reflection of problem solving activities. The format and questions are very similar to a thinking journal format documented by John Barell (1990, p. 225). One journal page was designed for each lesson. The first journal page included seven questions. Because of time constraints, however, subsequent
thinking journal pages were restricted to five questions. A typical journal page is shown in Figure 4.

Because the subjects were unfamiliar with problem solving concepts, each journal page included one question that asked students to identify and describe the problem, one question that asked them to discuss general strategies, and at least one question that directed them to describe rules applicable to the specific problem that they had solved. It was expected that by repeating these questions each week, students would gradually gain understanding of the concepts of problem, strategy, and rule, and improve their ability to provide explicit answers. Other questions included those addressing understanding one's feelings in problem situations (like frustration, anger, or elation), and questions that directed students to evaluate or compare problem solving strategies. All thinking journal pages are included in Appendix B.

4. Non-reflective journals (control group)

Students in group B followed identical procedures, but they performed an alternate writing activity involving questions that focus on students' declarative, rather than procedural knowledge. Group B students answered questions based on the content of the computer activities, questions not designed to foster reflective thought. All non-reflective journal pages are included in Appendix C.
Control group journal entries were not rated for quantitative analysis.

Non-reflective journals were designed to be as educationally valuable as possible, given their experimentally limited purpose. Hunkins (1989) suggests a method of classifying or developing questions according to their function, based on Bloom's Taxonomy of Educational Objectives. The questions in the thinking (experimental) journals for Moptown Hotel deal mostly with procedural knowledge, and with evaluation-level thinking that is at one extreme of Bloom's taxonomy. The alternate (control) journals instead focus on declarative, knowledge-level thinking, which is at the other extreme. Hunkins refers to these kinds of questions as "cognitive questions, knowledge classification" (1989, p. 54); they are characterized by their key words, such as "what", "when", "who", "identify", and "list." Control group questions focus on knowledge of terminology, specific facts, trends and sequences, and classification and categories. Related opinion questions, including those dealing with students' opinions of journal writing, allow students to achieve more depth of thought, while maintaining distance from problem solving issues.

5. Journal writing procedures

The researcher observed and interacted with children in both treatment groups as they wrote. Interaction included
encouraging children, answering their questions, clarifying journal questions, and watching for children who were experiencing difficulty. The process is similar to one described by Bode (1989, p. 570), who suggests that young children write fifteen to twenty minutes daily, in a bound journal, and that the teacher rotate around the room as children write, writing back immediately as they finish. In this study, however, written responses to student journal entries were made after journals were collected.

Each student received a personal, written response from the researcher, in order to encourage and support children in both treatment groups. The teacher's role in supporting journal writing is to establish a personal relationship with the student, in addition to offering information, assistance, or encouragement (Staton, 1984, p. 479). Staton states that "partial success in accomplishing [this] first goal is an enabling condition for being able to intervene in the student's train of thought to demonstrate a different, more reasonable way of construing events." To this end, personal responses were written in a brief, but open and conversational style. The student was addressed often by first name. Responses included encouraging comments, probing questions, reflecting and rewording the student's comments, clarification for students who misunderstood questions or concepts, and requests for more complete answers for those who offered only minimal responses.
Students were given an opportunity to read the responses prior to the following writing session, but the design did not allow for a continual dialogue (there was no opportunity to follow a line of thought from one session to the next).

Contrary to those who recommend both lack of punitive and beneficial consequences, foil stars were placed next to exemplary student responses, and Moptown Hotel bookmarks were given prior to each session for students who had written extensive or thoughtful responses (rewards were distributed equally between groups). There was insufficient time to allow rapport to build slowly through the development of a personal relationship; rewards were successful in quickly increasing student motivation.

Instruments

1. Post-tests of problem solving Performance and Transfer

Students completed eight written problem solving worksheets, used as posttest, reprinted or adapted from those included with Moptown Hotel. Four worksheets have the same content as Moptown Hotel computer activities, and were used to assess problem solving performance. Four worksheets have different content, but require the same skills; they were used to assess the student's ability to transfer skills to other kinds of problems. All performance and transfer instruments are included in appendix D. The post-tests followed thinking journal or alternate writing activities
for each treatment groups. Tests were administered during the same day as instruction for the first two occasions, except for a single class in week 2 that was tested the following day. Tests for the last two occasions were administered immediately following the journal-writing session.

Worksheets were collected, scored, and recorded for each individual. Each item in lesson one and lesson three posttests, both Performance and Transfer, was scored as either correct or incorrect (0 or 1), since no partially correct responses are possible. Each item in lesson two and lesson five posttests, both Performance and Transfer, was scored such that students received credit for partially correct solutions (0 through 4 points for each item). Determination of partial credit is not subjective. In fact, since students can generate many possible correct or partially correct solutions, tests were scored using the computer software program listed in Appendix F.

Validity of the Performance post-tests was determined by analyzing the degree to which the written form is equivalent to the computer exercise. Each test was selected from the sets of worksheets included with Moptown Hotel according to this criteria. Each test item was equivalent to a computer counterpart, in both structure and content. The major differences between written tests and computer exercises were in mechanical aspects, i.e., the student identified
Moppets by typing letters for each trait on the computer, but selected Moppets from sets of Moppet cards for written exercises.

Validity of the Transfer post-tests was determined by analyzing the degree to which the written test requires and benefits from the same strategies and rules used to solve the computer problem. Each test was selected from the sets of worksheets included with Moptown Hotel according to this criteria. The degree to which test items are isomorphic to computer problems varies, and efforts were made to adjust transfer problems to make them more equivalent.

Reliability of post-tests was established by dividing each test into odd-numbered and even-numbered sets, calculating a Pearson product-moment correlation coefficient between scores for the two halves, and calculating a coefficient of internal consistency, using the Spearman-Brown prophecy formula (Kerlinger, 1986, p. 412-413). There were two expectations regarding post-test reliability. First, it was recognized that reliability can be problematic with very short tests, and post-tests for Moptown Hotel ranged from three to ten items each. Secondly, it was expected that reliability of the instruments would increase with later testing occasions, because (1) later test forms were redesigned to minimize careless errors, and (2) post-testing was administered by the researcher, assisted by classroom teachers, instead of by classroom teachers alone,
for the last two testing occasions. Coefficients of internal consistency increased from .701 for Performance and .811 for Transfer in week 1, to .923 and .884, respectively, in week 3 (Table 2). It is not possible to compute a meaningful coefficient for the Transfer test for week 5, using the split-half method, because it contains only three items; however, correlations between its individual test items are in the same range as items for other transfer tests.

Table 2. Results of Split-Half Analyses of Internal Consistency

<table>
<thead>
<tr>
<th>Week</th>
<th>Performance</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.701</td>
<td>0.811</td>
</tr>
<tr>
<td>2</td>
<td>0.794</td>
<td>0.817</td>
</tr>
<tr>
<td>3</td>
<td>0.923</td>
<td>0.884</td>
</tr>
<tr>
<td>5</td>
<td>0.867</td>
<td></td>
</tr>
</tbody>
</table>

2. Rating procedures for analysis of student journals

Degree of metacognitive and reflective thought (Reflectivity) was operationally defined as a numeric rating from zero to three, based on an analysis of student writing in thinking journals. The rating scale is defined such that the endpoints match those of a model suggested by Swartz and Perkins, in which there is a continuum of "levels of thought..."
that are increasingly metacognitive" (Barell, 1990, p.211).
The model includes the following four levels:

**Tacit use**: Making decisions without really thinking about it.

**Aware use**: Being conscious of one's own decision-making processes.

**Strategic use**: Organizing one's thinking by selecting strategies for decision making.

**Reflective use**: Reflecting upon one's thinking, pondering how to proceed, and how to improve.

Because of the age and inexperience of the subjects of this study, and despite the fact that journal questions were designed to elicit reflective behavior, journal entries seldom exhibited signs of strategic or reflective use, and for the most part varied from complete lack of awareness through various degrees of descriptiveness regarding thinking and problem solving processes. Therefore, a four-point rating scale was designed to measure easily discriminated features of the writing of fourth grade students. The ratings are listed below. Since students were not given strict time limits on journal writing, and because it was observed that students often avoided questions they could not answer, unanswered questions were rated "unaware."
Level

0 Unaware

The response contains no indication of awareness of thinking or problem solving processes, rules, problem definitions, or feelings; or a direct statement by the child to that effect (i.e., "I don't know").

1 Aware / Mentioning

The response indicates that the child is aware that thinking or problem solving processes occurred, that there were rules that could be applied, that there was a problem to be solved, or that certain feelings resulted; however, the descriptions rely solely on reference to labels or obvious and non-informative surface characteristics.

2 Descriptive

The response provides explicit descriptions of thinking or problem solving processes, rules, problem definitions, or feelings.

3 Reflective

The response provides (1) explicit evaluation, reasoning, or justification of problem solving processes or rules, (2) explicit comparison or consideration of alternatives, or (3) statements that recognize the value or effectiveness of evaluating,
comparing, or justifying one's problem solving strategies (consistent with the definition of a reflective comment, below).

Explicit statements are those that refer to specific actions, procedures, conditions, rules, problem statements, or reasoning, regardless of the completeness or accuracy of the description. The scale is intended to measure, as objectively as possible, the student's ability to recall and articulate thinking and problem solving processes, whatever those processes were, and not to measure the correctness of the student's thinking.

The above written guidelines and a comprehensive list of generalized "template" responses were prepared, for constant reference by the researcher while rating responses in the thirty Method A journals (see Appendix E). In order to reduce bias due to irrelevant surface characteristics, such as handwriting, spelling, and grammar, the 660 Journal A responses were typed and corrected, and entered into a database. Student names were omitted. Spelling and ungrammatical language were corrected, unnecessary capitalization was removed, missing punctuation and articles were added, instances of "&" and "+" were changed to "and", and obvious missing words inserted. The result was a set of responses that are visually and grammatically consistent, allowing the evaluator to easily focus on the child's thinking. Responses were sorted by journal page and
question number, and printed so that the evaluator could analyze responses for an entire journal page at once, for greater consistency. An example of a corrected journal response is shown in Figure 5.

Figure 5. Example Student Journal Response

4. Did you find it important to look at the last Moppet? Why? What would happen if you didn't?

"The last Moppet is like your goal. You try to make the Moppet before him look like him, with the exception of 1 (2) characteristic(s). If I didn't look at it, I'd keep going through the problem until I got it out of luck."

All journals were analyzed and rated by the researcher. A random sample of sixty-six journal responses (10% of the total) was taken. A second rater was provided a brief overview of the Moptown Hotel problem set, and an independent rating of the sample set was performed, using the same written guidelines described above. Despite the second rater's unfamiliarity the Moptown Hotel problems,
inter-rater reliability, measured by calculating a Pearson's correlation coefficient, was .7811.

**Procedures for Analysis of Data**

The analyses of problem solving performance and transfer of skills each involved examining two treatment levels, \( X_1 \) and \( X_2 \), of a single independent variable, instructional method. The proposal had called for t-tests for significant difference between means, but it was discovered that through random occurrence, a greater number of female students were assigned to the treatment group than would be expected, and a greater number of males to the control group. Results of a cross tabulation and \( \chi^2 \) (chi-squared) goodness-of-fit test are summarized in Table 3. The cross tabulation indicates that 2 more females than would be expected were assigned to the experimental treatment, and 2 more males to the control group. However, the Pearson \( \chi^2 \) value is not significant, indicating that the level of variance would occur in a random sample 23.5% of the time \( (\chi^2 = 1.40318, p = .23619) \). Nonetheless, the fact that the control and experimental groups were out of balance with respect to gender requires that a factorial procedure be used to separate variance between that caused by the experimental independent variable, and that attributable to gender differences.
Table 3. **Chi-squared Goodness-of-fit Test for Random Assignment**

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Expected Value</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>METHOD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Journal</td>
<td>19</td>
<td>16.7</td>
</tr>
<tr>
<td>Gender</td>
<td>Control</td>
<td>16</td>
<td>18.3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>35</td>
<td>35.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Expected Value</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>11</td>
<td>13.3</td>
<td>-2.3</td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
<td>14.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>28.4%</td>
<td>44.4%</td>
</tr>
</tbody>
</table>

Column Total 30 33 63

Total 47.6% 52.4% 100.0%

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>Value</th>
<th>DF</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>1.40318</td>
<td>1</td>
<td>.23619</td>
</tr>
</tbody>
</table>

Because directional t-tests were replaced with non-directional factorial tests, research hypotheses 1 and 2 were converted to the null form ($H_0: \mu_1 = \mu_2$). Thus, results reported are more conservative than if directional tests had been used. Null hypotheses are:

1. There will be no significant difference on tests of problem solving performance between students who write thinking journals, and students who do not.
2. There will be no significant difference on tests of problem solving transfer between students who write thinking journals, and students who do not.
Research hypotheses 3 and 4, stated in the null form ($H_0: \beta = 0$) are as follows:

3. There will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving performance.

4. There will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving transfer.

1. Analysis of problem solving performance ($Performance, Y_1$)

Each student received four Performance ratings, one for each problem solving performance worksheet. An individual mean Performance rating was calculated for each student, by averaging that student's four Performance ratings. A grand mean Performance rating was calculated for each treatment group, by averaging all individual mean Performance ratings for that treatment. If any score was missing, the student's score was omitted from comparison of grand means. A two by two (Method by Gender) Multivariate Analysis of Variance (MANOVA) was performed to test the hypothesis that students who write thinking journals will score higher on tests of problem solving performance than students who do not write thinking journals, and also to separate variance due to gender. Although researchers often use multiple regression
with effect coding in place of ANOVA when cell sizes are unequal, a single-dependent-variable MANOVA can be used instead; it produces equivalent results, and its data entry and reports are similar to ANOVA.

The hypothesis was also tested individually for each of the four testing occasions. For each occasion, a mean Performance rating was calculated for each treatment group by averaging all individual mean Performance ratings for that treatment. Means were compared by performing a two by two (Method by Gender) Multivariate Analysis of Variance (MANOVA) for each testing occasion.

For all tests of Performance, the level of significance below which the null hypothesis would be rejected was arbitrarily set at the .05 level.

2. Analysis of problem solving transfer (Transfer, Y2):

Each student received four Transfer ratings, one for each problem solving performance worksheet. An individual mean Transfer rating was calculated for each student, by averaging that students four Transfer ratings. A grand mean Transfer rating was calculated for each treatment group, by averaging all individual mean Transfer ratings for that treatment. If any score was missing, the student's score was omitted from comparison of grand means. A two by two (Method by Gender) Multivariate Analysis of Variance (MANOVA) was performed to test the hypothesis that students
who write thinking journals will score higher on tests of problem solving transfer than students who do not write thinking journals, and also to separate variance due to gender.

The hypothesis was also tested individually for each of the four testing occasions. For each occasion, a mean Transfer rating was calculated for each treatment group by averaging all individual mean Transfer ratings for that treatment. Means were compared by performing a two by two (Method by Gender) Multivariate Analysis of Variance (MANOVA) for each testing occasion.

For all tests of Transfer, the level of significance below which the null hypothesis would be rejected was arbitrarily set at the .05 level.

3. Metacognitive and reflective thought (Reflectivity)

At the end of the experiment, each statement written in the thirty Method A thinking journals were rated for degree of metacognitive and reflective thought (Reflectivity), according to the procedures and criteria described under "Instruments". Each student received a Reflectivity rating for each of four weekly journal pages, defined as the average rating of responses made by that student. Individual mean Reflectivity ratings were calculated for each Method A student, by averaging the student's four Reflectivity ratings.
The individual mean Reflectivity ratings were tested for association with individual mean Performance ratings to test the hypothesis that there will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving performance. Association was tested by computing a regression coefficient for the set of paired variables, and testing it for statistical significance, according to a procedure described by Hinkle, et al. (1988, p. 448-452). Gender was provided as a second predictor in the model. Individual regression tests were also performed for each individual journal writing and testing occasion.

The individual mean Reflectivity ratings were tested for association with individual mean Transfer ratings to test the hypothesis that there will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving transfer. Association was tested by computing a regression coefficient for the set of paired variables, and testing it for statistical significance. Gender was provided as a second predictor in the model. Individual regression tests were also performed for each individual journal writing and testing occasion.
For all tests of association, the level of significance below which the null hypothesis would be rejected was arbitrarily set at the .05 level.
CHAPTER IV

RESULTS AND DISCUSSION

Introduction

The purpose of this study was to compare scores on tests of problem solving performance and transfer, for students who received software instruction supported by student thinking journal activities, with scores of students who received the same instruction, but who performed non-reflective writing activities. The study also examined (1) the relationship between the degree of metacognitive and reflective thought displayed in students' journal entries, and their problem solving performance; and (2) the relationship between the degree of metacognitive and reflective thought displayed, and the transfer of learned problem solving skills to a related, but different problem solving task.

Results of the analysis of scores on tests of problem solving performance and transfer, and results of analysis of students' thinking journals, are presented in this chapter. Descriptive statistics were also calculated to establish whether the data satisfied the assumptions of the statistical procedures used. Major hypotheses were tested using multivariate analysis of variance (MANOVA) and
multiple linear regression. Statistical analysis was performed using the microcomputer version of Statistics Package for the Social Sciences (SPSS-PC+).

Descriptive Statistics for Hypotheses 1 and 2

The assumptions of analysis of variance (ANOVA) are (1) the observations are normally distributed on the dependent variable in each group, (2) the variances in the population for all groups are equal, and (3) the observations are independent (Stevens, 1986, p. 199). The assumptions of MANOVA, used for this study because of its unequal cell sizes, are equivalent to ANOVA when it is used to test only one dependent variable. The effects in MANOVA of the combination of unequal cell sizes and unequal variance is well documented, an important consideration for this study.

The null hypothesis that scores on dependent variables for each group are normally distributed in the population was tested by performing a Shapiro-Wilks test for each group (female journal, female control, male journal, and male control). The tests were performed for both dependent variables, Performance and Transfer. Because the significance figures are high for seven of the eight tests (.26 < p < .97), the null hypothesis was retained for these groups of scores, and the assumption of normal distributions was satisfied. A significant test (p = .04) for one cell, female subjects, Method A (Journal) treatment, measure of
Transfer, indicates lack of normal distribution. However, since analysis of variance is robust for violations of normality (Stevens, 1986) lack of normal distribution in one cell can be tolerated.

The null hypothesis that the variances of all groups are equal in the population ($H_0: \sigma_1 = \sigma_2 = \ldots = \sigma_k$) was tested by performing a Levene test of homogeneity of variance for each dependent variable. For both tests, the null hypothesis was rejected, indicating that the variances are unequal. Therefore, the requirement of homogeneity of variance was not satisfied.

MANOVA is robust with respect to homogeneity of variance, as long as the cell sizes are approximately equal (largest / smallest < 1.5) (Stevens, 1986, p. 199). If the cell sizes are unequal, and variances are unequal, then the actual $\alpha$ can differ from the reported $\alpha$. If large group sizes are associated with smaller variances, the test becomes more liberal (e.g., the reported, or nominal $\alpha$ might be smaller than the actual $\alpha$, increasing the chance of a type I error. If large group sizes are associated with larger variances, the test becomes more conservative. The effects are slight, however, unless the group sizes differ greatly; at a 2:1 ratio, a reported $\alpha = .01$ would range from .007 to .02; a reported $\alpha = .05$ would range from .035 to .088 (Stevens, 1986, p. 218).
Table 4. **Cell Sizes and Variances for MANOVA Analyses**

<table>
<thead>
<tr>
<th>CELL</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, Group A</td>
<td>79.306</td>
<td>116.834</td>
<td>10.809</td>
<td>15</td>
</tr>
<tr>
<td>Female, Group B</td>
<td>77.563</td>
<td>147.016</td>
<td>12.125</td>
<td>10</td>
</tr>
<tr>
<td>Male, Group A</td>
<td>69.776</td>
<td>189.228</td>
<td>13.756</td>
<td>8</td>
</tr>
<tr>
<td>Male, Group B</td>
<td>59.917</td>
<td>579.943</td>
<td>24.082</td>
<td>15</td>
</tr>
</tbody>
</table>

**Transfer, Weeks 1,2,3,5**

**Performance, Weeks 1,2,3,5**

The cell sizes for the comparison of grand means, for both **Performance** and **Transfer** tests, vary by no more than 1.9:1, only slightly outside the range specified by Stevens. In both cases (**Performance** and **Transfer**), the most extreme differences among variances are in groups of approximately equal size (Table 4). Therefore, it is concluded that the difference between nominal and actual significance values is minimal.

It is assumed that measures are independent, meaning that performance by one subject does not influence performance of another subject. For this study, treatment groups were formed by random assignment, and measured performance on post-tests were assumed to be independent, to the degree that performance of public school children can be independent. In practice, there is a tendency for children to influence each other's performance; in fact, it has been
documented that this can be problematic in research of cooperative learning with microcomputers, since students working in pairs influence each other's performance (Stevens, 1986, p. 203). Collusion by subjects during testing is also a factor. As post-tests for this study were scored, signs of possible collusion were noted; also, sets of item scores for individual tests were sorted and inspected for signs of matching patterns. Both methods revealed only a small number of suspected problems. It is assumed, therefore, that the data for this study are independent.

Testing Hypothesis 1

H₀: μ₁ = μ₂: There will be no significant difference on tests of problem solving performance between students who write thinking journals, and students who do not.

Mean scores for individual testing occasions of problem solving Performance, and grand means for all four testing occasions, are summarized in Table 5. The same data are presented graphically in Figures 6 and 7.

A 2 by 2 factor Multivariate Analysis of Variance (MANOVA) was performed to determine whether significant differences exist between grand mean Performance scores, or if significant interactions between gender and method exist. An individual mean Performance rating was calculated for
Table 5. *Performance and Grand Mean Performance Scores*

<table>
<thead>
<tr>
<th>Performance</th>
<th>Female</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>74.3</td>
<td>55.5</td>
<td>97.1</td>
<td>90.1</td>
<td>79.25</td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>76.9</td>
<td>63.0</td>
<td>93.8</td>
<td>77.5</td>
<td>77.79</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Male</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>59.1</td>
<td>48.3</td>
<td>92.5</td>
<td>55.6</td>
<td>63.88</td>
</tr>
<tr>
<td>Control Group</td>
<td>50.0</td>
<td>57.9</td>
<td>95.0</td>
<td>72.5</td>
<td>68.85</td>
</tr>
</tbody>
</table>

Figure 6. *Problem Solving Performance for Females*

Figure 7. *Problem Solving Performance for Males*
each student, by averaging that students four Performance ratings. Individual mean Performance ratings were entered into a 2 by 2 MANOVA, with each score identified by Method (Journal or Control) and Gender. The results of comparisons of group means are summarized in Table 6. Results indicate that while females scored significantly higher than males in tests of problem solving performance ($p < .05$), there was no significant difference associated with Method ($p = .514$). Therefore, the null hypothesis that there is no significant difference on tests of problem solving performance between students who write thinking journals, and students who do not, is retained. There was no significant interaction effect between Gender and Method ($p = .829$), indicating that for problem solving performance, there was no difference in the effect of journal writing on Performance between male and female students.

Table 6. Results of MANOVA for Mean Problem Solving Performance

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>11606.69</td>
<td>46</td>
<td>252.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>1654.39</td>
<td>1</td>
<td>1654.39</td>
<td>6.56</td>
<td>.014</td>
</tr>
<tr>
<td>METHOD</td>
<td>109.30</td>
<td>1</td>
<td>109.30</td>
<td>.43</td>
<td>.514</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>11.94</td>
<td>1</td>
<td>11.94</td>
<td>.05</td>
<td>.829</td>
</tr>
</tbody>
</table>
Because of the expectation that early journal sessions might be less effective than subsequent ones, the hypothesis was also tested individually for each of the four testing occasions, using four separate MANOVAs. However, when a series of individual tests are used in this manner, the chance of at least one type I error increases over that of a single test. If the criterion for rejecting has been set at the .05 level, then chance of at least one type I error occurring when there are four comparisons is $1 - (.95 \times .95 \times .95 \times .95) = .19$. The risk of type I error is lower, however, if a more conservative criterion is adopted, or if reported significance levels are lower. At the .01 level of significance, the chance of at least one type I error is only .039. This relationship was considered when interpreting the results of individual tests, with concern for reported alpha levels exceeding .01.

For each comparison, individual student Performance ratings were entered into a 2 by 2 MANOVA, with each score identified by Method (Journal or Control) and Gender. The results of comparisons of group means are summarized in Table 7.
Table 7. Results of MANOVA for Performance, Weeks 1 - 5

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>33253.36</td>
<td>56</td>
<td>593.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>6368.60</td>
<td>1</td>
<td>6368.60</td>
<td>10.72</td>
<td>.002</td>
</tr>
<tr>
<td>METHOD</td>
<td>151.74</td>
<td>1</td>
<td>151.74</td>
<td>.26</td>
<td>.615</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>487.79</td>
<td>1</td>
<td>487.79</td>
<td>.82</td>
<td>.369</td>
</tr>
</tbody>
</table>

Tests of Significance for Performance, week 2

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>51711.05</td>
<td>52</td>
<td>994.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>514.39</td>
<td>1</td>
<td>514.39</td>
<td>.52</td>
<td>.475</td>
</tr>
<tr>
<td>METHOD</td>
<td>995.41</td>
<td>1</td>
<td>995.41</td>
<td>1.00</td>
<td>.322</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>15.83</td>
<td>1</td>
<td>15.83</td>
<td>.02</td>
<td>.900</td>
</tr>
</tbody>
</table>

Tests of Significance for Performance, week 3

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>8602.94</td>
<td>54</td>
<td>159.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>38.02</td>
<td>1</td>
<td>38.02</td>
<td>.24</td>
<td>.627</td>
</tr>
<tr>
<td>METHOD</td>
<td>2.27</td>
<td>1</td>
<td>2.27</td>
<td>.01</td>
<td>.905</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>117.17</td>
<td>1</td>
<td>117.17</td>
<td>.74</td>
<td>.395</td>
</tr>
</tbody>
</table>

Tests of Significance for Performance, week 5

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>31492.88</td>
<td>53</td>
<td>594.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>5326.54</td>
<td>1</td>
<td>5326.54</td>
<td>8.96</td>
<td>.004</td>
</tr>
<tr>
<td>METHOD</td>
<td>63.33</td>
<td>1</td>
<td>63.33</td>
<td>.11</td>
<td>.745</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>2968.32</td>
<td>1</td>
<td>2968.32</td>
<td>5.00</td>
<td>.030</td>
</tr>
</tbody>
</table>

There was no significant interaction effect between Gender and Method for weeks one, two, and three, however, the test indicates a significant interaction effect between Gender and Method for week five (p < .05). For week five, Performance scores for females in the Journal group were higher than for females in the Control group, and Performance scores for males in the Journal group were lower.
than for males in the Control group. The interaction plot illustrates that there was a pattern of ordinal interaction, that is, the cell means for females are always greater than for males (Figure 8).

Figure 8. Interaction Plot, Week Five, Performance

Analysis of main effects for Gender indicates that females scored significantly higher in problem solving performance than did males in week one (p < .01), there were no significant differences for week two (p = .475) and week three (p = .627), and the main effects are discounted for week five because of the interaction already mentioned. Significance levels reported are sufficiently large or small, such that the effects of using multiple tests on actual $\alpha$ are unimportant. A post hoc $\omega^2$ (omega squared) measure of association indicates that approximately 14% of the variation in the dependent variable, Performance, is accounted for by gender.
Analysis of main effects for Method indicates that there were no significant differences at the .05 level between Journal and Control group scores for any weekly treatment. For both main effects tests, significance levels reported are sufficiently large such that the effects of using multiple tests on actual $\alpha$ are unimportant. Therefore, when tested individually for each occasion, the null hypothesis that there is no significant difference on tests of problem solving performance between students who write thinking journals, and students who do not, is retained.

Testing Hypothesis 2

$H_0: \mu_1 = \mu_2$: There will be no significant difference on tests of problem solving transfer between students who write thinking journals, and students who do not.

Scores for individual testing occasions of problem solving Transfer, and grand means for all four testing occasions, are summarized in Table 8. The same data are presented graphically in Figures 9 and 10.

A 2 by 2 factor Multivariate Analysis of Variance (MANOVA) was performed to determine whether significant differences exist between grand mean Transfer scores, or if significant interactions between gender and method exist. An individual mean Transfer rating was calculated for each student, by averaging that students four Transfer ratings. Individual mean Transfer ratings were entered into a 2 by 2
Table 8. Transfer and Grand Mean Transfer Scores

<table>
<thead>
<tr>
<th></th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>90.4</td>
<td>70.6</td>
<td>75.9</td>
<td>82.8</td>
<td>79.92</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>88.9</td>
<td>62.5</td>
<td>76.9</td>
<td>69.4</td>
<td>74.43</td>
<td></td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>90.9</td>
<td>38.8</td>
<td>71.0</td>
<td>74.2</td>
<td>68.71</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>78.4</td>
<td>38.3</td>
<td>74.0</td>
<td>46.7</td>
<td>59.36</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Problem Solving Transfer for Females

Figure 10 - Problem Solving Transfer for Males
MANOVA, with each score identified by Method (Journal or Control) and Gender. The results of comparisons of group means are summarized in Table 9. Results indicate that while females scored significantly higher in tests of problem solving transfer than did males (p < .01), there was no significant difference associated with Method (p = .255). Therefore, the null hypothesis that there is no significant difference on tests of problem solving transfer between students who write thinking journals, and students who do not, is retained. There was no significant interaction effect between Gender and Method, indicating that for problem solving transfer, there was no difference in the effect of journal writing on Transfer between male and female students.

Table 9. Results of MANOVA for Mean Problem Solving Transfer

<table>
<thead>
<tr>
<th>MANOVA * * ANALYSIS OF VARIANCE * *</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests of Significance for Transfer, weeks 1-5</td>
<td>SS</td>
<td>DF</td>
<td>MS</td>
<td>F</td>
</tr>
<tr>
<td>Source of Variation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WITHIN CELLS</td>
<td>12402.38</td>
<td>44</td>
<td>281.87</td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td>2375.44</td>
<td>1</td>
<td>2375.44</td>
<td>8.43</td>
</tr>
<tr>
<td>METHOD</td>
<td>375.67</td>
<td>1</td>
<td>375.67</td>
<td>1.33</td>
</tr>
<tr>
<td>GENDER BY METHOD</td>
<td>183.84</td>
<td>1</td>
<td>183.84</td>
<td>.65</td>
</tr>
</tbody>
</table>

The hypothesis was also tested individually for each of the four testing occasions, using four separate MANOVAs. As in the case of hypothesis testing for
Performance, the results were interpreted with possible differences between observed and actual \( \alpha \) in mind.

Table 10. **Results of MANOVA for Transfer, Weeks 1 - 5**

| MANOVA ** ANALYSIS OF VARIANCE ** | Tests of Significance for Transfer, week 1 | | Tests of Significance for Transfer, week 2 | | Tests of Significance for Transfer, week 3 | | Tests of Significance for Transfer, week 5 |
|-------------------------------------|------------------------------------------|---|------------------------------------------|---|------------------------------------------|---|
| Source of Variation                | SS            | DF | MS            | F   | Sig of F        | Source of Variation                | SS            | DF | MS            | F   | Sig of F        | Source of Variation                | SS            | DF | MS            | F   | Sig of F        | Source of Variation                | SS            | DF | MS            | F   | Sig of F        |
| WITHIN CELLS                       | 25987.48      | 55 | 472.50        | .73 | .398           | WITHIN CELLS                       | 32221.20      | 51 | 631.79        | 16.42 | .001           | WITHIN CELLS                       | 23605.51      | 54 | 437.14        | .48 | .492           | WITHIN CELLS                       | 41442.68      | 53 | 781.94        | 4.33 | .042           |
| GENDER                             | 343.01        | 1  | 343.01        | 1.44 | .235           | GENDER                             | 10371.73      | 1  | 10371.73      | 1   | .541           | GENDER                             | 208.95        | 1  | 208.95        | .13 | .723           | GENDER                             | 3386.43       | 1  | 3386.43       | 7.32 | .009           |
| METHOD                             | 680.14        | 1  | 680.14        | .90 | .347           | METHOD                             | 239.19        | 1  | 239.19        | .38 | .581           | METHOD                             | 55.35         | 1  | 55.35         | .03 | .859           | METHOD                             | 5725.36       | 1  | 5725.36       | .87 | .355           |
| GENDER BY METHOD                   | 424.74        | 1  | 424.74        | .31 | .541           | GENDER BY METHOD                   | 194.61        | 1  | 194.61        | .31 | .581           | GENDER BY METHOD                   | 13.99         | 1  | 13.99         | .03 | .859           | GENDER BY METHOD                   | 680.62        | 1  | 680.62        | .87 | .355           |

For each comparison, individual student Transfer ratings were entered into a 2 by 2 MANOVA, with each score identified by Method (Journal or Control) and Gender. The results of comparisons of group means are summarized in Table 10. No significant interaction effects were reported,
so main effects could be examined. Results indicate that females scored significantly higher than males in the test of problem solving Transfer for week two (p < .001), and possibly higher than males in week five (p = .042, which is within the area of concern due to multiple tests), but there were no significant gender differences for weeks one and three. For the first three weekly testing occasions, there were no significant differences between Journal and Control treatments.

However, there was a significant difference in measures of Transfer between Journal and Control groups for week five; students who wrote thinking journals scored significantly higher than students who did not (p < .01). The .01 significance level reported is sufficiently small such that the effects of using multiple tests on actual α are unimportant. Therefore, when tested individually, for the last occasion, the null hypothesis that there is no significant difference on tests of problem solving transfer between students who write thinking journals, and students who do not, is rejected.

A post hoc $\omega^2$ measure of association indicates that approximately 15% of the variation in the dependent variable, Transfer, is associated with levels of the two independent variables, Method and Gender. Method (journal writing versus control) was more highly associated with Transfer than gender; Method accounts for approximately 10%,
twice that associated with gender. However, the total (15%) indicates that a relatively small proportion of the variation in the dependent variable was explained. Measures of association are useful, however, they cannot be used as an indication of practical significance. There are many factors that can limit measures of association (Stevens, 1986, p. 135), including use of dichotomous factors, such as gender; the use of experimental designs rather than correlational designs; and the fact that most behaviors have multiple causes that cannot be fully explained with a single independent variable. This latter limitation was evident in the results reported in Table 10, in that there are high levels of "within cells" variation.

An examination of the group means for week five indicates that students who wrote thinking journals scored significantly higher in problem solving Transfer than students who did not write thinking journals. Non-statistical examination of individual group scores for week five indicates that mean Transfer scores were higher for the Journal group than for the Control group in all groups tested. They were higher for the Journal treatment for both males and females (Table 8 - males, 74.2 vs. 46.7; females, 82.8 vs. 69.4), and they were higher for the Journal treatment in all three classes (Table 11).
Table 11. Transfer Score for Individual Classes. Week Five

<table>
<thead>
<tr>
<th>Scores of Problem solving Transfer, Week five</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1, Journals 77.78</td>
</tr>
<tr>
<td>Class 1, Control 64.81</td>
</tr>
<tr>
<td>Class 2, Journals 82.41</td>
</tr>
<tr>
<td>Class 2, Control 68.33</td>
</tr>
<tr>
<td>Class 3, Journals 78.47</td>
</tr>
<tr>
<td>Class 3, Control 43.18</td>
</tr>
</tbody>
</table>

For individual observations, there were no significant interaction effects between Gender and Method, indicating that there was no difference in the effect of journal writing on Transfer between male and female students.

**Testing Hypothesis 3**

$H_0: \beta = 0$: There will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving performance.

*Individual mean Reflectivity ratings* were tested for association with *individual mean Performance ratings* to test the hypothesis that there will be no significant relationship between reflectivity and problem solving performance. Association was tested using multiple linear regression analysis; values were entered for both Reflectivity and Gender as predictors of problem solving Performance.
Table 12. Results of Multiple Linear Regression Analysis. 
Reflectivity with Performance

<table>
<thead>
<tr>
<th><strong>MULTIPLE REGRESSION</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
</tr>
<tr>
<td>Performance, Weeks 1,2,3,5</td>
</tr>
<tr>
<td><strong>Multiple R</strong></td>
</tr>
<tr>
<td><strong>R Square</strong></td>
</tr>
</tbody>
</table>

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>2183.56858</td>
<td>1091.78429</td>
</tr>
<tr>
<td>Residual</td>
<td>21</td>
<td>5277.36892</td>
<td>251.30328</td>
</tr>
</tbody>
</table>

**F** = 4.34449  Signif F = .0264

--- Variables in the Equation ---

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>T</th>
<th>Sig T</th>
</tr>
</thead>
<tbody>
<tr>
<td>JOURNALS</td>
<td>29.951610</td>
<td>13.348901</td>
<td>.413166</td>
<td>2.244</td>
<td>.0358</td>
</tr>
<tr>
<td>GENDER</td>
<td>7.002184</td>
<td>3.353170</td>
<td>.384528</td>
<td>2.088</td>
<td>.0491</td>
</tr>
</tbody>
</table>

The results are summarized in Table 12. The results indicate that both Reflectivity (listed as "JOURNALS" in SPSS reports) and Gender are significant predictors of problem solving performance ($\beta = .413$, $p < .05$). Multiple R indicates the degree to which the combined predictor variables are correlated with the observed values of the criterion variable; the multiple R value listed in Table 12 indicates a moderately high degree of correlation between the two predictor variables and problem solving Performance ($R = .541$). $R^2$ indicates the "proportion of the variation in the criterion variables that can be attributed to the variation of the combined predictor variables" (Hinkle, p. .
The $R^2$ value listed in Table 12 indicates that the two predictor values predict approximately twenty-nine percent of the variation in the criterion variable, problem solving Performance. The significant beta for the gender variable reflects higher Performance scores for females ($\beta = .385$, $p < .05$), already established by the MANOVA used to test hypothesis 1.

Visual analysis of plots generated by SPSS with the regression analysis for hypothesis 3 indicate that the assumptions for multiple regression analysis were satisfied. The assumptions of multiple regression are that the conditional distributions are both normal and homoscedastic (the standard deviations of conditional distributions are equal) (Hinkle, p. 452). These assumptions were verified using SPSS "RESIDUALS" and "PLOTS" procedures.

The assumption of normality is tested visually by plotting a histogram of standard residuals; the distribution should appear fairly normal. This is difficult to discriminate with a small number of cases (once missing scores for weekly journals and posttests are omitted, there are only 24 cases for the regression analysis of mean scores). The residual plot for this test is inconclusive. Most of the residual values for the regression model summarized in Table 12 follow a normal distribution pattern. If non-normality were assumed, it could be corrected by transforming the data, however, there are not enough points
to reliably determine a method of transformation. Therefore, it was assumed that the residuals were sufficiently normal.

Homoscedasticity is tested by plotting residuals against predicted values, and residuals against the values of the independent variables. If variance is constant, the plots of residuals produce a uniform "spread", neither increasing nor decreasing with values of the independent variables or predicted values. Plots of residuals for the regression analysis for hypothesis 3 show no such patterns, therefore, the homoscedasticity requirement was satisfied. The same residual plots can also be used to support the assumption that the regression statement models a linear relationship; if the points fall into a pattern other than a uniform horizontal band (i.e., curvature or diagonal, the relationship might be non-linear. The plots show no particular pattern, therefore, linearity was established.

A final consideration is that the predictor variables selected should be highly correlated with the criterion variable \textit{Performance}, and must not be correlated with one another. For the current analysis, it must be established that gender cannot be used to predict Reflectivity, and vise-versa. Table 13 lists the mean reflectivity ratings of male and female students, and indicates that Reflectivity ratings are almost identical across the groups. Therefore, there is no association between gender and \textit{Reflectivity}.\n
Table 13. Reflectivity Scores, Weeks 1 - 5

<table>
<thead>
<tr>
<th></th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 5</th>
<th>Weeks 1-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.95</td>
<td>1.16</td>
<td>1.48</td>
<td>1.38</td>
<td>1.24</td>
</tr>
<tr>
<td>St.Dev.</td>
<td>0.32</td>
<td>0.39</td>
<td>0.25</td>
<td>0.42</td>
<td>0.26</td>
</tr>
<tr>
<td>Male Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.92</td>
<td>1.18</td>
<td>1.40</td>
<td>1.42</td>
<td>1.26</td>
</tr>
<tr>
<td>St.Dev.</td>
<td>0.28</td>
<td>0.35</td>
<td>0.28</td>
<td>0.33</td>
<td>0.23</td>
</tr>
<tr>
<td>All Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.94</td>
<td>1.17</td>
<td>1.45</td>
<td>1.39</td>
<td>1.25</td>
</tr>
<tr>
<td>St.Dev.</td>
<td>0.30</td>
<td>0.37</td>
<td>0.26</td>
<td>0.38</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The T value of the coefficient for journal Reflectivity as a predictor of Performance (3.878) was significant (p < .05). The multiple R indicates that the model predicts a moderately high proportion of the variation in the criterion variable (Performance), and the assumptions of linear regression were satisfied. Therefore, the null hypothesis was rejected; it was established that students who had higher Reflectivity scores also had higher problem solving Performance scores. It must be cautioned that no causality can be inferred from these data; the two variables are associated, indicating that more reflective writers were also better problem solvers, but it cannot be inferred that greater reflective thinking improves problem solving performance.

Weekly Reflectivity scores were tested for association with their corresponding tests of Performance (i.e., Reflectivity rating for week one with Performance rating for week one), using individual regression analyses for each of the four occasions. However, no individual regression
revealed significant association for Reflectivity at the .05 level. This indicates that individual weekly Reflectivity scores are not strong predictors of scores for the Performance posttests administered the same day.

Post hoc analyses of the correlations between weekly Reflectivity ratings and scores on weekly Performance posttests, as well as for mean Reflectivity ratings and mean Performance scores, were performed. The analyses are useful in understanding the source of the associations revealed by the regression analysis. The results, summarized in Table 14, indicate that, overall, Reflectivity ratings were moderately and significantly correlated with Performance (r = .38, p < .05). Also, the Reflectivity ratings of males, after week one, were highly correlated with scores of Performance, while those of females were not. Overall correlation for males, across all occasions, was .87, p < .01.

Table 14. Correlations. Reflectivity with Performance

<table>
<thead>
<tr>
<th>Correlations (r) - Journal ratings with Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week:</td>
</tr>
<tr>
<td>All group A</td>
</tr>
<tr>
<td>0.1712</td>
</tr>
<tr>
<td>Female only</td>
</tr>
<tr>
<td>Male only</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
Testing Hypothesis 4

H₀: β = 0: There will be no significant relationship between the degree of metacognitive and reflective thinking evident in student journal entries, and scores on tests of problem solving transfer.

Individual mean Reflectivity ratings were tested for association with individual mean Transfer ratings to test the hypothesis that there will be no significant relationship between reflectivity and problem solving transfer. Association was tested using multiple linear regression analysis; values were entered for both Reflectivity and Gender as predictors of problem solving Transfer.

Table 15. Results of Multiple Linear Regression Analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>T</th>
<th>Sig T</th>
</tr>
</thead>
<tbody>
<tr>
<td>JOURNALS</td>
<td>22.855149</td>
<td>8.979322</td>
<td>.455376</td>
<td>2.545</td>
<td>.0193</td>
</tr>
<tr>
<td>GENDER</td>
<td>5.940208</td>
<td>2.329654</td>
<td>.456184</td>
<td>2.550</td>
<td>.0191</td>
</tr>
</tbody>
</table>
The results are summarized in Table 15. The results indicate that both Reflectivity and Gender are significant predictors of problem solving transfer (Reflectivity, $\beta = .455, p < .05$; Gender, $\beta = .456, p < .05$). The multiple $R$ value listed in Table 15 indicates a moderately high level of correlation between the two predictor variables and problem solving Transfer ($R = .607$). The $R^2$ value listed in Table 15 indicates that the two predictor values predict approximately thirty-seven percent of the variation in the criterion variable, problem solving Transfer. The significant beta for the gender variable reflects the higher Transfer scores for females, already established by the MANOVA used to test hypothesis 2.

Visual analysis of plots generated by SPSS with the regression analysis for hypothesis 4 indicate that the assumptions for multiple regression analysis were satisfied. The assumption that the conditional distributions are normal was tested visually by plotting a histogram of standard residuals, although it has already been stated that it is difficult to distinguish patterns with a small number of cases. Most points on the residual plot fall in a normal pattern, but the visual analysis for this test was also inconclusive, and there are not enough points to reliably determine a method of transformation. Therefore, it was assumed that the residuals were sufficiently normal.
Homoscedasticity (constant variance) was tested by plotting residuals against predicted values, and residuals against the values of the independent variables. Plots of residuals for the regression analysis for hypothesis 4 show no patterns of increasing or decreasing "spread", therefore, the requirement of homoscedasticity was satisfied. The same residual plots were also used to test the assumption that the regression statement models a linear relationship; the points do not fall into a non-horizontal pattern, therefore, linearity was established.

It was determined in the previous section that the predictor variables selected are not correlated among themselves. Table 13 (above) lists the mean reflectivity ratings of male and female students, and indicates that Reflectivity ratings are equal across the groups. Therefore, there is no association between gender and Reflectivity.

The T value of the coefficient for journal Reflectivity as a predictor of Transfer (2.545) was significant (p < .05). The multiple R indicates that the model predicts a moderately high proportion of the variation in the criterion variable (Transfer), and the assumptions of linear regression were satisfied. Therefore, the null hypothesis was rejected; it was established that students who had higher Reflectivity scores also had higher problem solving Transfer scores. Again, no causality can be inferred from
these data; the two variables are associated, indicating that more reflective writers were also better at transferring problem solving skills, but it cannot be inferred that greater reflective thinking improves the possibility of successful transfer of problem solving skills.

Weekly Reflectivity scores were tested for association with their corresponding tests of Transfer (i.e., Reflectivity rating for week one with Transfer rating for week one), using individual regression analyses for each of the four occasions. Individual regression analyses for weeks 1, 2, and 3 produced no significant associations for Reflectivity at the .05 level. The regression analysis for week five produced a significant T for Reflectivity ($\beta = .388$, $p < .05$), although the overall model is weak; the F value is not significant (significance of $F = .0823$), and the $R^2$ value indicates that only about 19 percent of the variation in the criterion variable is associated with the predictor variables.

However, the results of this test are of interest when the results of earlier hypothesis testing is recalled. Hypotheses 1 and 2 stated that the scores of problem solving posttests would be no different for journal writing students than for non-journal writing students; the null hypothesis was rejected for only one of the eight MANOVAs performed to test these hypotheses. The null hypothesis was rejected for
the test of Transfer, week five, suggesting that journal writing had a positive effect on Transfer for the final occasion only. It is also evident from the regression analyses for separate occasions that only one of the eight tests produced a significant T for the Reflectivity variable; for the final occasion only, the Reflectivity ratings of student journal writing was a significant predictor of problem solving transfer. This alignment of results will be examined when drawing conclusions regarding the main hypotheses of this study.

Table 16. Results of Multiple Linear Regression Analysis. Reflectivity with Transfer. Week 5

<table>
<thead>
<tr>
<th><strong>MULTIPLE REGRESSION</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable: Transfer, Week 5</td>
</tr>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>2200.38256</td>
<td>1100.19128</td>
</tr>
<tr>
<td>Residual</td>
<td>24</td>
<td>9512.58041</td>
<td>396.35752</td>
</tr>
<tr>
<td>F =</td>
<td>2.77575</td>
<td>Signif F = 0.0823</td>
<td></td>
</tr>
</tbody>
</table>

------------------- Variables in the Equation -------------------

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>T</th>
<th>Sig T</th>
</tr>
</thead>
<tbody>
<tr>
<td>JOURNAL</td>
<td>22.380757</td>
<td>10.723804</td>
<td>0.387656</td>
<td>2.087</td>
<td>0.0477</td>
</tr>
<tr>
<td>GENDER</td>
<td>5.496769</td>
<td>4.005731</td>
<td>0.254805</td>
<td>1.372</td>
<td>0.1827</td>
</tr>
</tbody>
</table>
A post hoc analysis of the correlations between both weekly Reflectivity ratings and scores on weekly Transfer posttests, and mean Reflectivity ratings and mean Transfer scores, was performed. As when testing hypothesis 3, the analysis is useful in understanding the source of the associations revealed by the regression analysis. The results, summarized in Table 17, indicate that the Reflectivity ratings of females were not correlated with Transfer for any testing occasion, nor were they for the mean of all testing occasions. The Reflectivity ratings of males were moderately, but not significantly correlated with scores of Transfer for the first two testing occasions, and not correlated for the third. However, Reflectivity ratings for males were highly and positively correlated with Transfer for the final occasion ($r = .82, p < .001$). The high correlation among males, despite a low, but positive correlation among females, produced a moderate positive correlation for all group A students for the final occasion, $r = .34, p = .052$. When both groups are combined and mean scores analyzed, the result is a significant positive relationship between mean Reflectivity and mean Transfer, $r = .4034, p < .05$).
Table 17. Correlations: Reflectivity with Transfer

<table>
<thead>
<tr>
<th>Correlations (r) - Journal ratings with Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week:</td>
</tr>
<tr>
<td>All group A</td>
</tr>
<tr>
<td>Female only</td>
</tr>
<tr>
<td>Male only</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
† p = .052

The results drawn from the data in Table 17 are that (1) the significant association between mean Reflectivity and mean Transfer identified in the regression equation for all testing occasions combined (Table 15); and (2) the significant association between Reflectivity and Transfer in the regression equation for the final testing occasion (Table 16) were both largely due to a very high relationship between Reflectivity and Transfer for males; Reflectivity scores for females did not contribute substantially.

Reflectivity and Journal Writing

Quantitative analysis of journal data revealed that the quantity and quality of student journal writing improved as the experiment progressed. The results of a word count of student thinking journal responses are summarized in Table 18. The average number of words written by Journal Method students per journal session increased slightly, from 87.3 words to 92.9 words, although the first journal contained...
more questions. The average number of words per journal, for five-question journals, increased from 84 in week two to 92.9 in week five. The length of individual responses increased by almost fifty percent, from 12.5 words per response in week one, to 18.6 words per response in week five.

Table 18. Changes in the Length of Students' Journal Responses

<table>
<thead>
<tr>
<th>Week</th>
<th>Journals</th>
<th>Words / Journal</th>
<th>Questions</th>
<th>Mean Response Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39</td>
<td>87.3</td>
<td>7</td>
<td>12.5 words per response</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>84.0</td>
<td>5</td>
<td>16.8 words per response</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>91.5</td>
<td>5</td>
<td>18.3 words per response</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>92.9</td>
<td>5</td>
<td>18.6 words per response</td>
</tr>
</tbody>
</table>

The quality and efficacy of student thinking journal writing, for this study, is assumed to be related to the measure of Reflectivity used in the tests of hypotheses 3 and 4. Mean student Reflectivity ratings, which range from 0 (unaware) to 3 (reflective), increased from .94 in week one, to 1.25 in week five. T-tests of dependent samples reveal that the weekly increases in reflectivity for weeks 1 through 3 were significant (week one to week two, \( p < .01 \); week two to week three, \( p < .001 \)). The slight decline in week five is not significant \( (p = .483) \). The data in Table 19 also indicate that reflectivity ratings for males and females were almost identical for all four weekly measures.
Table 19. **Reflectivity Scores, Weeks 1 - 5**

<table>
<thead>
<tr>
<th></th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Students</td>
<td>0.95</td>
<td>1.16</td>
<td>1.48</td>
<td>1.38</td>
<td>1.24</td>
</tr>
<tr>
<td>Male Students</td>
<td>0.92</td>
<td>1.18</td>
<td>1.40</td>
<td>1.42</td>
<td>1.26</td>
</tr>
<tr>
<td>All Group A Students</td>
<td>0.94</td>
<td>1.17</td>
<td>1.45</td>
<td>1.39</td>
<td>1.25</td>
</tr>
</tbody>
</table>

**Overview of Significant Findings**

It was defined that problem solving **Performance** refers to achievement on posttests containing problems of the same structure and content as those used in instruction, while problem solving **Transfer** refers to achievement on posttests containing problems of the same structure as those used in instruction, but different content. **Transfer** posttests for this study involved word sequences rather than the graphical Moppet figures used in instruction. **Reflectivity** was defined as a numerical rating of the degree of metacognitive and reflective thought evident in students thinking journal responses.

**Research question 1:** Will students who write thinking journals perform better on tests of problem solving **Performance** than students who do not write thinking journals?

1. There were no significant differences in scores on tests of problem solving performance, either when averaged or for individual occasions, between students who wrote thinking journals, and students who did not.
2. Average Performance scores for females were higher than for males. However, there were no differences in the effect of journal writing on measured problem solving Performance between male and female students, except for a single occasion, week 5. Performance scores for females in the Journal group for week 5 were higher than for females in the Control group, while Performance scores for males in the Journal group were lower than for males in the Control group.

Research question 2: Will students who write thinking journals perform better on tests of problem solving Transfer than students who do not write thinking journals?

1. There was no significant difference between the average Transfer scores of students who wrote thinking journals, and students who did not.

2. For the first three testing occasions, there were no significant differences between Transfer scores of students who wrote thinking journals, and students who did not.

3. In the final session, students who wrote thinking journals scored significantly higher on tests of problem solving Transfer than students who did not write thinking journals. Transfer scores were higher for both male and female journal-writing students.

4. The average Transfer scores for females were higher than for males. Weekly Transfer scores for females were
higher than for males in weeks two and five, but there were no significant differences for weeks one and three.

5. There were no differences in the effect of journal writing on problem solving Transfer between male and female students.

**Research question 3:** Is there a relationship between the degree of metacognitive and reflective thinking in student journal responses, and scores on tests of problem solving Performance?

1. Students who had higher average Reflectivity scores also had higher average problem solving Performance scores.

2. When Reflectivity scores for both genders are combined, individual weekly Reflectivity scores are not strong predictors of scores for the Performance posttests administered the same day.

3. When Reflectivity scores for males and females are analyzed separately, individual weekly Reflectivity ratings of males were highly correlated with scores of Performance, while those of females were not.

**Research question 4:** Is there a relationship between the degree of metacognitive and reflective thinking in student journal responses, and scores on tests of problem solving Transfer?

1. Students who had higher average Reflectivity scores also had higher average problem solving Transfer scores.
2. When Reflectivity scores for both genders are combined, individual weekly Reflectivity scores are not strong predictors of scores for the Transfer posttests administered the same day.

3. When Reflectivity scores for males and females are analyzed separately, the Reflectivity ratings of females are not correlated with Transfer for any testing occasion, nor are they for the mean of all testing occasions.

4. When Reflectivity scores for males and females are analyzed separately, the Reflectivity ratings of males were not significantly correlated with scores of Transfer for the first three testing occasions. However, Reflectivity ratings for males were highly and positively correlated with Transfer for the final occasion (week 5).

Journal writing: Several descriptive statistics documented student progress in journal writing activities:

1. The average number of words per journal increased from 87.3 in week 2 to 92.9 in week 5.

2. The length of individual responses increased by almost fifty percent, from 12.5 words per response in week 1, to 18.6 words per response in week 5.

3. Mean student Reflectivity ratings increased from .94 in week one, to 1.25 in week 5. The weekly increases in reflectivity for weeks 1 through 3 were significant at the .01 level.
4. Reflectivity ratings for males and females were almost identical for all four weekly measures.

Discussion

1. Problem solving ability: Performance and Transfer

This study provides no evidence that writing thinking journals improves problem solving Performance. This is inconsistent with several studies that reported that children demonstrate higher levels of thinking and problem solving ability when they verbalize their activities and their thinking (Dickson, 1982; Olson & Ives, 1983, cited in Dickson, 1985; Metwali, 1979, cited in Silver & Thompson, 1984).

The repeat of lesson two content for lesson five could have influenced the outcome for week five. This complicates the interpretation of the lack of significant difference for Performance. Some degree of practice effect due to previous exposure to the content of the lesson is likely, although it was equal in both groups. Since the study did not examine gains scores, there is no equivalent of "pretest-posttest interaction." The effect of prior exposure to the content on outcomes is unknown - it could have decreased or increased the probability of significant difference; it can only be concluded that the study revealed no significant advantage for journal writing students when Performance scores were compared.
The results regarding Transfer were mixed. It cannot be concluded that journal writing improves the ability to transfer skills to other kinds of problems. The study does provide evidence, however, that under some conditions, writing in thinking journals can improve transfer. This conditionally confirms a report that written reflection of problem solving activities following practice with a computer problem resulted in clearer development of problem-solving procedures (Levin, et al., 1986). The results are also consistent with the observation that studies of thinking skills instruction lacking instruction in metacognition fail to report generalized transfer, while studies that report transfer include instruction in metacognitive skills (Belmont & Butterfield, 1977; Meichenbaum, 1985). It also supports research indicating that students can successfully transfer problem solving skills from computer problems to other kinds of problems (Swan, 1989; Melnik 1986).

The "conditional" effectiveness of journal writing could mean several things, given the information that can be derived from this brief series of treatments. First, it was observed that students react differently to different instructional problems; this suggests that a strategy that is effective for one kind of problem might be ineffective for another. There were many consistent observations across the four weekly Moptown Hotel lessons, including persistent
gender differences between means for both Performance and Transfer measures. However, even this persistent difference disappeared during the third week of instruction, as the means of all four gender/method groups converged to a single point. An analysis of the differences between lesson three and the other lessons revealed that the other lessons require that the student actively produce a response by combining factors until the right Moppet or word has been created. Lesson three, on the other hand, requires only that the student look at a set of figures and choose one, a much less complex and more passive task. It is possible that this kind of difference in task complexity resulted not only in gender differences disappearing, but in a loss of journal effectiveness as well. Future research should be designed to determine which kinds of problem tasks benefit from reflective writing, and which kinds do not.

Secondly, a common sense assumption is that some students need specific instruction and practice in writing thinking journals before any benefits occur. The experimental evidence from this study supports this "practice effect" assumption to some extent. A significant difference occurred in only one of the four measures of Transfer - the last one. Unfortunately, no clear trend can be documented to further support this hypothesis; a uniformly increasing advantage for journal-writing students did not develop (or did not have time to develop). However,
the assumption is supported by the fact that students did improve their journal writing proficiency, in words per response, and in measured Reflectivity. Also, descriptive data gathered while observing students working in journals suggests that substantial practice was required, that early journal formats were ineffective, and that students were still progressing at the end of the study. Several observations made during the first week can be contrasted with those of the final weeks of instruction:

Week 1:
1. "Too much interaction - need instructions to reduce collaboration."
2. "Many 'what does this mean?' questions... need some explanation."
3. "Some students needed clarification regarding the difference between problem and solution."
4. "Some students did not know how to respond to the 'feelings' question... several didn't understand the 'what would you do differently' question, and still could not answer it when they did."
5. "The child who came in late wrote only one response (he writes with much difficulty). Another child displayed extreme anxiety over the questions..."

Week 3:
1. "Very few problems understanding questions. Again, I interacted with children as they wrote to get them to
clarify their thoughts, although I did not observe as many problems. I did see some evidence of students just 'parroting' what we said in class...

2. "I am beginning to see that the journaling activity is not having the effect that I expected. I attribute this to... the [fact that the] children began at a 'tacit' level of metacognitive ability"

Week 5:

1. "Several students had problems with the wording on the two 'rule' questions. The emphasis of the wording does not match the way I presented the concept... no one asked questions or seemed stumped on the other 'A' questions"

2. "I noticed several children having trouble remembering (again) the second part of the rule (getting the new traits from the last Moppet)."

It is also likely that since lesson five was a second treatment of the material for lesson two, there was some degree of practice effect due to previous exposure to the content of the lesson and posttest formats (they were not the same problems). This alters the conditions of the experiment; unlike other occasions, students were familiar with the lesson content (it is noted, however, that the condition does not threaten the reliability of results,
since both experimental and control students were exposed equally to both lessons).

There is also the question of whether journal writing activities for lesson two influenced the outcome of lesson five, due to prior journal-writing for the same content, and beyond that expected due to additional journal-writing practice. There was evidence in classroom discussions, practice sessions, and posttest answers that some students remembered what they had done in the previous session (lesson two), since they were discussing and applying strategies that had not been discussed in lesson five. It cannot be determined, however, whether experimental group students remembered or comprehended these strategies better than control group students, and if so, whether it was because they had written about them in the previous thinking journal treatment. However, it would be expected that, if journal writing in week two had influenced Transfer in week five, Reflectivity in week two would be a good predictor of Reflectivity in week five. The correlation matrix used to produce table 17 reveals no such conclusion; it indicates that Reflectivity scores for week five were highly correlated with scores of Transfer for week five, while Reflectivity scores for week two were not significantly correlated with scores of Transfer for week five, suggesting either that a journal practice effect did not occur, or that it was only a partial influence on the outcome.
There are many alternate explanations for the sudden occurrence of significant differences in the final treatment, unrelated to any possible "practice effects." There were several differences in the final weekly treatment. First, the posttests for final lesson had been improved, based on prior concerns of reliability and observed difficulties with the instruments. In fact, the measured reliability (internal consistency) of the posttests improved substantially over the course of the treatments; it is possible, then, that the significant results for the final lesson were influenced by more reliable measurement. Also, the journal format for the final lesson was much more direct and explicit in its questions than previous journals (refer to Appendix B). It had become apparent that to be effective, the questions must be focused toward the specific thinking desired. Therefore, it is possible that the journals were more effective because they had changed, and not because the students had developed.

Secondly, there were two deliberate changes made to the original problem set used in lesson two. It had been noted that the transfer "word" problems provided with Moptown Hotel and used for lesson two posttests could not be successfully solved by merely applying the strategies used in class - the transfer word problems were in fact supersets of the computer problems, not isomorphs, requiring creative problem solving to produce or adapt rules to the new
situations. For lesson five, the problems were redesigned so students could apply the same strategies and procedures used to solve the equivalent Moppet problems practiced in class.

A second modification was that the cosmetic and mechanical features of the worksheet intentionally mimicked that of the Performance worksheet containing the Moppet problems, providing a visual cue to the student that the two problem sets were similar. This is a form of scaffolding, intended to support students in the early process of learning to transfer skills; the cues would have eventually been removed as students learned to make the association on their own. Previous posttests had not provided this support - they had, for instance, presented Moppet problems in a horizontal format and word problems in a vertical format. It had been observed that on previous occasions, students tended to react to differences in surface detail between any two problems, including differences between on-line computer problems and their supposed Performance posttest equivalents, as though there was simply no relationship between the two. The following observation was recorded after one of the instructional sessions:

I have seen evidence that children do not carry strategies and rules from one task to another, because they do not relate the tasks to one another. To some extent, this is even true of the
"performance" worksheets, even though they are based on the same problem. Some children have successfully performed the computer task, then explained that "they just don't get it" when given the same kind of problem on paper, reacting as though they had never seen the problem before...

When the problem is different, as in the transfer worksheets, I have seen no evidence that children see the problem as related to the computer task in any way - they approach it as a brand new problem.

This observation reflects one of the more basic tenets of general problem solving theory, that the problem representations of young and/or novice problem solvers are based on surface features, syntax, or spatial proximity, and as such, are not very effective (Duffield, 1991). It also supports the principle that teachers should explicitly point out the similarities among different contexts to encourage transfer of problem solving skills. It is possible, then, that when students were cued to the similarity between problems, those who had written journal exercises had additional knowledge at their disposal for solving the problems. In essence, the test forms were more effective at drawing potential transfer ability from the students.

The interpretation of an isolated significant effect in an experiment in which eight occasions offered the
possibility of significant differences is problematic. It is possible that several, or all of these factors, in some combination, influenced the significant findings for the Transfer test for lesson five. However, the data gathered in this study can support no single conclusion over another, nor can it absolutely exclude the influence of other confounding variables, given its stated limitations of sample selection and control.

2. Reflective writing and its relation to problem solving ability

   Journal writing skills improved as students gained experience with the specific kinds of journals used in this study, and as the formats of the journals were refined. Evidence to this effect included longer responses, and increases in the specificity of answers, indicated by an increase in measured Reflectivity. This supports another researcher's observation that "as students continue to write in their journals, their entries become longer and their sentence structures become more complex" (Danielson, 1988, p. 24), and it reflects a qualitative analysis of student journal writing that indicated that students' metacognitive abilities were enhanced through journal writing activities (Linn, 1987). However, it was also noted above that students remained quite limited in their reflective writing ability. This observation is consistent with research
suggesting that young children are limited in their metacognitive knowledge (Flavell, 1979; Eylon & Linn, 1988).

The relationship between Reflectivity and problem solving ability was analyzed because it could be argued that if reflective journal writing is responsible for improved problem solving ability, then better problem solvers should also be more reflective journal writers. If there were no such relationship, then it would be reasonable to question why journal writing should have any positive impact on problem solving ability. Such a situation would have questioned the meaning of any significant findings regarding journal writing and its effect on problem solving ability. However, the expected relationship exists. Significant relationships between mean Reflectivity and mean Performance, and between mean Reflectivity and mean Transfer were observed.

As expected, then, students who demonstrate greater Reflectivity tend to be better at both problem solving performance and transfer. However, it appears that the strength of this relationship is almost entirely due to male problem solvers. Although males scored lower on tests of problem solving skills, there is a stronger relationship between their reflectivity and their problem solving ability than for females. Also, it appears that reflectivity averaged over a period of time is a better predictor of a student's problems solving ability than individual
reflectivity scores. Generally, individual scores are not reliable predictors of individual performances.

This experiment was not designed to test causality regarding reflectivity and its impact on problem solving performance and transfer. However, thinking journals were designed to evoke reflective behavior, the control journals were specifically designed to not evoke such behavior, and it was observed that in the final observation, students in the experimental group outperformed the control group. It was also observed that, on average, students (particularly male students) who were more reflective were also better problem solvers. Thus, it is possible to infer, but not conclude, that, on average, writing that produces greater degrees of metacognitive and reflective thinking is more helpful to problem solvers than writing that does not.

However, it was also noted that when individual occasions were tested, there were generally no significant relationships between Reflectivity and Transfer. It is possible to ignore this result - individual tests can be less powerful, producing insignificant results, while combined tests of the same variables produce significance (Stevens, 1986, p. 114). However, it is of interest that the only significant relationship between Reflectivity and problem solving ability corresponded with with the only significant difference among means. There are probably many possible explanations, not the least of which is that the
alignment is a chance occurrence. The following discussion is speculative, and is offered as a possible explanation, not a conclusion.

If the inverse of the relationship is examined, e.g., for most occasions, reflectivity is not associated with success in problem solving, it could be hypothesized that for earlier problems, for any of a variety of reasons, students simply did not apply what they knew - knowledge they had demonstrated by their responses in their thinking journals. There is precedence for this hypothesis in the literature on problem solving. It has been found that ineffective problem solvers often do not apply relevant information to problems, even when they have the needed information (Picus, et al., 1983, p.7; Lockhead, 1981, p. 68; Silver & Thompson, 1984, p. 535-537). This can be due to several factors. First, novice problem-solvers are weak in discriminating the salient features of a problem representation (Eylon & Linn, 1988, p. 277), and may therefore fail to recognize the ordinary cues that help more experienced problem solvers retrieve necessary information. Also, novices are often deficient in the conditional knowledge that provides a selective retrieval mechanism; learners use conditional knowledge to determine when certain components of their declarative and procedural knowledge should be applied to a given task (Lippert, 1988, p. 1-2). Finally, when students are provided with multiple ways of
representing procedural knowledge, as was the case in the Transfer worksheets for this study, they have been found to have developed separate, disconnected systems of knowledge (Lockhead, 1981); students may fail to apply their knowledge because it is not connected with the representation at hand.

Of course, there is nothing in the quantitative data to support this hypothesis, but one recorded classroom observation offers evidence that students did not always apply their knowledge, and documents several more reasons for their doing so:

1. I see strong evidence that children sit and listen to discussions about strategies and rules, then do not use them when they work at the computer or when they do worksheets [posttests]... I have tried to carry the discussion into the activities, in my monitoring and guidance, and children often can tell me the strategies and rules they should be using, but they often continue to guess.

2. Some of the problems can be worked successfully at an "intuitive" level. If a child can work a problem without overtly applying rules or strategies, he or she will continue to do that.

It is certainly feasible, in light of these and other similar observations, and the literature on novice problem
solvers, that for the earlier tests, students simply did not apply what they knew at a verbal level, such that their success could not be reliably predicted from the information written in their journals. However, in the final occasion, faced with a difficult transfer problem, one containing visual cues that would evoke a connection with the Moppet problems, it is possible that those who were aware of the processes that had been practiced, and who had just written about them in their journals, applied the information, resulting in a high correlation between reflectivity and success, and perhaps, a significant advantage for journal writing students.

3. Gender differences

Female students generally outperformed males students in problem solving ability (both Performance and Transfer). However, it is usually recognized that the rate of mental development differs between females and males in this age group, so the result is of limited application, except that it supports research by Johnson (1984) that refuted previous assumptions that males are generally better problem solvers than females, assumptions largely attributed to work by Donald Taylor of Stanford University. Gender differences noted in this study also underline the importance of controlling for gender differences in future studies of problem solving ability.
Why a strong relationship between Reflectivity and problem solving ability never existed for female students is a question that cannot be answered from the data in this study. However, a series of experiments by Johnson (1984) found a male advantage in problem solving, limited, however, to word problems in mathematics, and not extending to other domains. The author hypothesized that males excel in the ability to translate verbal representations into ones that can be attacked analytically and mathematically, and that this may be related to sex differences in mathematics aptitude and spatial ability. The strong association observed in this study between male reflective writing and problem solving ability, and the lack of such a relationship for females, is perhaps related to this hypothesis.
CHAPTER V

SUMMARY AND CONCLUSIONS

This study examined the effects of instruction using generic, or content-free problem solving software, software that provides practice in general problem solving skills, rather than those specific to any particular subject. The use of such software is often assumed to improve students' problem solving ability, however, such software can be ineffective unless supported by other teaching or learning activities. Research on teaching, and general problem solving theory, suggests that children are more likely to learn problem solving skills, and more likely to transfer those skills to other kinds of problems, if they write or talk about their problem solving activities, thereby learning to reflect on their own problem solving behavior. Since journal writing provides students with an opportunity to verbalize, and to interact with an adult who can provide support and guidance in the endeavor, and since specially designed thinking journal activities have been attributed with encouraging reflective thinking, it was hypothesized that instruction using thinking journal activities with problem solving software would lead to better problem
solving performance, and greater transfer of problem solving skills, compared with problem solving software used alone.

The study compared scores on separate tests of problem solving performance and transfer, for students who received software instruction supported by student thinking journal activities, with scores of students who received the same instruction, but who performed non-reflective writing activities. The subjects of the study were sixty-three public school fourth grade students, in three classes of between nineteen and twenty-four students each. Classroom instruction was provided by the researcher.

Problem solving performance was defined as achievement on posttests containing problems of the same structure and content as those used in instruction. Problem solving transfer was defined as achievement on posttests containing problems of the same structure as those used in instruction, but different content. Transfer posttests involved word sequences rather than the graphical Moppet figures used in instruction.

The study also examined the relationship between the degree of metacognitive and reflective thought displayed in students' journal entries, and their measured problem solving ability for both performance and transfer. Reflectivity was defined as a numerical rating of the degree of metacognitive and reflective thought evident in students thinking journal responses.
It was hypothesized that students who wrote thinking journals would perform better on tests of problem solving performance than students who did not write thinking journals. Results of multivariate analysis of variance (MANOVA) indicate that there were no significant differences in scores on tests of problem solving performance, when averaged or for individual occasions, between students who wrote thinking journals, and students who did not. Although female students performed better than males in tests of problem solving performance, journal writing was generally no more likely to improve problem solving performance for one gender than for the other.

It was also hypothesized that students who wrote thinking journals would perform better on tests of problem solving transfer than students who did not write thinking journals. Results indicate that there was no significant difference between the average transfer scores of students who wrote thinking journals, and students who did not, nor were there differences for the first three testing occasions when analyzed separately. In the final session, however, students who wrote thinking journals scored significantly higher on tests of problem solving transfer than students who did not write thinking journals (p < .01); transfer scores were higher for both male and female journal-writing students.
On average, female students performed better than males in tests of problem solving transfer. However, journal writing was as likely to improve problem solving transfer for one gender as for the other.

Furthermore, it was hypothesized that there would be a positive, significant relationship between the degree of metacognitive and reflective thinking in student journal responses, and scores on tests of problem solving performance and problem solving transfer. Lack of such a relationship would have questioned the influence of reflective thinking journal writing upon the outcome of the experiments. However, results of regression analyses support the hypothesis, finding that students who had higher average reflectivity scores also had higher average problem solving performance and transfer scores (for both measures, p < .05).

Generally, however, individual weekly reflectivity scores were not found to be reliable predictors of scores for either problem solving performance or transfer. The single exception was that in the final session, it was found that students who had higher reflectivity scores also had higher problem solving transfer scores (p < .05).

It was also found through post hoc analyses of correlation coefficients, separated by gender, that the significant relationship between reflectivity and scores of
problem solving ability was largely attributed to males and not females.

Other statistics indicate that during the course of the four week treatment, students who wrote thinking journals progressed in their journal writing abilities. The average number of words per journal increased, and the length of individual responses increased by almost fifty percent, from 12.5 words per response in week 1, to 18.6 words per response in week 5. Also, student reflectivity ratings increased significantly during each of the first three weeks (p < .01); they ranged from .94 in week one, to 1.25 in week 5, on a scale of 0 through 4. Reflectivity ratings for males and females were almost identical for all four weekly measures.

Conclusions

This study asked two basic questions. First, is journal writing an effective teaching strategy when used to support learning with generic problem solving software? Secondly, is the metacognitive and reflective thinking encouraged in journal writing an important factor in successful problem solving behavior?

1. The impact of journal writing on problem solving ability

This study produced no results that would indicate that, generally, journal writing helps students learn to solve problems. It was concluded, however, that under the right
conditions, and for the right kinds of problems, thinking journal writing of the kind used in this study might help students clarify their understanding of their own thinking processes, resulting in improved ability to apply problem solving skills to another problem of the same type. This conclusion must be tempered with the knowledge that when averaged over four testing occasions, neither tests of problem solving performance nor tests of transfer revealed significant advantages for journal writing students. Also, the meaning of "the right conditions" and "the right problems" must be clarified.

It is likely that thinking journals are effective for some kinds of problem situations, but not others. The study found a significant advantage for journal writing students in a single test of problem solving transfer, in which students applied skills learned with one problem representation to a problem set using another representation. There was no such advantage when students worked problems like those practiced in class. It cannot be concluded that such an advantage for this latter case could not have occurred. It can be concluded, however, that, under some conditions, journal writing can help children transfer skills to other kinds of problems.

The second observation was that students reacted differently in problems that required active production of solutions, than in problems that required only passive
selection of answers. In the latter type of problem, students in all groups performed the same, regardless of the strategies employed. This suggests that some kinds of problems are better suited for journal writing activities than others.

The results of the study are inconclusive regarding the "conditions" under which journal writing can be effective. It can be stated with certainty, however, that many students were not ready to benefit from journal writing activities when the study began. This implies, and informal observations made by the researcher suggest, that journal writing is ineffective unless (1) effort and time is devoted to instructing students in journal writing procedures and the terminology and concepts required, and (2) journal formats and questions are designed to provide the scaffolding students need as they gradually learn these concepts and skills.

2. The contribution of reflective thinking

Exploratory studies often produce serendipitous results, often more promising or interesting than those originally envisioned (Kirk, 1982). This study was designed to measure the effects of only one independent variable, journal writing. However, the inclusion of gender as a factor, due to unforeseen influences on the dependent variable, revealed that while the journal writing experience had the same
effect on problem solving behavior for both gender groups, it is possible that males apply metacognitive awareness gained through journal writing differently than females. Males who wrote more reflective comments in their journals also performed better on both kinds of problem solving tests, while for females, the metacognitive thinking displayed in journals (which was as high as for males) was unrelated to their performance. Rather than providing an answer, this poses two important questions. First, since female problem solving behavior also improved due to journal writing activities, what features of journal writing, other than the kind of reflective thought measured in this study, produced the effect? Secondly, should practitioners and researchers view journal writing activities differently for male and female elementary school children? If so, how?

The data collected for this experiment suggest that journal writing can be used to provide instruction at a personal and individualized level that both increases the child's metacognitive awareness and increases the child's ability to verbalize procedural matters. Individualized instruction is an important facet of thinking skills instruction, especially for younger students who are more easily distracted by irrelevant factors, more likely to exhibit problems with cognitive load, and who some researchers believe are not at a sufficient level of cognitive development to warrant thinking skills instruction.
at all. Individualized instruction is difficult to manage within the realities of public school instruction. Journal writing, therefore, can be an important tool for educators who must somehow teach children to "learn how to learn."

As is the case in almost any experimental study, there are limitations to the application of these conclusions. The sample for this study was not randomly selected, therefore, there is no claim that generalizations beyond the sample population are valid, although there is also no specific information that would suggest that the subjects were atypical. Also, measures of problem solving ability were paper-and-pencil exercises that followed very closely the form of the instructional software and of materials used in classroom discussion, and which were administered immediately following instruction. This might or might not reflect actual classroom practice, and certainly does not reflect "real world" problem solving, so one should be cautious in stating that the strategies used "increased problem solving ability."

**Implications for Problem Solving Instruction**

Journal writing can be a very important tool for supporting the development of children's problem solving skills, because it is very difficult to teach thinking and problem solving skills to large groups of young students. The instructional treatment for this study involved
approximately equal proportions of classroom discussion and computer practice, but it was repeatedly observed that classroom instruction must be constantly refined due to the demands of thinking skills instruction, and that even when this was done, most students needed additional individual guidance in order to begin to assimilate and apply the information discussed. The power of journal writing in assisting the teacher in this task is threefold. First, journals provide teachers direct and dynamic communication with the student, allowing access to the student's understanding of the concepts discussed and practiced in class, rather than relying on less descriptive and static measures such as posttests or quizzes (Kuhrt & Farris, 1990; Bode, 1989). Secondly, journals help teachers make subtle adjustments in their teaching (Sanders, 1985); this was evident in the changes made to lesson plans and materials as the current study progressed. Finally, journal writing allows the teacher to provide the personalized attention required of effective problem solving instruction, beyond that possible during busy practice sessions or classroom discussions. "It is through dialog journal writing that education can be personalized. From a teacher's perspective, that is very empowering" (Bode, 1989, p. 570).

Several specific recommendations for the use of journal writing in problem solving instruction can be made on the
basis of both quantitative and descriptive information gathered through this study.

1. Since journal writing produced significant results in only a single lesson, and since there were other indicators that students reacted differently to different kinds of problems, teachers should be aware that journal writing might be helpful in some problem situations, and not in others. The theoretical application of this principle is that teachers should adopt strategies that produce results for specific instructional situations, rather than globally adopting a strategy like journal writing across many situations. This does not necessarily match practice, however, since many schools adopt an instructional tool like journal writing as an across-the-board philosophy, and it is difficult to say that doing so, in the aggregate, is necessarily poor practice. The practical application of this observation, then, is that teachers should be watchful for situations in which journal writing is unproductive, and students' journals responses themselves might be a source of the information needed to make this decision.

2. A corollary to this principle is that journal writing might be a more powerful tool in supporting the transfer of skills, rather than for improving performance in the original problem solving situation. However, such a conclusion would involve overgeneralizing the results of
this study, so the question remains to be addressed by future research.

3. It is apparent that young students need instruction and extended practice before journal writing becomes productive, and before any benefits of journal writing can accrue. Therefore, it is essential that the concepts and terminology needed to understand thinking journal questions be explicitly discussed with students before the journals are used in instruction.

4. The process of familiarizing students with the problem solving concepts needed for effective thinking journal writing can be made more efficient through the design of journal questions that include scaffolding devices or cues. In this study, the journal format for the final lesson was much more direct and explicit in its questions than prior journals, and this might have contributed to higher scores for journal writing students. Journals should initially contain questions that lead the student in the direction desired. For example, the final journal posed the question "What do you do when the Moppet you put in the second box is exactly like the last Moppet? Explain. Which trait do you change next?"

However, a successful problem solver must not only know the processes and strategies required to solve a problem. It is also important that the problem solver know the questions appropriate for each step in the process. An
important metacognitive strategy is self-questioning; many strategies are aimed at developing children's knowledge of different kinds of questions (Hunkins, 1989, p. 9, 19). Thinking journals intended to foster metacognition should ultimately be directed toward the child's internalization of a set of questions that will guide future independent learning. Therefore, as students become familiar with problem solving behavior, scaffolding devices or cues should be gradually removed, so that the student learns to mentally ask the specific question in response to more general inquiries, like "How did you solve the problem?"

5. Problem solving software often includes non-computer, paper-and-pencil exercises to help students learn to transfer problem solving skills. The visual design and mechanics of these worksheets should also provide a different kind of scaffolding, to help beginning students recognize that there is a relationship with previous problems. The "word sequence" posttests used as measures of transfer were visually similar to those of the equivalent Moppet worksheets, providing visual cues to the student that there were similarities between the two problem sets. This kind of scaffolding should be also be gradually removed so that students learn to recognize similarities on their own.
Suggestions for Further Research

1. Future research of problem solving skills for elementary aged students should block, factor, or otherwise control for gender differences. Since it is not known if the difference in the relationship between reflectivity and problem solving behavior between young males and females is a transitory or permanent effect, research of any age group, and especially of journal writing behavior, should both control for and document differences between male and female subjects.

2. Future studies should address the question of what features or benefits of journal writing, other than the kind of reflective thinking measured in this study, can contribute to more successful problem solving behavior. Are these other effects, perhaps other important thinking patterns unidentified in this study, also helpful to problem-solving students?

3. Prior research had produced several metrics of reflective thought, however, these were found ineffective because they had been designed for older and more experienced students. Future studies of young children should continue to develop, adapt and refine formal measures of written and spoken metacognitive and reflective thinking, so that knowledge of why and how verbalization helps clarify thought can be examined for younger populations.
4. Finally, actual classroom journal writing does not lend itself to "positivist" experimental research. "This apparently uniform, observable 'treatment' of having students write every day in a dialog journal to the same teachers, is in fact never a uniform or singular treatment for the acquisition of reasoning or anything else." (Staton, 1984, p. 501). Therefore, it is suggested that qualitative research methods be used to further study the questions posed above, particularly those related to the way in which students apply their metacognitive knowledge.
APPENDIX A

LEARNING TASK ANALYSIS
Intellectual & Cognitive Skills Learning Task Analysis
Moptown Hotel Activities 1 - 5

1. Who's Next Door?
   - Generate an analogy using Moppets in a 2x2 array
     - Classifies "analogy" by giving examples defined concept
     - Adopt a strategy of changing the same trait cognitive strategy
     - Classifies "array" by providing a definition defined concept
     - Adopt a strategy of focusing on one feature at a time cognitive strategy
     - State the four Moppet traits verbal information
     - Identify red and blue Moppets concrete concept
     - Identify Bibbits and Gribbits concrete concept
     - Identify fat and thin Moppets concrete concept

2. Change Me
   - Generate one- and two-difference Moppet trains
     - problem solving / complex rule
   - Demonstrate making a series of Moppets that differ by a specified number of traits
     - rule
   - Adopt a strategy for controlling variables cognitive strategy
   - Adopt a strategy for planning ahead cognitive strategy
   - Demonstrate changing a Moppet by a specified number of traits rule

3. Spot Me
   - Demonstrate selecting 1 variable (trait) at a time rule
   - Adopt a strategy of hill-climbing / means -ends analysis cognitive strategy
   - Adopt a strategy of looking ahead cognitive strategy
   - Adopt a strategy for controlling variables cognitive strategy
   - Demonstrate the selection of the Moppet with the unique trait complex rule

4. Whose Birthday?
   - Demonstrate the selection of the secret Moppet problem solving / complex rule
   - Use negative information to infer positive information rule
   - Adopt a strategy of eliminating traits from a list of traits cognitive strategy
   - Demonstrate making a series of Moppets that differ by a specified number of traits
     - rule
   - Demonstrate changing a Moppet by a specified number of traits rule

5. Secret Pal
   - Demonstrate locating the Secret Pal in five or fewer guesses
     - problem solving / complex rule
   - Adopt a strategy of eliminating traits from a list of traits cognitive strategy
   - Adopt a strategy of controlling variables cognitive strategy
   - Adopt a strategy for planning ahead cognitive strategy
   - Demonstrate selecting the Secret Pal in five or fewer guesses
     - problem solving / complex rule

- *-.*, -.*-.*-.*-.*, -.*-.*-.*, -.*-.*-.*-.*, -.*-.*-.*-.*
APPENDIX B

THINKING JOURNAL QUESTIONS
Journal Activity 1 - *Who's Next Door?*

Name ____________________________________________

☐ Answer all of these questions. Use the blank paper provided in your journal. Number each answer.

1. What was the problem you had to solve?

2. Explain the strategy that you used to solve the problem.

3. How did you discover the strategy? Describe your thinking.

4. What would you do differently next time? Why?

5. How did you feel working through the problem? Why?

6. What do you do (what is your thinking) when you complete an analogy like "bee is to honey as spider is to ___"? Explain.

7. What are you learning about your thinking and feeling? Do you think it is important? Why?
Journal Activity 2 - Change Me

Name

☐ Answer all of these questions. Read each question carefully!

1. What was the problem you had to solve?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

2. What was the rule that you used to solve it? Explain.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

3. How did you discover this rule - in other words, what was your strategy? Describe your thinking.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
4. Did you find it important to look at the last Moppet? **Why?** What would happen if you didn't?

5. What other strategies could you use next time? **Why** would this be better?
Journal Activity 3 - Spot Me

Name __________________________

☐ Answer all of these questions. Read each question carefully!

1. What was the **problem** you had to solve?

   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________

2. What is the **rule** for spotting the Moppet with the greatest difference? **Explain.**

   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________

3. **How** did you discover this rule - in other words, what was your **strategy**? Describe your thinking.

   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________
   ____________________________________________

GO TO NEXT PAGE
4. What other strategies could you use next time? Why would this be better?

5. How did you feel working through the problem? Why?
Journal Activity 5 - Change Me

Name ____________________________________________

☐ Answer all of these questions. Read each question carefully. Be sure you notice the important words in **bold** print!

1. What was the **problem** you had to solve?

____________________________________________________________________________________________________

____________________________________________________________________________________________________

____________________________________________________________________________________________________

2. **Exactly how** do you solve a RULE 1 problem? **Explain all the steps** you performed. **How many** traits do you change? **Which** traits do you choose?

____________________________________________________________________________________________________

____________________________________________________________________________________________________

____________________________________________________________________________________________________

____________________________________________________________________________________________________

3. What do you **do** when the Moppet you put in the second box is **exactly** like the last Moppet? **Explain.** **Which trait** do you change next?

____________________________________________________________________________________________________

____________________________________________________________________________________________________

____________________________________________________________________________________________________

____________________________________________________________________________________________________

GO TO NEXT PAGE  ⚫
4. We discussed several strategies you could use to make Change Me problems easier. Which of these strategies did you use? Explain what they are and how you used them.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

5. Which of these strategies helped you the most? Which helped you the least? Explain why you think so.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
APPENDIX C

NON-REFLECTIVE JOURNAL QUESTIONS
Journal Activity 1 - Who's Next Door?

Name  ____________________________________________

☐ Answer all of these questions. Use the blank paper provided in your journal. Number each answer.

1. What is an array?

2. There are 16 different Moppets in all. If there were no red Moppets (only blue ones), how many different Moppets would there be? How did you find out? Explain.

3. Moppet traits are color, shape, size, and "Bibbit or Gribbit." If you designed a game like Moptown Hotel, what different traits could your computer characters have?

4. What are the rules that you follow when you use the computer lab? Why is it important to have rules when many people work together?

5. Is Moptown Hotel as much fun as your favorite arcade game? Why?
Journal Activity 2 - Change Me

Name

☐ Answer all of these questions. Read each question carefully!

1. What is the difference between a Bibbit and a Gribbit? Is there a rule you can use to remember which is which?

2. Does the music that Moptown Hotel plays when you get the answer right make it easier to play the game? Does it make it more fun? Why?

3. What does it mean when the computer displays the words "RULE: 1" or "RULE: 2" at the top of the screen? Explain.
4. What is an analogy? Explain.

5. How many different traits do Moppets have? List them.
Journal Activity 3 - Spot Me

Name ________________________________

☐ Answer all of these questions. Read each question carefully!

1. Moppets have four traits. What does the word trait mean to you? Write your definition in a complete sentence.

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

2. Humans have traits, too, for example, hair color. List at least four other human traits.

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

List two traits that both Moppets and humans have.

__________________________________________________________________________
__________________________________________________________________________

3. Do you like working with computers? Why? What kinds of things can you do well with computers?

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
4. Do you like to write in this journal? _____ Why? _____

5. Do you prefer writing in this kind of journal, with its own questions, or would you rather just write whatever you want in a spiral notebook journal? Why? Which is easier?

6. If you had a computer to use all the time in your classroom for your regular schoolwork, what would you use it for? If you could take it home with you, how would you use it?

Do you think these things will happen in the future? Why?
Journal Activity 5 - Change Me

Name ________________________________________

☐ Answer all of these questions. Read each question carefully. Be sure you notice the important words in **bold print**!

1. **Why** do you think that we have a journaling activity after our computer practice? **Explain.**

_________________________________________________________________________
_________________________________________________________________________
_________________________________________________________________________

2. Do these journaling activities help you **understand** your own own thinking? **How? Why?**

_________________________________________________________________________
_________________________________________________________________________
_________________________________________________________________________
_________________________________________________________________________

3. Do these journaling activities help you learn to **explain** your thinking to others? **How? Why?**

_________________________________________________________________________
_________________________________________________________________________
_________________________________________________________________________
_________________________________________________________________________

GO TO NEXT PAGE
4. Have you written journals in other classes? How are they different from the Moptown Hotel journals? Is this better or worse for you? Why?

5. How are they the same as the Moptown Hotel journals?
APPENDIX D

PERFORMANCE AND TRANSFER POSTTESTS
Who's Next? (Part 1)
First color the Moppets. Then use the Moppet cards to decide who lives next door. Write the number of the Moppet in the blank.

1. 

2. 

3. 

4. 

Name: ________________________________
Who's Next? (Part 2)
First color the Moppets. Then use the Moppet cards to decide who lives next door. Write the number of the Moppet in the blank.

1. 

2. 

3. 

4. 

Name: ________________________
Moppet Cards

Color all the Moppets (R = red, B = blue).
Keep this page. You will need it for several Moptown Hotel activities.

Red Moppets

1 2 3 4

5 6 7 8

Blue Moppets

9 10 11 12

13 14 15 16
What's Missing?

Thinking about Similarities

Fill in the missing word in each set.

1. big little  
   long 

2. bird nest  
   pig 

3. hand arm  
   foot 

4. sugar sweet  
   lemon 

5. night moon  
   day 

6. dog bone  
   bird
Moppet Trains
First color all the Moppets. Then use the Moppet cards to complete this puzzle. Write the number of the missing Moppet in each empty space.

RULE: 1
For each Moppet, change only one trait to find the next one. Be sure the last Moppet has only one trait different from the third Moppet.

1. RULE: 1

CONTINUE ON NEXT PAGE
Moppet Trains

3. RULE: 1

RULE: 2
For each Moppet, change two traits to find the next one. Be sure the last Moppet has two traits different from the third Moppet.

4. RULE: 2
Change a Letter

Change one letter of the word in A to make a new word in B. Then change one letter of the word in B so it has only one letter different from the words in B and D. Put it in C.

1. bed

2. dish

Try it again. You may also add or delete letters.
More Word Changes

In each column, turn the word in A into the word in D by making one change in the word for each step.

1.  
   **first**
   A
   B
   C
   D

2.  
   **fun**
   A
   B
   C
   D

   **last**
   **done**
Find the Unique Moppet

Who is Different? I

First color all the Moppets. Then write the letter of the Moppet who has a trait that no other Moppet has.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

1. Who is the most different in the row above? _______________________
2. Why? _______________________

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

3. Who is the most different in the row above? _______________________
4. Why? _______________________
Who is Different? III

First color all the Moppets. Then write the letter of the Moppet who has a trait that no other Moppet has.

1. Who is the most different in the row above? ________________
2. Why? ________________

3. Who is the most different in the row above? ________________
4. Why? ________________
Which is the Most Different?

Looking at Words

1. Which word is the most different? 
2. Why?

3. Which word is the most different? 
4. Why?
Looking at Sentences

Look at the sentences in each box and decide which is the most different. Tell why you think so.

1.
A. You will go to the store.
B. The store is where you will go.
C. Will you go to the store?

Which is the most different?  
Why?

2.
A. John can see the plane on the runway.
B. Can John see the plane in the sky?
C. The plane on the runway is what John can see.
D. Can John see the plane on the runway?

Which is the most different?  
Why?

3.
A. Is the ball in the can?
B. The ball is on the can.
C. Is the ball on the can?
D. The can is what the ball is on.

Which is the most different?  
Why?
Name: __________________________

**Moppet Trains**
First **color** all the Moppets. Then use the Moppet cards to complete this puzzle. Write the **number** of the missing Moppet in each empty space.

**RULE: 1**
For each Moppet, change only **one** trait to find the next one. Be sure the last Moppet has only one trait different from the third Moppet.

1. **RULE: 1**

   ![Diagram 1](image)

2. **RULE: 1**

   ![Diagram 2](image)

CONTINUE ON NEXT PAGE ⇩
# Moppet Trains

3. **RULE: 1**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. **RULE: 1**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Change a Letter

Change one letter of the word in A to make a new word in B. Then change one letter of the word in B so it has only one letter different from the words in B and D. Put it in C.

1. 

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>bed</td>
<td>____</td>
<td>____</td>
<td>sat</td>
</tr>
</tbody>
</table>

2. 

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>pop</td>
<td>____</td>
<td>____</td>
<td>sad</td>
</tr>
</tbody>
</table>

3. 

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>play</td>
<td>____</td>
<td>____</td>
<td>gram</td>
</tr>
</tbody>
</table>
APPENDIX E

TEMPLATES FOR RATING JOURNAL RESPONSES
**Coding Method for Thinking Journal Responses**

<table>
<thead>
<tr>
<th>Level</th>
<th>Term</th>
<th>Form(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tacit / Unaware</td>
<td>No answer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unclear or indeterminate phrase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes or no</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] don't know</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] don't know, it just came to me</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] don't remember</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I {just} did it</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I don't know how I feel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I can't think of anything (different than what I did)</td>
</tr>
<tr>
<td>1</td>
<td>Aware / Mentioning</td>
<td>The [rule / strategy / plan] was [label / inexplicit process]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] [label / inexplicit process]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] [label / inexplicit process] because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] did what [he / you] [said / taught]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] did what [he / you] [said / taught] and it worked well</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The problem was label</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The problem was inexplicit problem statement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I feel feeling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I feel feeling, but I don't know why</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I feel feeling because inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Nothing&quot; (e.g., I would change nothing)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I wouldn't do anything differently because inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I [would / would not] inexplicit process because inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If inexplicit condition then (a conclusion)</td>
</tr>
<tr>
<td>2</td>
<td>Explicit / Strategic</td>
<td>The [rule / strategy / plan] was explicit process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] explicit process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] label which is explicit process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] explicit process because inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[I / we] [used / did] label which is explicit process because inexplicit reason</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The problem was explicit problem statement</td>
</tr>
</tbody>
</table>
I feel explicit statement of feelings
I feel feeling because explicit reason
I [would / would not] explicit process because inexplicit reason
I would not guess {next time} because inexplicit reason
If explicit condition then (a conclusion)
It is important to {do} reflective process because inexplicit reason

3 Reflective

[I / we] [used / did] [label / inexplicit process / explicit process] because explicit reason
[label / inexplicit process / explicit process] is [good / bad] because explicit reason
[label / inexplicit process / explicit process] was [better / worse] than another process because explicit reason
I wouldn't do anything differently because explicit reason
I would not guess {next time} because explicit reason
I [would / would not] [label / inexplicit process / explicit process] because explicit reason
reflective process is [good / bad] because explicit reason
it is important to {do} reflective process because explicit reason

labels - (1) names of strategies and rules: "hill climbing", "looking ahead", "being organized", "my rule", "my strategy", ...
(2) names of concepts: "analogies", "arrays", "strategy", "rule"
(2) names of Moptown Hotel games, i.e., "who's next door", "secret pal", "change me"

explicit reasons - specific rationale or justification for a particular process, action, or feeling

inexplicit reasons - because it worked/didn't work, because I got the right/wrong answer, because I thought so, because I want to do better, because it's fun, ...

explicit processes - specific actions, steps, processes, ways of thinking

inexplicit processes - read the question, looked at the figures, filled in the blank, guessed, used instinct, answered the question, looked at the colors, shape, did problems, "though of what to do", looked closer, worked carefully, did what you said, ...
reflective processes - any reference to the processes of evaluating, comparing, or justifying problem solving strategies or rules

feelings - terms like good, bad, angry, frustrated, bored, awesome, confident, ...

explicit statements of feeling - an explicit description of specific kinds of feelings

explicit conditions - an explicitly described set of circumstances or conditions that must be true for some conclusion to be made

inexplicit conditions - if "it looked right", if "it just 'popped out' at me", ...

explicit problem statement - identification of all or part of the objectives of the problem

inexplicit problem statement - to answer the question, to fill in the blank, to fill in the missing Moppet, to find out what's missing, to choose the right answer, to find out what traits go in the box
APPENDIX F

COMPUTER PROGRAM FOR SCORING LESSON 2 AND 5 POSTTESTS
1 REM Moppet Train Check Program
2 REM
3 REM red = &h1000, blue = &h0000
4 REM tall = &h0100, short = &h0000
5 REM fat = &h0010, thin = &h0000
6 REM gribbit = &h0001, bibbit = &h0000
7 REM
8 CLS : INPUT "RULE (Type 1 or 2)"); ru
9 IF ru < 1 OR ru > 2 THEN GOTO 8
10 DIM mo(16)
100 DATA &h0011, &h0001, &h0110, &h0111, &h0010, &h0100, &h0101, &h0000
101 DATA &h1011, &h1001, &h1110, &h1111, &h1010, &h1100, &h1101, &h1000
200 FOR i = 1 TO 16
210 READ mo(i)
220 NEXT i
500 CLS : PRINT "RULE : "); ru: PRINT
1000 INPUT "First Moppet:"); m(1)
1005 IF m(1) < 1 OR m(1) > 16 THEN SOUND 500, 4: GOTO 1000
1010 INPUT "Last Moppet:"); m(4)
1015 IF m(4) < 1 OR m(4) > 16 THEN SOUND 500, 4: GOTO 1010
1020 PRINT
1100 INPUT "Second Moppet:"); m(2)
1105 IF m(2) < 1 OR m(2) > 16 THEN SOUND 500, 4: GOTO 1100
1110 INPUT "Third Moppet:"); m(3)
1115 IF m(3) < 1 OR m(3) > 16 THEN SOUND 500, 4: GOTO 1110
1120 cc = 0
1130 FOR i = 1 TO 3
1140 ts = mo(m(i)) XOR mo(m(i + 1))
1150 DC = 0
1160 IF (ts AND &h1) = &h1 THEN DC = DC + 1
1170 IF (ts AND &h10) = &h10 THEN DC = DC + 1
1180 IF (ts AND &h100) = &h100 THEN DC = DC + 1
1190 IF (ts AND &h1000) = &h1000 THEN DC = DC + 1
1200 IF NOT (DC = ru) THEN SOUND 1000, 2: PRINT "There are "; DC; " differences between Moppets "; i; "; " and "; i + 1: GOTO 1220
1210 PRINT "Moppets "; i; " and "; i + 1; " are CORRECT.": cc = cc + 1
1220 NEXT i
1230 IF cc = 3 THEN PRINT "PROBLEM SOLVED CORRECTLY!"
1235 PRINT : PRINT "Score = "); cc
1240 PRINT : GOTO 1100
BIBLIOGRAPHY


