EFFECTS OF A COMPUTER-ASSISTED TUTORING SYSTEM ON ACQUISITION, MAINTENANCE, AND GENERALIZATION OF TIME-TELLING SKILLS OF ELEMENTARY SCHOOL STUDENTS WITH BEHAVIOR DISORDERS

DISSERTATION

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

By
Yao Ma, M. ED.

The Ohio State University
2006

Dissertation Committee:
Professor Ralph Gardner III, Adviser
Professor Timothy E. Heron
Professor Diane M. Sainato

Approved by:

Adviser
College of Education
ABSTRACT

This study examined the effects of a Computer-Assisted Tutoring System (CATS) on the accuracy, speed, maintenance, and generalization of time-telling skills with elementary school students with behavior disorders. The tutoring program was adapted based on the peer-tutoring model created by Cooke, Heron, and Heward (1983). The peer tutoring folders were replaced by the computer and specially designed software, created by the researcher, was used. This study was a systematic replication of a master thesis study completed by McKain (2004).
ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisors Dr. Ralph Gardner and Dr. Timothy Heron for their guidance, insight, patience and kindness during the course of my graduate studies. I would never have achieved my goal without your support. I am so fortunate to have worked so closely with you.

I also want to thank Dr. Diane Sainato for her involvement in my research study, for reading the many drafts of this work, and for providing valuable comments.

I would like to give special thanks to my peer graduate students Christopher Yawn and Guoying Zeng, who drove almost every week in a snowy winter to the intervention site to be my observers. I cannot thank you more.

I would like to dedicate this work to my parents, sister, and brother-in-law, who have been a source of love, support, and encouragement throughout these years. I owe my success to them all. And finally, this is for my husband Yong who shared all critical moments with me through this journey. These four and a half years have been priceless in our lives.
VITA

September 04, 1975..............................................Born – Huhhot, Inner Mongolia

1998............................B. A. English for Science & Technology, Tianjin University
Tianjin, China

2001............................M.Ed. Special Education, Millersville University of Pennsylvania
Millersville, PA

FIELDS OF STUDY

Major Field:  Education
Applied Behavior Analysis
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Vita</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Review of the Literature</td>
<td>6</td>
</tr>
<tr>
<td><strong>Effective Teaching Practices:</strong></td>
<td>7</td>
</tr>
<tr>
<td>Direct Instruction (DI)</td>
<td>7</td>
</tr>
<tr>
<td>Behavior Analysis Practices</td>
<td>10</td>
</tr>
<tr>
<td>- Active student response and increase opportunities to respond</td>
<td>11</td>
</tr>
<tr>
<td>- Immediate and frequent feedback</td>
<td>12</td>
</tr>
<tr>
<td><strong>Goal Setting and Cues of Performance</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>Charting and Self-evaluation</strong></td>
<td>16</td>
</tr>
<tr>
<td><strong>Computer-Assisted Instruction (CAI)</strong></td>
<td>19</td>
</tr>
<tr>
<td><strong>Peer Tutoring</strong></td>
<td>20</td>
</tr>
<tr>
<td>- Tutoring Formats</td>
<td>20</td>
</tr>
<tr>
<td>- Cross-aged Tutoring</td>
<td>20</td>
</tr>
<tr>
<td>- One-to-one Tutoring</td>
<td>21</td>
</tr>
<tr>
<td>- Small-Group Tutoring</td>
<td>21</td>
</tr>
<tr>
<td>- Home-Based Tutoring</td>
<td>21</td>
</tr>
<tr>
<td>- Classwide Peer Tutoring</td>
<td>22</td>
</tr>
<tr>
<td><strong>Branches of Comprehensive Peer Tutoring Systems</strong></td>
<td>23</td>
</tr>
<tr>
<td>- Juniper Gardens Children’s Project</td>
<td>24</td>
</tr>
<tr>
<td>- State University of New York at Fredonia</td>
<td>26</td>
</tr>
<tr>
<td>- Vanderbilt University</td>
<td>28</td>
</tr>
<tr>
<td>- The Ohio State University</td>
<td>29</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>McKain’s Computer-assisted Peer-tutoring Study</td>
<td>35</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>43</td>
</tr>
<tr>
<td>Research Questions</td>
<td>44</td>
</tr>
<tr>
<td>Glossary</td>
<td>46</td>
</tr>
<tr>
<td>3. Method</td>
<td>48</td>
</tr>
<tr>
<td>Participants</td>
<td>48</td>
</tr>
<tr>
<td>Settings</td>
<td>49</td>
</tr>
<tr>
<td>Definition and Measurement of Dependent Variables</td>
<td>50</td>
</tr>
<tr>
<td>Accuracy of Time-telling Skills</td>
<td>50</td>
</tr>
<tr>
<td>Speed of Time-telling Skills</td>
<td>52</td>
</tr>
<tr>
<td>Maintenance of Time-telling Accuracy</td>
<td>53</td>
</tr>
<tr>
<td>Maintenance of Time-telling Speed</td>
<td>53</td>
</tr>
<tr>
<td>Generality of the Accuracy of the Four Time-telling Skills</td>
<td>54</td>
</tr>
<tr>
<td>Generality of the Speed of the Four Time-telling Skills</td>
<td>54</td>
</tr>
<tr>
<td>Participant Opinions</td>
<td>55</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>55</td>
</tr>
<tr>
<td>Peer Tutoring System</td>
<td>55</td>
</tr>
<tr>
<td>Computer-assisted Tutoring Program</td>
<td>56</td>
</tr>
<tr>
<td>Participants’ Tutoring Behaviors</td>
<td>59</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>60</td>
</tr>
<tr>
<td>Materials</td>
<td>61</td>
</tr>
<tr>
<td>Procedures</td>
<td>63</td>
</tr>
<tr>
<td>Comparing Group Level Measurement</td>
<td>63</td>
</tr>
<tr>
<td>Tutor Training</td>
<td>64</td>
</tr>
<tr>
<td>Pretest</td>
<td>66</td>
</tr>
<tr>
<td>Goal Setting (Criterion Determination)</td>
<td>68</td>
</tr>
<tr>
<td>Instruction on Time-telling</td>
<td>69</td>
</tr>
<tr>
<td>Tutoring Session</td>
<td>69</td>
</tr>
<tr>
<td>Daily Assessment</td>
<td>71</td>
</tr>
<tr>
<td>Trial</td>
<td>72</td>
</tr>
<tr>
<td>Maintenance Assessment</td>
<td>73</td>
</tr>
<tr>
<td>Generality Assessment</td>
<td>74</td>
</tr>
<tr>
<td>Procedural Integrity and Interobserver Agreement (IOA)</td>
<td>74</td>
</tr>
<tr>
<td>Integrity and IOA on Computer’s Program</td>
<td>74</td>
</tr>
<tr>
<td>Integrity and IOA on Tutor’s Behavior</td>
<td>76</td>
</tr>
<tr>
<td>4. Results</td>
<td>78</td>
</tr>
<tr>
<td>Interobserver Agreement</td>
<td>78</td>
</tr>
<tr>
<td>Interobserver Agreement on Tutor’s Behaviors</td>
<td>78</td>
</tr>
<tr>
<td>Interobserver Agreement on Program Procedure</td>
<td>79</td>
</tr>
<tr>
<td>Procedure Integrity</td>
<td>80</td>
</tr>
</tbody>
</table>
Limitations of the Study ........................................................................................................ 148
  Participant Characteristics ................................................................................................ 148
  Lack of Comparing Instruction ...................................................................................... 149
  Skill Grouping ................................................................................................................. 149
  Time of Study .................................................................................................................. 151
Suggestions for Future Research ..................................................................................... 152
  Accuracy Control ............................................................................................................. 152
  Generalization .................................................................................................................. 153
  Comparing to Conventional Instruction ........................................................................ 154
  Involving External Reinforcers ...................................................................................... 155
  Exploring efficient implementation of the program ................................................... 156
  Exploring the efficiency of tutoring behaviors .............................................................. 156
  Teaching Functions of Telling Time (Estimation and Calculation) ................................. 158
Summary ............................................................................................................................ 159

List of References ............................................................................................................... 160

Appendices:

A. Informational Letter and Consent Form ................................................................. 172
B. Sample Illustration of Printed Data Sheet in a Daily Assessment ............................ 175
C. Social Validity Survey to Participants .................................................................... 175
D. Software Interface Samples ...................................................................................... 177
E. Interobserver Agreement (IOA) Data Collection Form ........................................... 179
F. Interobserver Agreement (IOA) on Program Procedure ............................................. 185
G. Scripts for Training .................................................................................................... 189
H. Scripts for Time-telling Instruction .......................................................................... 196
LIST OF TABLES

Table                                                                 Page

3. 1. Summary of the components and features of the computer software .............. 59
3. 2. Schedule of time-telling skills. ........................................................................ 71
4. 1. Percent IOA on procedure integrity across participants. ................................. 80
4. 2. Percent tutors’ procedure integrity across participants.................................... 82
4. 3. Student 1’s baseline, Level 1, and Trial............................................................. 87
4. 4. Ranges, means, and criteria of accuracies by Student 1................................. 89
4. 5. Ranges, means, and criteria of speed by Student 1.......................................... 90
4. 6. Ranges and means of speeds by Student 1 in maintenance and generality......... 94
4. 7. Student 2’s baseline, Level 1, and Trial............................................................. 99
4. 8. Ranges, means, and criteria of accuracies by Student 2..................................... 101
4. 9. Ranges, means, and criteria of speed by Student 2.......................................... 102
4. 10. Ranges and means of speed by Student 2 in maintenance and generality......... 105
4. 11. baseline, Level 1, and Trial............................................................................. 109
4. 12. Ranges, means, and criteria of accuracies by Student 3................................. 110
4. 13. Ranges, means, and criteria of speed by Student 3......................................... 111
4. 14. Ranges and means of speed by Student 3 in maintenance and generality......... 113
4. 15. baseline, Level 1, and Trial................................................................. 118

4. 16. Ranges, means, and criteria of accuracies by Student 4. ......................... 120

4. 17. Ranges, means, and criteria of speed by Student 4............................... 121

4. 18. Ranges and means of speeds by Student 4 in maintenance and generality.. 123

4. 19. The percentage difference and difference in seconds between the average speeds in Maintenance assessments and the final goals of all participants............... 129

4. 20. The percentage difference and difference in seconds between the average speeds in Generality assessments and the final goals of all participants....................... 130


<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. 1: Image used for Student 3.</td>
<td>62</td>
</tr>
<tr>
<td>4. 1: Accuracy scores across skills by Student 1.</td>
<td>85</td>
</tr>
<tr>
<td>4. 2: Speed scores across skills by Student 1.</td>
<td>86</td>
</tr>
<tr>
<td>4. 3: Accuracy scores across skills by Student 2.</td>
<td>96</td>
</tr>
<tr>
<td>4. 4: Speed scores across skills by Student 2.</td>
<td>97</td>
</tr>
<tr>
<td>4. 5: Accuracy scores across skills by Student 3.</td>
<td>107</td>
</tr>
<tr>
<td>4. 6: Speed scores across skills by Student 3.</td>
<td>108</td>
</tr>
<tr>
<td>4. 7: Accuracy scores across skills by Student 4.</td>
<td>115</td>
</tr>
<tr>
<td>4. 8: Speed scores across skills by Student 4.</td>
<td>116</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

The ability to tell time accurately is a basic life skill that enables people to conduct their daily lives efficiently and effectively. For school aged children, learning to tell time accurately can help them follow school schedules, establish a sense of time elapse, and later learn using time properly. To meet the National Council of Teachers of Mathematics (NCTM) goals that all students achieve higher levels of competence in math and each student will learn to think mathematically, it is important for students to acquire basic skills (Miller & Mercer, 1997). The ability to tell time has been a well established curriculum goal and basic skill for elementary-aged students, including those enrolled in kindergarten and first grade. Children with emotional and behavioral disabilities often have skill deficits when compared to typical peers. However, a review of recent literature found that much emphasis has been placed on problem-solving and conceptual understanding, while little focus has been on teaching basic skills to students with learning and cognitive disabilities (Cawley, Parmar, Foley, Salmon, & Roy, 2001). The NCTM standards also suggest that teachers require adequate resources and a curriculum rich with opportunity to practice. In the Standards, it is suggested that "electronic
technologies – calculators and computers” are essential tools for teaching learning, and doing mathematics. Technology offers teachers options for adapting instruction to special student needs." (NCTM, 2000, p. 24, 25). This study focused on using reciprocal peer tutoring to teach accurate and quick time-telling skills to children with behavior disorders.

Over the years a number of efforts have been initiated to establish directions in math education and special education as a whole to promote effective practices and involve technology into school learning. Since the 1997 Amendments of the Individuals with Disabilities Education Act of 1997 (IDEA ’97), accurate evaluation and outcome have become important issues in the composition of the major document – the Individualized Educational Program (IEP). IDEA ’97 specifies what is to be included in an IEP (20 U.S.C.S. § 1414(d) (1) (A)). The IEP team must generate “measurable annual goals”, including either benchmarks or intermediate short-term objectives. The IEP requires information on interventions (and related materials) that a specific education program is going to be carried out, and projections on how efficient the interventions will be in terms of the time spent and their effectiveness. The IEP becomes a measure of accountability for teachers and schools. Thu, special education law is calling for teaching practices that are engaged in making teaching and outcomes connected, and in having improvement in students’ academic performance the priority. Holding educational interventions accountable is further affirmed in the No Child Left Behind Act of 2002 (NCLB), which was signed into law on January 2002, by President George W. Bush. The NCLB is a major revision of the Elementary and Secondary Education Act (ESEA).
NCLB requires a system of accountability to measure intervention results in school and in academic studies, which leaves major implications in the areas of assessment and technology. The emphasis on effectiveness and efficiency of special education has never been stronger than in the latest revision of IDEA. Congress added new language in IDEA 2004 to assure students with disabilities to improve their “academic achievement and functional performance”, including the use of “scientifically based instructional practices, to the maximum extent possible (20 U.S.C. § 1400(e)).” The revision further describes “scientifically based instructional practices” when used in a child’s IEP as “the special education and related services and supplementary aids and services, based on peer-reviewed research to the extent practicable (20 U.S.C. § 1414(d) (1) (A)).” IDEA 2004 also adds a statement to support “the development and use of technology, including assistive technology devices and assistive technology services, to maximize accessibility for children with disabilities (20 U.S.C. § 1400 (h)).

Long before the ‘97 IDEA amendment, two instructional methods that value the process of instruction and measurable outcomes were identified in a nationally tested Follow Through Project: Direct Instruction and behavior analysis. Follow Through was one of the largest and most expensive social experiments ever conducted (McDaniels, 1975). Follow Through was designed to determine innovative teaching practices in schools serving low-income students. The longitudinal studies evaluated approaches from a wide range of educational philosophies. It was found that DI Follow Through students achieved at a level much higher than is typical for student with similar demographic characteristics. In math computation, Follow Through students instructed using behavior
analysis practices also achieved comparable level (Gersten & Carnine, 1984).

Computer-assisted instruction (CAI) most often refers to drill-and-practice, tutorial, or simulation activities offered either by computers or as supplements to traditional, teacher-directed instruction (Bangert-Drowns, Kulik, & Kulik, 1985). CAI was first developed to enhance basic skill acquisition and practices (Irish, 2002). Along with the advanced computer technology development, CAI has come to the current multimedia software format dominated with sounds, videos, and animations. It has always been an important issue to find out the impact of CAI on educational effectiveness. Numerous studies carried out to assess CAI programs in diverse academic subjects have affirmed its effectiveness (Kulik & Kulik, 1989; Majstek & Wilson, 1989; Renshaw, Taylor, & Reynolds, 1998; Sinclair, Renshaw, & Taylor, 2004; Taylor, Renshaw, & Jensen, 1997).

Similarly, the use of peer students to enhance academic and social performance has been well documented (e.g., Arreaga-Mayer, Greenwood, & Utley, 1994; Harrison, 2002, Heron, Heward, Cooke, & Hill, 1983; Maheady & Sainato, 1985). Peer tutoring is one of the five peer-mediated systematic classroom interventions (i.e., cross-age, one to one, small-group, home-based, and classwide peer tutoring) widely applied to improve students’ academic and social performances. Peer tutoring involves multiple strategies based on behavior analysis principles and practices (e.g., positive reinforcement, constant and immediate feedback, increasing active student responding). As in math instruction, peer tutoring has been found to be beneficial in improving learning of math through the use of pre-arranged and tested procedures, rote memorization, oral and written drills, and
practice using flashcards (Miller, Butler, & Lee, 1998). Additionally, the tutor’s corrective feedback seemed to be a critical component in the tutoring of students with learning disabilities (Barbetta, Miller, Peters, Heron, & Cochran, 1991; Koury & Browder, 1986; Maheady & Harper, 1987).

Combining peer tutoring and CAI, McKain (2004) conducted a study to solve specific problems encountered when the tutor was no more adept than the tutees and needed to hear the answer in a peer tutoring program. In McKain’s study, young children of the same age with developmental disabilities were trained to tutor each other to tell time to the hour, half hour, and quarter hour. A personal computer program was used to display the time on a simulated clock face. Tutoring in the study followed the procedure defined by Cooke, Heron, and Heward (1983). The particular function of the computer program was to present clock times randomly to the tutoring dyad. As the reciprocal peer-tutoring format did not require the tutor to be more knowledgeable on the tutoring content than the tutee, the computer program provided audio answers to the tutor who can hear the answer by wearing earphones. The study showed the positive effects of the computerized peer-tutoring program on student learning. In addition, McKain discussed problems she encountered during the study, and provided suggestions for program revision.
Children with emotional and behavioral disorders usually have academic skill deficits compared to their typically developed peers (Cullinan, 2002). The interaction between poor academic achievement and disruptive behavior can lead to further antisocial conducts (Hallenbeck & Kauffman, 1995). Several teaching practices have been identified as being effective in improving school students’ academic performance across settings, subjects, ages, and ability levels. For example, positive reinforcement has been effectively used with preschool children with language delays (Vollmer, Northup, & Ringdahl, 1996), with middle school students with attention deficit hyperactivity disorder (Evans, Ferre, & Ford, 1995), with students with severe disabilities and aggressive behaviors (Kee, Hill, & Weist, 1999), with pupils with emotional and behavioral disorders (EBD) (Frey & George-Nichols, 2003), and with secondary school students with and without disabilities (Cavanaugh, Heward & Donelson, 1996). Further, reinforcement has been used to increase students’ response rate among elementary school students with moderate and severe mental disabilities (Ketterer, Schuster, Collins & Morse, 2000), and students with learning disabilities and ESL students (Davis &
O’Neill, 2004), on group behaviors (Brigham, Bakken, Scruggs, & Mastropieri, 1992), and on teachers’ behaviors (Gross & Ekstrand, 1983). The effective strategies have been examined separately and combined as packages. With the development of computer technology, some of these packages can be simulated by computer. Computerized educational programs, although based on confirmed effective strategies, are subject to research as well.

This chapter will review the literature on effective teaching practices, specifically, Direct Instruction and behavior analysis practices used in education. Next, other strategies to be included in the peer-tutoring program will be reviewed. These strategies include goal setting, cueing of performance, charting, and self-evaluation. The peer-tutoring program described by Cooke et al. (1983), which is the program that had been applied in current study, and its effectiveness will then be reviewed, followed by literature on computer-assisted instruction (CAI). McKain’s study (2004) of a computerized peer-tutoring program will be discussed in detail as its procedures formed the overall basis for this study. This chapter will also review literature on changing criterion design. Last, the purpose of the study, research questions, and glossary will be presented.

Effective Teaching Practices

Direct Instruction (DI)

The Direct Instruction (DI) Curriculum is a comprehensive instructional program that “integrates effective teaching practices with sophisticated curriculum design, classroom organization and management, and careful monitoring of student progress, as
well as extensive staff development” (Stein, Carnine, & Dixon, 1998, p. 228). The term "Direct Instruction" refers to a rigorously developed and highly scripted method for teaching in fast-pace, which provides constant interactions between learners and the instructor. The fundamental premise underlying DI is to teach more in less time and to control the details of the curriculum (Carnine, Grossen, & Silbert, 1995). The major characteristics of the DI approach include learner-tested curriculum design, scripted lessons (for less teacher talking, more fluent and consistent instruction, and more student active responses), immediate feedback, and high rates of student engagement. These characteristics embody the spirits of the revised IDEA in a student’s IEP composition and implementation. That is, the DI approach provides a well-designed curriculum (curriculum design), and applies highly efficient teaching strategies and materials to set the occasion for the consistency and validity of implementation (effective presentations), and the evaluation of outcomes.

The outstanding effectiveness of DI over other widely applied educational strategies was confirmed in the Follow Through Project, the largest (covering 79,000 students in 180 communities) and the most expensive (costing over 600 million US dollars) educational study ever conducted (McDaniels, 1975). Follow Through was designed to determine innovative teaching practices in schools serving low-income students from kindergarten to grade three. The original rationale to conduct the project was based on the observation that Head Start students began losing the advantages from Head Start by third grade. The longitudinal study evaluated approaches from a wide range of educational philosophies, such as behavior analysis, wholistic learning, and
student-centered learning. The area measured fell in the categories of basic skills, higher order thinking skills (cognitive-conceptual), and self-esteem (affective-cognitive). The results from multiples studies on all the educational programs examined were analyzed by the Stanford Research Institute (SRI) and Abt Associates. It was found that DI Follow Through students achieved at a level much higher in all categories and academic subjects (reading, arithmetic, spelling, and language) than was typical for students with similar demographic characteristics. In math computation, students instructed using behavior analysis practices also achieved comparable level (Gersten & Carnine, 1984).

The key feature that differentiates DI from traditional teacher-directed instruction models is a highly sophisticated analysis of the curriculum instead of using ready made textbooks (Stein et al. 1998). Experts of DI carefully analyzed components and learning procedures of each skill required in a given academic task, and composed a curriculum that systematically develops important background knowledge and explicitly applies and links the background knowledge to new content. Curriculum is examined to identify a few critical blocks of content, the “Big Ideas”. Teaching the critical content raises the efficiency of learning when the students acquire the greatest amount of knowledge that is most critical within the content area (i.e., the so-called “Big Ideas”) in a limited period of time. The Big Ideas are then studied in detail to determine their importance in terms of the frequency with which the content appears within a given grade level, the relationship between the content and the curriculum, and instruction and practicing time. The Big Ideas are then sequenced to be instructed accordingly. DI curriculum includes generalizable strategies to expand skills learned to a broader range of problem types.
through presenting enough examples of a full range of types concerning one skill content. Once the Big Ideas and sequences are determined, effective instructional strategies are identified to present the curriculum content explicitly. The DI curriculum uses scripted lessons to assure that teachers use clear and consistent language, and to have more opportunities to monitor students and provide feedback while teaching.

Behavior Analysis Practices

Another set of approaches that values process of instruction and measurable outcomes identified in Follow Through Project was strategies based on behavior analysis principles (National Assessment of Educational Progress, 1979). In defining applied behavior analysis (ABA), Cooper, Heron, and Heward (1987) distinguished it from other disciplines by emphasizing its focus, goals, and methodology:

“Applied behavior analysis focuses on objectively defined, observable behaviors of social significance; it seeks to improve the behavior under study while demonstrating a reliable relationship between the procedures employed and the behavioral improvement; and it uses the methods of science—description, quantification, and analysis” (p. 2).

ABA’s theoretical basis and accumulated research through the years have had extensive evidence in their effectiveness in improving children’s behavior and learning. Practices based on behavior analysis research being used in classroom instructions and exercises have shown consistent effects across various situations. These practices have been applied comprehensively in some curriculum packages. The following is a brief discussion of practice of active student response and immediate and frequent feedback, which are derived from behavioral theories and are relevant to the current study.
Researchers found that student academic achievement has a positive relationship with the amount of active participation in the learning process (Cavanaugh & Heward, 1996; Gardner, Heward, & Grossi, 1994; Shin, Deno, Robinson, & Marston, 2000). That is, the more responses students make during a lesson, the more they learn. These studies also presented evidence that opportunities for student to respond in academic activities can be created, and the amount of these opportunities could be raised to positively affect students’ academic success.

Many studies have examined the effects of using response cards on students’ response rate, academic performances, and classroom behaviors and obtained positive results (Heward, 1994; Narayan, Heward, Gardner, Courson, & Omness, 1990). Researchers found that comparing the use of response cards and hand raising, students had more opportunities to respond using the response cards, and their accuracy rates were higher. Students also reported preference on the format of response card that required more responses from the students (Gardner, Heward, & Grossi, 1994; Narayan et al., 1990). Similar results have been obtained among elementary school students with moderate and severe mental disabilities (Ketterer et al, 2000), secondary school students with and without disabilities (Cavanaugh et al., 1996), and students with learning disabilities and English as Second Language (ESL) students (Davis & O’Neill, 2004). These studies concluded that (a) a functional relationship exists between active responding and the acquisition of academic information, (b) active responding increases student responding and accuracy, and (c) involvement increases student learning and on-
task behavior in studies with elementary and preschool children. More recently, Davis and O’Neill (2004) extended research on active student responding by examining the effects of response cards on academic and off-task responding during writing instruction in a resource room among secondary school students with learning disabilities. Six participants completed the study. The researchers measured the percentage of trials on which the participants made academic responses when response cards were available or absent, participants’ correct response rates, and the rate of off-task behaviors. The result showed that all participants achieved higher average levels of correct academic responding during response-card conditions (M=91%) than in the hand-raising conditions (M=74%), and the group average weekly quiz scores were substantially higher during the response-card conditions (M=88%) than in the hand-raising conditions (M=19%). The measures of social validity showed that the participants’ written answers listed more negative aspects of hand-raising than response cards, although the participants’ oral statements indicated that they liked hand-raising more. Still all the participants expressed their satisfaction with respect to their academic improvements.

Immediate and Frequent Feedback

Researchers suggest a continuous schedule of reinforcement at the initial stage of acquiring new behaviors (Cooper et al., 1987). That is, for each correct response, the student should be awarded a reinforcer; for any incorrect response, the student should be corrected or redirected immediately. The correction should be made every time an incorrect response is emitted. The schedule of reinforcement will be gradually thinned later when the occurrence of the target behavior reaches a certain level.
A significant amount of literature has demonstrated that immediate and frequent feedback is positively associated with student learning (Fuchs, 1986, Gersten, Carnine, & Woodward, 1987). Researchers found that students achieved higher academic performance levels when the instructor provided immediate corrective feedback for their errors (Gersten, 1987). Robinson DePascale, and Roberts (1989) found that consistent feedback helped students with learning disabilities complete more problems and improved accuracy from 73% to 94%. These researchers addressed the specific functions of immediate and consistent feedback on the performance of students with learning disabilities: since these students often exhibit less of an attention span, tend to make more careless mistakes, and participate in fewer academic activities. Strategies such as immediate and frequent feedback that can increase the opportunities of students with learning disabilities to be engaged in academic activities are beneficial and critical. Blankenship and Lilly (1981) discussed the functions of feedback and indicated that feedback may help students to distinguish between correct and incorrect responses, and to recognize their learning progress. Larrivee (1986) identified several instructional behaviors that are significantly correlated to the academic performance of students with disabilities attending general education classes as (a) frequent positive feedback, (b) corrective feedback contingent on incorrect student responses, and (c) supportive responses to achievers of all levels.

**Goal Setting and Cues of Performance**

Goal setting involves identifying a particular criterion level (i.e., the goal) for a target behavior, and consistently comparing the criterion level with the performance of
the target behavior to examine if the behavior is improving towards the goal. Goal setting has been used in many studies to increase or decrease a variety of behaviors in various fields (e.g., athletics, education) (Langeland, Johnson, & Mawhinney, 1998; Lerner & Locke, 1995; Ward & Carnes, 2002). For individuals with cognitive disabilities, Copeland and Hughes (2002) reviewed literature on goal-setting and found that in conjunction with other instructional strategies, use of goal setting strategies was associated with increased rate or accuracy of performance of children and adults with mild to severe mental retardation. When reviewing experimental studies on goal orientations with students with mental retardation, these researchers also found that there is no significant difference in task performance between self-selected goals and assigned goals for this population.

Consistent with the findings of Copeland and Hughes, many other studies revealed that goal-setting techniques are more effective when combined with other behavioral approaches (e.g., summary feedback, self-monitoring, direct instruction, self-graphing, behavior contracting) than when used separately (Bandura & Cervone, 1983; Ruth, 1996; Trammel & Schloss, 1994). Trammel and Schloss’s study investigated the effects of combining self-monitoring techniques with goal-setting and self-graphing on the homework completion of secondary level students. Eight students aged 13 years 9 months to 16 years with learning disabilities were selected to participate due to their consistent failure to complete daily homework assignments from multiple teachers. Students were instructed to complete an assignment sheet daily and self-record if each assignment was completed and submitted. Students’ homework was considered
completed when accuracy reached 70% or higher. In the following phase, in addition to the self-recording, the students were taught to self-graph the data of their progress of homework completion. The graphs were posted in the resource room. Students were also instructed to set a goal every three days. Each goal was set by the student and it was to the average level for the preceding three days. A multiple baseline design across subjects was used to assess the effectiveness of the intervention package. An increasing rate of assignment completion (from 0 to 4 to 4 to 6 assignments handed in daily) was found across all subjects. The increment continued in goal-setting and graphing phase. Each subject maintained to complete of at least five assignments daily when no goal, rewards, or visual reminders were presented.

Rachlin (1978) suggested that for individuals with cognitive deficits, an abstract concept such as a preset goal might be difficult to comprehend. A number of studies on this group circumvented the problem by applying visual or audio cues to represent goals, define goal levels, or provide feedback on participants’ progress towards a goal (e.g., Cole & Gardner, 1988; Grossi & Heward, 1998; Kokubun, 1999; Mullen & Martin, 1988).

Kokubun (1999) examined the effects of setting visually cued goals on students’ standing broad jump performance. The subjects were 30 teenagers of an average age of 16-2 years and with intellectual disabilities. Subjects’ jumping performances were measured under conditions with and without goals. Each subject completed 4 jumps, with the first and the fourth jumps without goal and the second and third with goals. Under the condition in which no goals were set, students were instructed to jump as far as they
could. Under condition in which goals were set, students were provided with an assigned goal and were instructed to jump to reach the goal. Each goal was set 20cm farther than a distance measured in previously trial of each student. The goal was shown visually using a strip of plastic tape laid on the floor. The tape was removed under the “without-goal” condition. The result of performance under the “goal” condition was a mean of 108 cm, while that for the “without-goal” condition was 102 cm. The differences between the two conditions were statistically significant \((F = 15.3, P < 0.005)\). The author obtained similar results with slightly older students (20 to 24 years) without cognitive disabilities (twenty students; mean for goal condition was 162.5 cm and for without-goal was 154.2 cm; \(F = 54.0, P < 0.005\)). Data also demonstrated a significant relationship between the condition difference and behavior regulation score measured during jumping. The study did not find comparative relationship between students’ jumping performance with their chronological age or intelligence level (IQ scores). However, the author discussed the relatively little effects on participants with Down syndrome, which indicated that these children’s deficient motor ability may have diminished the effect of goal setting.

**Charting and Self-evaluation**

According to Van Houten (1980), plotting one’s own behavior performance on graphs can provide feedback to the individual in study. Such feedback presented immediately after the observational period shows an ongoing record of the behavior. In addition, charting itself has been an effective intervention used in academic and behavior therapy objectives (Levin & Carr, 2001; Wolfe, Heron, & Goddard, 2000)
Charting can assist the teaching of self-evaluation skills. Positive results have been found in research that combine self-monitoring / self-evaluation with goal-setting and charting procedures (Copeland & Hughes, 2002). In Grossi and Heward’s study (1998) helping adults with developmental disabilities increase work productivity in restaurant training, result showed that accuracy or rate of task completion behaviors increased when participants monitored their performance, compared with preset goals, and verbally stated the result to a supervisor of their progress.

Research indicated that students with disabilities can be taught to evaluate their learning progress toward pre-set goals by reviewing graphs, checklists or charts (Smith & Nelson, 1997). In research using a multicomponent intervention package to enhance schoolwork performance of four teenage students with mental retardation (Copeland, Hughes, Agran, Wehmeyer, & Fowler, 2002), researchers visualized the goal (e.g., achieve a certain grade) by 1) providing rationales for a goal and defining the goal to the students, 2) showing the students a list of scores earned during baseline, 3) asking students to set a performance goal, and 4) using a goal-evaluation sheet (Grossi & Heward, 1998) to provide the students with a visual representation of their goals to achieve the final goal. Combining with multiple behavioral strategies (i.e., modified assignments, direct instruction, self-monitoring), all students’ overall school worksheet performance increased dramatically (e.g., 3% to 86%, 6% to 98%), and continued when procedures were withdrawn. Three of the four participants responded positively to the function of the goal-evaluation sheet, which, as one participant reported, helped them to see their grades (i.e., the final goal).
Snyder and Bambara (1997) taught students to view their goals of applying “classroom survival skills” (e.g., handing schoolwork in on time, raising hand) using a checklist. Three adolescents were taught to use their checklists to count the number of skills they performed daily and to compare it with a pre-set goal. Similar to the study of Copeland, Hughes, Agran, Wehmeyer, and Fowler’ study (2002), using visual cues (e.g., charts, checklist, graphs) could be assisting components in a treatment package that contained multiple strategies. Results indicated students’ generalization of the learned skills to other settings and maintaining over time.

In a study conducted by Wolfe, Heron, and Goddard (2000) on four elementary boys (aged 9) with learning disabilities who performed at a very low level on writing, one target objective was to enhance the students’ written language performance. A five-phase changing criterion procedure was used. Goals (criterion levels) for each phase were assigned for students according to their work in previous phase. During treatment interval, students were instructed to write an essay for 10 minutes, then count and graph the number of words written. If the number met or exceeded the goal for that day, the student could place a star on a chart posted on the wall, which was connected to a token system used in the classroom. The researchers found that teaching self-monitoring alone to increase students’ on-task behavior could not increase written language performance. After the changing criterion with public posting, the writing skills of three of the four participants increased.

Besides visually presenting the effects of the program and providing feedback to students, using graphs can help teacher’s decision making. Graphs provide a direct and
ongoing view of students’ performance and progress using the program (Johnston & Pennypacker, 1980). For teachers, visual analysis takes little time to learn and is easy to use.

Computer-Assisted Instruction (CAI)

Computer assisted instruction (CAI) is one form of assistive technology (AT). The IDEA defines AT for educators as any product in forms of an object or a system that are either designed and produced commercially, or created for individuals to increase, maintain, or improve functional capabilities of people with disabilities. Complicated as it sounds, however, AT does not equal computerized, expensive devices and software that require extensive training. AT options range from low-tech applications (e.g., tape player) to high-tech applications (e.g., voice detective system). It is very often that high- and low-tech applications co-exist in practical settings (Aston, 2000; Polloway, Patton, & Serna, 2001). CAI, standing at the high-tech end of AT, teaches directly through the computer (Edward, Norton, Taylor, Weiss, & Dusseldorp, 1975).

Since its introduction to classroom education in the 1960’s, CAI has been examined repeatedly as to its effectiveness with diverse academic subjects and age groups. CAI has shown its effectiveness to enhance rote memorization (Kulik, Bangert, & Williams, 1983; Kulik & Kulik, 1989, Majstek & Wilson, 1989) and higher-order problem solving skills (Renshaw, Taylor, & Reynolds, 1998; Taylor, Renshaw, & Jensen, 1997; Wilson, Cassella, & Wilson, 1989). The impact of CAI on practice and drill is better documented than that on problem solving (Sinclair, Renshaw, & Taylor, 2004). Basic skill practices can be better fulfilled through computer software, which efficiently
frees up resources for conceptual learning at basic levels and presents massed practice in a more efficient way than teacher-directed, traditional practice. For higher-order problem solving skills, it is suggested to software designers that a more complicated learning program should be simplified to a selection of shorter, stand-alone, and more focused exercises (Huff, 1998). The reason is that these quickly accessible exercises are more likely to be incorporated into classroom teaching and thus the advantages of CAI are more likely to be taken from the applications. Also, focused tasks can free up students from distracting elements and concentrate on a particular problem, especially higher-order problems.

Peer Tutoring

Peer (classwide) tutoring system is one of the five commonly used systematic tutoring variations under peer-mediated approaches that have been identified as effective and have been subjected to extensive study (e.g., Cooke, Heron, & Heward, 1983; Delquadri, Greenwood, Stretton, & Hall, 1983). The other four tutoring formats include cross-aged, one-to-one, small group and home-based tutoring (Miller, Barbetta, & Heron, 1994). Each of the five tutoring formats, including classwide peer tutoring, will be briefly discussed.

Tutoring Formats

Cross-aged Tutoring

Cross-aged tutoring involves individualized instruction by an elder tutor to a younger learner. Usually the age difference between the students is approximately two or more years. Often both tutor and learner are from the same school; in some cases,
however, junior-high or high-school students from other local schools serve as tutors for elementary students (Barbetta, Miller, Peters, Heron, & Cochran, 1991).

**One-to-one Tutoring**

One-to-one tutoring involves intensive tutoring to students with particular remedial needs in specific academic area. Only selected student dyads are engaged in tutoring using this format. A flexible schedule and tutoring contents can be adopted by the dyads. The tutors can be students with higher skill level in the subject. It is indicated, however, that students being trained properly to use a structured tutoring system can serve competently as tutors without originally high skill level in the tutoring content (Franca, Kerr, Reitz, & Lambert, 1990).

**Small-Group Tutoring**

Small-group tutoring can be conducted in two ways. In one way, the small-group involves students who need remedial practices. These students practice on extra school work through tutoring. Another way involves the whole class in peer tutoring on a rotational basis. Different from classwide format, students take turns on tutoring activities in small-groups on a daily or weekly basis, while the rest of the class engage in usual school work and other class activities (Miller, Barbetta, Drevon, Martz, & Heron, 1996).

**Home-Based Tutoring**

As the name suggests, home-based tutoring is an extension of school practices to after school periods at home, where parents or elder siblings serve as tutors (Barbetta & Heron, 1991). Home-based tutoring is yet to be thoroughly studied; still, according to a
few studies, parents can serve as effective tutors for their children (Barbeta & Heron, 1991; Elksnin & Elksnin, 1991).

\textit{Classwide Peer Tutoring}

Classwide peer tutoring involves all students in the class engaging in tutoring activities simultaneously during regular school day in typical instructional settings (Miller et al., 1996). It is a tutoring format that has been widely studied and its effectiveness has been well documented (e.g., Cooke et al., 1983; Delquardi, Greenwood, Whorton, Carta, & Hall, 1986; Greenwood, 1991). The advantages of this tutoring format include incorporating a wide range of subject areas, and diverse student ability and age levels. Still, this format of tutoring requires more teacher time and effort on material reparation, tutor training, and data analysis on daily basis.

Not all tutoring activities that involve peer students or siblings as tutors fall automatically in one of the formats discussed above. Peer tutoring is defined as “any formal and comprehensive approach to teach students to prompt, praise, test, and chart the academic, social or non-traditional skills of their partners on a daily basis” (Heron & Harris, 2001, p. 452). Behavior analysis principles and practices have contributed to the development of a peer tutoring system, which distinguishes peer tutoring from incidental tutoring arrangements in the following areas:

1) Program design – each step in a peer tutoring system is described and taught directly and clearly to tutors and tutees, and the steps are executed following a pre-arranged logical sequence. The effective practices used in tutoring are also applied in tutor training sessions. Tutor training is a critical component in peer tutoring, in which
tutors are taught to conduct tutoring sessions consistently and systematically following
the pre-designed instructional formats.

2) Effective presentations of materials – instead of using any convenient
instructional methods, peer tutoring systems employ effective practices (e.g. active
student response, student’s opportunities to respond, immediate and frequent feedback,
positive reinforcement), which are supported by extensive behavioral research. These
practices are established on behavioral principles. By employing these highly effective
teaching strategies, peer tutoring sets the occasion the consistency and validity of its
implementation;

3) Systematic evaluation – peer-tutoring systems evaluate tutee’s learning
progress systematically to build maintenance. Learning and generality are not assumed.
Instead, outcomes of a peer tutoring system are decided through systematic measurement.
A peer tutoring system pretests tutees to decide tutoring materials, daily tests and records
tutees’ performance after each practice session, and tests tutees on the same materials
after a period of time for maintenance. Results of these tests result in the adaptations of
materials and procedures for particular tutees when needed.

Branches of Comprehensive Peer Tutoring Systems

Group learning, the format that peer tutoring is taking, has been discussed by
Maheady (1997) in his article instructing teachers on teaching groups with multiple
ability levels. Maheady suggested that grouping is one of the alterable instructional
factors that "can powerfully influence positively or negatively the levels of individual
student engagement and hence academic progress" (p. 325). With more students with
disabilities attending general education classes, effective grouping practices will help teachers to include all students in the classroom into class activities and address each student’s needs. Maheady identified four centers in the United States that have conducted extensive research on peer tutoring and have created their own tutoring system. These centers are: Juniper Gardens Children’s Project at Kansas City, State University of New York at Fredonia, Vanderbilt University, and The Ohio State University. This section will provide a brief introduction of the tutoring system developed at each center.

*Juniper Gardens Children’s Project*

The Class-wide Peer Tutoring (CWPT) program developed by researchers at the Juniper Gardens Children’s Project in Kansas was initially designed to prevent academic failure in culturally diverse children from low socioeconomic families (Delquadri et al., 1983). CWPT is a system in which all the students in one class are organized into tutor-tutee pairs and work together on competing teams. Tutees earn points for their team by responding to the tasks presented by their tutors. Greenwood, Delquadri, and Carta (1988) summarized the core procedures of CWPT: (a) teacher’s review of materials and introduction of new content to be learned, (b) clarification of content materials to be tutored in each session (e.g., reading passages, spelling word lists, or math fact lists), (c) arrangement of new partners each week, using partner pairing strategies, (d) students’ reciprocal roles in each session (i.e., students serve as both tutor and tutee), (e) teams competing for the highest team point total, (g) contingent individual tutee point earning, (h) tutors providing immediate error correction, (i) public posting of individual and team scores, and (j) social reward for the winning team.
During the first decade of success in demonstrating the effectiveness of the program among students who are at-risk in general education classrooms, CWPT was also extended to special education settings where the program was replicated and analyzed. A number of studies demonstrated improvements in the oral reading rates of students with learning disabilities, the majority of whom were students from minority groups, or students to whom English is a second language (ESL) (Greenwood, Delquadri, & Hall, 1984). CWPT was also found superior over teacher-mediated instructions for ESL students, and students come from minority groups (Madrid, Terry, Greenwood, Wahley, & Webber, 1991). Further, CWPT has been studied and demonstrated to be effective with students of diverse ability levels, ages, and settings (e.g., elementary-aged students with learning disabilities (Mathes, Fuchs, Fuchs, Henley, & Sanders, 1994); elementary-aged students with mild retardation in inclusive settings (Mortweet, 1995; Sideridis, 1994); secondary students with mild disabilities (Maheady, Sacca, & Harper, 1988); children with Attention Deficit Hyperactivity Disorder (ADHD) in general education setting (DuPaul & Henningson, 1993), and students with autism (Kamps, Barbetta, Leonard, & Delquadri, 1994)).

Standard procedures in CWPT also facilitate students with disabilities being included in general educations environments. First, CWPT, as a class activity organizing format, is not limited to particular class content. It can be adapted to commercial curriculum materials and individualized content as well. Second, the program includes an active and influential resource in the classroom, the peers, into the teaching process. Third, the group contingency arrangement provides extensive reinforcers to students with
disabilities not exclusively contingent upon individual’s performance, but on the collective performance of the individual's partner and team. Fourth, students have opportunities to interact widely with all other students in the classroom through the weekly pair-changing procedure. Students take reciprocal roles during each session provides students with opportunities to learn the social and teaching skills and also greatly motives the students (Greenwood & Delquadri, 1995).

The efficiency and accuracy of CWPT program are greatly enhanced when computerized software is developed for teachers to store and rank scores, analyze and arrange process, plan partner pairing, and time the tutoring sessions accurately. The program can also identify students whose performance improved over prior sessions and can prompt the teacher to provide praise accordingly at that particular point in the session. In addition, the software provides teachers with data interpretation and suggestions (Greenwood, Finney, Terry, Arreaga-Mayer, Carta, Delquadri, Walker, Innocenti, Lignugaris-Kraft, Harper, & Clifton, 1993).

State University of New York at Fredonia

Inspired by the CWPT program, researchers at the State University of New York at Fredonia created Classwide Student Tutoring Teams (CSTT) by combining components of the CWPT with the Teams-Games-Tournaments (TGT) program developed by Slavin and colleagues at the Johns Hopkins University (Slavin, 1986, 1990). The elements incorporated in CSTT that are similar to CWPT include structured tutoring procedures, daily point earning procedure, public posting of scores, and group contingent reinforcement. Different from CWPT, students are grouped into small
heterogeneous learning teams of three to five members, instead of paired dyads. Grouping is arranged according to the teacher’s private ranking of each student’s academic performance level. When assigning groups, the teacher positions students of different rankings evenly into the groups to keep relatively equal numbers of high-, low-, and intermediate-achieving students in each group. Each student in the group takes turns as the tutor and dictates the learning content during the 30 minute practice and testing session. The tutor is responsible for checking tutee’s responses, presenting correct answers when mistakes occur, directing tutees to self-correct mistakes, and providing points accordingly. Groups earn points for correct responses, error corrections, and following procedures correctly. The teacher’s major responsibility in the program is preparing a study guide with a series of questions and answers (usually between 10 to 30 items) and distributing each group of a deck of numbered cards corresponding to the numbered questions (Harper, Maheady, Mallette, & Karnes, 1999).

Maheady and his colleagues (1987) explored the effectiveness of CSTT on math performance of six classes of 9th and 10th grade students with lower academic achievement, and found that with CSTT instruction, the students’ weekly math quiz scores increased from 62% to 82%. Participants with learning and behavior disorders are reported to perform at the same level of their non-disabled peers. Only one out of 93 students received a failing grade, according to the result, comparing to 45% of all students failed under traditional teacher-led instruction.

To facilitate teachers using CSTT, a videotape and a manual have been published to instruct teachers how to use the program (Harper, Maheady, Mallette, & Sacca, 1992).
The authors stated that one of the important goals of CSTT is to teach students cooperative skills (Harper et al., 1999). Arguments that often occur at the initial stages when using CSTT among group members are suggested to be ignored as opportunities for the students to learn what behaviors work and what does not. Similar to other tutoring programs, CSTT is designed for reviewing purposes, so that students should have access to the learning content fully before program is used.

Vanderbilt University

Peer-Assisted Learning Strategies (PALS) was developed by Dr. Doug Fuchs and Dr. Lynn Fuchs, Professors of special education at Vanderbilt Kennedy Center, based on Class-Wide Peer Tutoring (CWPT) developed at Juniper Gardens (Fuchs, Fuchs, Thompson, Svenson, Yen, Otaiba, Yang, McMaster, Prentice, Kazdan, & Saenz, 2001). The original intention of developing a modified version of CWPT is to enrich tutoring activities, retain the reciprocal teaching format, and keep the feasibility of CWPT. PALS provides reading and math programs from kindergarten to high school, with a similar format but various motivational strategies that are age appropriate. PALS is designed as a class organizing format that does not require special reading materials; in PALS math, additional materials are available for kindergarten through grade six. It is recommended that PALS be implemented during 25- to 35- minute activities two to four times a week. Like CWPT and CSTT, PALS is one version of peer tutoring that groups students into smaller learning units and provides students with awarded points for desired academic and social behaviors (Fuchs et al., 2001). The grouping procedure is different in PALS, in that the teacher identifies (a) students and specific skills the students need to improve,
and (2) students who are most appropriate in terms of academic performance and social skills to help the students’ needs. Using this information, the teacher pairs students in the class, so that partners work simultaneously on different activities that address specific skill area. Pairs are not permanent but changed regularly. In each session, students take turns to be the “coach” and the “player”, which are corresponding roles of tutor and tutee in CWPT and CSTT.

A number of studies using PALS on elementary students with disabilities have found that students with disabilities who participated in the PALS program achieved higher scores than comparison group students with comparable disabilities in math computation and math concept applications (Fuchs, Fuchs, Hamlett, Phillips, & Bentz, 1994; Fuchs, Fuchs, Phillips, Hamlett, & Kams, 1995). PALS has also been shown to improve reading abilities among low- and average-achieving students and students with learning disabilities (e.g., Fuchs, Fuchs, Mathes, & Simmons, 1997; Simmons, Fuchs, Fuchs, Hodge, & Mathes, 1994). Limited studies have applied PALS with secondary students with disabilities and found promising results suggesting that PALS can be effective in increasing the mathematics skills (e.g., basic algebra problem solving skills, Allsopp, 1997) of this population.

*The Ohio State University*

The classwide peer-tutoring model created by Cooke, Heward, and Heron at The Ohio State University involves dyadic peer tutoring practice. The procedure usually involves pretest, tutoring and posttests. Pretests are given to examine tutee’s knowledge of the content to be learned. Only the contents not known are selected to tutor. Content is
printed on flashcards. When tutoring, the tutor presents flashcards of content to be learned one at a time to the tutee, and prompts an answer. The tutor praises the tutee for correct responses, and prompts for a second time contingent on an incorrect response. The tutor gives corrective feedback if the tutee fails to answer correctly for a second time, and prompts the tutee to repeat the answer. Tutoring is followed by assessment when the tutor presents the same flashcards one at a time to the tutee without feedback. The tutor places cards that the tutee answered correctly in one pile and those that were incorrect are placed in another pile. Mastery of one card is determined when the tutee responds correctly to the card for two to three consecutive sessions. Mastered cards are withdrawn from the practicing cycle and new unknown cards are added to the set once all cards in the original set are mastered. It is a non-replacement system (Cooke et al., 1983; Miller et al., 1994; Heron et al., 2006).

Miller, Barbetta, and Heron (1994) created the term START tutoring to describe the main features of a systematic tutoring approach. START stands for Select a tutoring format, Train the tutors, Arrange the environment, Run the program, and Test for effectiveness. For each step of START tutoring, the teacher can choose from a list of substeps or alternatives to conduct. For example, in selecting a tutoring format there are five options (i.e. classwide, cross-age, small group, one-to-one dyads, and home-based, p.267), and to run the program, five components are to be administered (i.e. pretesting, practice, testing, performance tracking and charting, and maintenance testing, p.272). In the following, each of the five steps of START tutoring will be described:
(a) Pretesting. In pretesting, the tutor presents one item (e.g., word, math problem, fact) at a time on a list or sheet to the tutee. Unknown items (i.e., no response in given time, usually three to five seconds, or wrong responses) are then identified and are printed onto flashcards. Pretesting is conducted repeatedly throughout the tutoring as new items need to be identified when the tutee masters the cards in practice. Modifications can be made in pretest to cater to students’ specific needs. For example, the teacher can reduce the set size from ten unknown items to five when particular frustration is observed in the tutee, or when necessary, print the definition of a term on the back of the card.

(b) Practicing. During practice, tutor and tutee use a folder with an attached tracking graph and pockets labeling “Go”, “Stop”, and “Star”. The pockets contain working flashcards, mastered cards, and star stickers respectively. The tutor presents flashcards to the tutee from “Go” pockets one at a time, continuously reinforces the tutee’s correct response, prompts a second try if incorrect or no response is provided by the tutee, and models the correct answer immediately after the tutee’s second wrong attempt. Tutoring dyads are required to go through the practice card set as many times as possible within the allotted practice time; cards are shuffled after each round to randomize order. In classwide peer tutoring, small-group and one-to-one dyad tutoring, after about 5 to 10 minutes, the tutors and tutees switch roles. A modification often applied in this stage is Tutor Huddle. This is arranged for tutors who are still in the acquisition phase. In the early time of the tutoring day, tutors will participate in group practice that addresses the items to be tutored that day. Other modifications may include varying number of cards to be practiced, using key phrases to cue, error correction
variations (e.g., “Try again”, “No”, “Wrong”), and using a time delay procedure to transfer stimulus control from a prompt to the natural stimulus.

(c) Testing. After each practice session, the tutee is tested on the cards that were practiced. With no feedback, the tutor presents the cards one at a time and marks card with “O” or “X” for correct and incorrect respectively. Cards with three consecutive “O”s (i.e., the tutee correctly answered in three sessions in a row) are placed in the “Stop” folder as mastered items. Teachers can arrange modifications in this session by changing the criteria for mastery level (e.g., from correct answer in three consecutive days to five), using symbols to signify correct and incorrect responses for younger students. Teacher can also decide if students will practice the same set of items until mastery before new items are introduced, or to implement a card-replacement system which will replace mastered cards with new cards daily to keep the number of practicing items constant (e.g., 10 cards each day).

(d) Performance tracking and charting. Performance tracking and charting involve having the tutee graph his or her daily progress, a visualized feedback and reinforcement for young students to track the total number of items mastered. The teacher can apply a variety of reinforcers and schedules to present the reinforcers (e.g., using contingent and non-contingent reinforcement, group-oriented contingencies).

(e) Maintenance testing. To measure if the students have maintained the skills learned in tutoring, the teacher collects and dates cards mastered by the student, then moves from the “Stop” pocket for the maintenance test usually seven days after acquisition. For cards not mastered in a maintenance test, that card is added to the next
set of cards to be inserted in the “Go” pocket, but cards do not exceed a 10-card set. Cards for which the tutee has shown mastery in this test will be removed from the system. The interval between the completion of the tutoring items and the maintenance test can be longer or shorter. If necessary, the teacher can have parents involved in giving maintenance test to the students at home.

The model-lead-test method brought by Carnine and Silbert (1979) has been used in many tutoring programs to train students and has been found effective (e.g., Cooke et al., 1983; Maheady & Sainato, 1985). This method allows tutors to practice new tutoring skills, provides feedback on tutors’ performance, and gives teachers a method for evaluation. Other methods are also effective to teach tutors to follow prescribed tutoring protocols (e.g., role playing, token economy, instructional packages, videotape presentation) (Miller et al., 1994). Tutor training typically includes instruction in: a) transitions (to and from tutoring), b) practice procedures, c) rules for tutor and tutee behavior, d) reinforcement strategies for correct responses, e) error correction procedures, f) methods of gaining the teacher’s attention, g) testing procedures, and h) record keeping for tutee performance (Miller et al., 1994).

The peer-tutoring model has been applied and has demonstrated its effectiveness in research across settings (e.g., self-contained, resource, inclusive, after-school settings) (Gardner, Cartledge, Siedl, Woolsey, Schley, & Utley, 2001; Heron et al., 1983; Mortweet, 1995) and academic subjects (e.g., reading, math, spelling, ESL) (Arreaga-Mayer et al., 1994; Arreaga-Mayer, 1998; Barbeta et al., 1991; Delquadri et al., 1983). A peer-tutoring program can be modified to meet specific needs if necessary. As Heron and
colleagues (2006) suggested, modifications to the design, structure, and response mode of
general educational materials should be considered when necessary for students with
special needs. Adaptations of materials have effectively assisted in traditional tutoring
programs. For example, there are studies using picture-based icons on the back of cards
as cues for tutors to provide feedback (Wright, Cavanaugh, Sainato, & Heward, 1995;
Harrison, 2002). Harrison (2002) first adapted a “card holder” to hold flashcards, which
allowed the tutor to use both hands to write when necessary. In the same study,
commercial clip art materials were also adapted to provide referent to the tutors who are
deaf; and other cues and prompts served the same function are also found in other studies
(Miller et al., 1994; Miller et al., 1996).

With the multimedia advantages that a computer program have, tutoring programs
using computer-adapted materials were found especially useful when the tutor was no
more adept than the tutees and needed to hear the answer (McKain, 2004). In McKain’s
study, young children of same age with developmental disabilities were trained to tutor
each other to tell hour, half hour, and quarter hour. A personal computer program was
used to display the time on a simulated clock face. All other tutoring steps in the tutoring
used the procedure defined by Cooke et al. (1983). The particular function of the
computer program was to present clock scenes randomly to the tutoring dyad. The tutor,
as in traditional tutoring programs, prompted the tutee for answer (“What time is it?”),
provided feedback (“Good job”, or “Try Again”, or “Say (time)”), and recorded results
by clicking corresponding buttons (green button for correct and red button for incorrect).
Specifically, as the reciprocal peer-tutoring format did not require the tutor to be more
knowledgeable on the tutoring content than the tutee, the computer program provided audio answers, which the tutor can hear by wearing an earphone. McKain’s study will be discussed in depth in the following section.

**McKain’s Computer-Assisted Peer-Tutoring Study**

McKain conducted a computer-assisted peer-tutoring program for her thesis study (2004). The study combined the peer tutoring model created by Cooke et al. (1983) and a set of computer software to teach young students time-telling skills. The software was developed by the present experimenter. Six first grade students with cognitive disabilities were trained to provide individual tutoring using a set of computer software. Students were trained using a “model-lead-test format” (i.e., through skill demonstration, teacher modeling, group role playing, and student practice). Tutoring behaviors included “initial prompting, prompting for incorrect response, praise, and printing results”. Seven to 17 sessions (approximately 20 minutes each) were conducted to train the students to use the program. On completion of training, students were paired to three tutoring dyads and tutored each other on hour, half hour, and quarter hour times. The set of computer software can present clock faces individually in a random order. When tutoring, one student in the pair would be in the role of tutor, and the other the tutee. The pair would switch its roles after one tutoring session. Throughout tutoring, the tutor would wear a pair of earphones to hear the time from the software program while the tutee did not. The tutor could replay the sound of the time if not sure by clicking a loudspeaker button. Contingent to the tutee’s response to the time of the clock face, the tutor provided feedback, “Good job” for correct answer and “Try again” for incorrect answer. The tutor
would tell the time and have the tutee repeat the time when the tutee was wrong for a second time. At the end of the tutoring session, the program would prompt the tutor to print the result, which recorded the tutor and tutee’s names, times included in the session, and correct answers the tutee made. Each student’s acquisition data were collected using a Large Judy Clock® by the researcher after tutoring. Acquisition in McKain’s study referred to the number of accurate responses of time a student made during one session.

Results of McKain’s study showed increases in acquisition for five participants. All participants maintained skills learned for hour times, and two participants for half-hour and quarter-hour times. However, functional relationships between the tutoring program and learning progress were weak for all participants. For at least one participant, prior knowledge seemed to maintain throughout the study; for another student, data had been “extremely variable” from 0 to 100% during both baseline and treatment sessions. McKain also reported longer time intervals to complete each tutoring session (over 20 minutes for a 24 trial session), which greatly frustrated the participants and the classroom teacher who supervised the tutoring.

When discussing the results, McKain suggested that:

1) mixed clock times included in the sets of software should be replaced with hours, half hours, and quarter hours grouped in separate sets;

2) twenty-four trials in one session may not be appropriate for students of this age with similar disabilities. When the trial number was lowered from 24 to eight, McKain found that the tutoring process became “much smoother” and students could complete tutoring in the allotted time;
3) partners should not be arranged arbitrarily, since at least two participants in her study became very dissatisfied with their partners that strongly influenced their view of the peer tutoring program in general. Suggested solutions are student’s choice, rotating partners, or holding sociometric testing prior to partner assignment;

4) teaching students to graph their own results daily could have been included in the tutoring program, to lessen the teacher’s duty and to practice on related skills;

5) some pre-skills should be taught before tutoring sessions, such as counting from 1 to 12, counting by 15s, 10s, or 5s if more advanced times are to be learned.

In addition to the above suggestions, McKain reported how time consuming the tutor training had been. Teacher and students’ frustration also came from repeated training that lasted for one month (seven to 17 sessions with 20 minutes each). The difficulties encountered during training and later reflected in tutoring included a lack of the repertoire (certain terms, such as “thirty”, “forty-five”, were not in students’ vocabulary), and lacking proficiency in tutoring protocol (e.g., forget to praise or provide feedback; clicking the wrong button). To improve proficiency in tutoring protocol, McKain provided external reinforcement (i.e., a ticket system) to tutor for correct tutoring behaviors. However, it was included from the social validity interview that the students did not like the procedure of picking external reinforcers after tutoring. Regardless, McKain found that “lacking of using proper tutoring procedure did not
appear to adversely affect skill acquisition”. She also found that to train the tutors to a high level of compliance on their tutoring behaviors was not necessary for tutoring effectiveness. This is consistent with suggestions by Miller, Barbetta, and Heron (1994) that tutors do not need to reach 100% accuracy before or during tutoring.

Time consuming problem encountered by McKain in her study may also suggest a possibly longer speed (i.e., lower speed) in time-telling among the participants. “Acquisition of telling time skills” was defined in McKain’s study as number of correct responses made by students. Mastery of a time was when a student responded correctly for three consecutive days. No speed or proficiency data were reported. Students’ proficiency of the skill may need further investigation.

The peer-tutoring model created by Cooke et al. (1983) and Miller and Heron (1994) required the tutors to give assessments to the tutees after tutoring practice. During testing, tutees responded under the same stimulus control (e.g., same tutor, physical location, materials, content). What differentiates testing from tutoring is that the tutor does not provide feedback (e.g., “Good job”, “Try again”, tell the answer and prompt for repeat) after each response from the tutee. McKain’s study revised the procedure in that the researcher provided testing after tutoring sessions using different materials (i.e., a Large Judy Clock® that has distinctive differences from the clock face in the software used during practice). The testing validity may be debatable in that students might have been tested on their generalizing skills under stimuli different from instruction.

McKain’s report did not provide analysis of factors that might have contributed to a tutor’s competency problem (i.e., forgetting to provide feedback before move to the
next question; click on the wrong button). She expected that technology would solve the problem in future research when the computer could be programmed to reinforce tutor’s correct responses. With limited revision of the software program, this may hardly be achieved without a sophisticated sound detecting and analyzing system. The decision to provide testing by the researcher instead of the student tutor might come from considerations of the students’ abilities: 1) they already made mistakes in tutoring, and 2) testing procedure may confuse them. It is true that by using a software program with similar functions as that used in McKain’s study, introducing the testing procedure to students may worsen the proficiency problem, since tutors will be required to respond differently under tutoring and testing conditions. For tutoring, prompting and feedback are critical; while for testing, correct scoring is the crucial task and no feedback should be provided. Without installing a more advanced device, the software program may be enhanced to solve these problems by adding stimuli that can simplify the procedure and help the user discriminate between the two conditions. Possible modifications may include: a) using differential layouts of the software window between tutoring and testing conditions, and b) replacing the “Correct” and “Incorrect” buttons with a “Next” button, to simplify the tutoring procedure. That is, tutoring will not require the tutors to score.

Another finding in McKain’s report that warranted further modification of the program was incidental learning during training. For example, in pre-testing student 3 showed little knowledge of hour or half-hour, while he progressed to mastery level for hours and a close to mastery level for half hours during baseline when he was trained to tutor. It took him only three sessions to reach mastery level for half-hour times when
tutoring started. Student 3’s incidental learning during training, although desirable for the student’s learning, weakened the result to demonstrate a functional relationship between treatment and outcome. Such incidental learning could have been prevented if the software program used for training had not been the same with that used for tutoring treatment. That is, training could teach only the tutoring procedure and skills that are irrelevant to telling time. Certain modifications of the experimental design may also help to reduce students’ exposure to content to be learned before treatment is in effect. A potential design that could serve in similar study is a changing criterion design.

Changing criterion design was introduced to applied behavior analysis literature in two papers coauthored by Hall (Hall & Fox, 1977; Hartmann & Hall, 1976). The design requires a baseline observation of a single target behavior, which is followed by a series of phases of treatment. Each phase is one step-wise change in criterion towards the desired rate of the target behavior (Hartmann & Hall, 1976).

The critical element of changing criterion designs is the systematic introduction of a criterion level of performance over successive phases so that the behavior under study is gradually shaped into a final level. Experimental control is demonstrated in this design by the simultaneous occurrence of the behavior and the change of the criterion that replicates the therapeutic change (Cooper et al., 1987). A changing criterion design can be used to evaluate the effect of reinforcement or extinction contingencies. That is, in order to obtain access to the reinforcer, the individual must meet the required criterion. Criteria are determined on the basis of the baseline level of the performance. When performance meets or exceeds the criterion and establishes certain stability, the
performance demands are increased to a higher criterion. This process repeats until the final criterion level is reached.

Different from multiple probe design used in analyzing a shaping program for multiple behaviors, changing criterion design is best applied to evaluate the effect of instruction techniques on a single target behavior in terms of improving the behavior on stepwise changes in its rate, frequency, accuracy, speed, or latency (Cooper et al., 1987).

To simultaneously allow efficient performance improvement in students and demonstrate experimental control of the design is challenging; all changing criterion designs may be problematic (Tawney & Gast, 1984). The challenge lays in selection of length of phases, and size and number of criterion changes (Cooper et al., 1987). For a decision on length of phases, Hartmann and Hall noted that “each treatment phase must be long enough to allow the rate of the target behavior to restabilize at a new and changed rate; it is stability after change has been achieved, and before introduction of the next change in criterion, that is crucial to producing a convincing demonstration of control” (p. 531). Similarly, the size of the criterion changes should be “large enough to be detectable, but not so large as to be unachievable” (Cooper et al., 1987, p. 220). Thus the experimental control can be demonstrated in that performance changes co-occur with and adjust to new criteria. Stronger experimental control may require more behavior changes to meet a new criterion. However, the number of behavior changes required should not be arbitrarily assigned as wished, but should be determined from careful analysis of the interrelation of the three crucial elements. With a constant length of time for a study,
there is a negative relation between the length of phases and the number of criterion changes. That is, when the time for an experiment is limited, more sessions allowed to demonstrate the improvement and stability of the behavior change will require fewer number of criterion changes, which may therefore require larger size of the criterion changes, and vice versa.

For identifying the final level of the performance of a behavior, Cooper and colleagues especially warn the use of artificial ceilings or floors on the rate or frequency of possible performance, as such arbitrarily assigned criterion may limit the individual’s potential of performance.

Literature has demonstrated that a changing criterion design can be used to effectively increase or decrease target behaviors of individuals with disabilities in various settings (e.g., Inge, Moon, & Parent (1993) on increasing production rate at industry setting; Bates (1989) on vocational training of individuals with profound disabilities; Wolfe et al. (2000) on academic performance of elementary students with mild disabilities). Changing criterion design has been applied alone (e.g., Hall & Fox, 1977; Hartmann & Hall, 1976) or in combination with other experimental designs. Schleien, Wehman, and Kiernan (1981) applied a combination of multiple baseline across subjects and changing criterion design to teach adults with severe developmental disabilities darting skills. In another study by Levin and Carr (2001) dealing with food selectivity among children with developmental disabilities, a reversal design was applied during the functional analysis to identify the variables responsible for problem behavior that accompanied food selectivity. During intervention, a multiple baseline design across
participants was employed. Within one of the treatment conditions, a changing criterion design was used because the condition required progressive increments of food assumption.

As McKain suggested, teaching pre-skills before tutoring is necessary for telling time. Learning to tell time for children is not an easy task since a number of discriminations must be made when telling time (Stein, Silbert, & Carnine, 1997), including a) direction in which clock hands move, b) discrimination between minute and hour hands, c) discrimination between minutes, which are not represented by numerals on the clock, and hours which are represented by the numerals on clock, and d) discrimination of vocabulary (e.g., before, after) (p. 370). For research goals set for participants in McKain’s study, the first two discriminations are relevant. An unmentioned element that might also contribute to students’ difficulty is learning to tell the hour when the hour hand is not directly pointing to a numeral. Students must choose one numeral from the two at each side of the hour hand.

Purpose of the Study

This study was a replication of McKain’s (2004) study with major modifications. The purpose of the present study was to examine the effectiveness of a Computer-Assisted Tutoring System (CATS) package on students’ acquisition, maintenance, and generalization of time-telling skills and tutoring skills. The CATS package was similar to that used in McKain’s (2004) study, but with modifications in software design, procedure, and data collection to address problems raised in McKain’s study report. The CATS package consists of instruction sessions, peer-tutoring, and program software. By
using training software, students were first trained to use tutoring and testing software. Training software was functionally similar to that used in tutoring, but on non-time-related items. Upon completion of training, students did tutoring practice and testing on a daily basis. Tutoring practice lasted for a number of minutes determined by the researcher before the program terminated automatically. Students were instructed to do as many problems correctly as possible in practice. Tutoring testing was also timed according to the criterion or goal assigned for that session, however would not self-terminate when the criterion time was up. The number of correct answers (i.e., accuracy) and speed used in the testing were recorded by the software for printing or reading from screen. Students’ generalization and maintenance of the skills acquired during the study were also measured.

Research Questions

1. What are the effects of the Computer-Assisted Tutoring System (CATS) on the accuracy of time-telling skills to the half hour, quarter hour times, and five-minute increments (See Glossary) with elementary school students with emotional and behavior disorders?

2. What are the effects of CATS on the speed of time-telling skills to the half hour, quarter hour times, and five-minute increments with elementary school students with emotional and behavior disorders?

3. What are the effects of CATS on students’ maintenance of accuracy of time-telling skills?
4. What are the effects of CATS on students’ maintenance of speed of time-telling skills?

5. What are the effects of CATS on students’ generalization of time-telling skills to untrained conditions in terms of accuracy?

6. What are the effects of CATS on students’ generalization of time-telling skills to untrained conditions in terms of speed?

7. What are the participating students’ opinions of the use of CATS?
Glossary

Accuracy – The number of correct responses a student emitted in an assessment. The value of the number is between (while include) 0 and 12.

Five-minute increments – The learning content of a time telling skill practiced in this study that includes 144 possible time items, which involve random combinations of 12 possible hours (i.e. 1 through 12) and 12 possible minutes (i.e., 05, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, and o’clock). Example: 3:25, 8:40, 5:15.

Generality – Generality of time-telling accuracy and generality of speed. They are the number of correct responses emitted and the number of seconds consumed by a student when presented 12 pictures of clock faces in an assessment. The content of a generality assessment is the same as that in the last assessment for which the students had previously received tutoring.

Half hour times – The learning content of a time telling skill practiced in this study that includes the following 12 time items: 1:30, 2:30, 3:30, 4:30, 5:30, 6:30, 7:30, 8:30, 9:30, 10:30, 11:30, 12:30.

Hour times – The learning content of a time telling skill pre-tested in this study that includes the following 12 time items: 1:00, 2:00, 3:00, 4:00, 5:00, 6:00, 7:00, 8:00, 9:00, 10:00, 11:00, 12:00.

Maintenance – Maintenance of time-telling accuracy and speed. They are the number of correct responses emitted and the number of seconds consumed by a student in an assessment given three, six, nine, and in a few occasions, 12 days after the tutoring practice terminated. The content of a maintenance assessment is the same as that in the last assessment for which the students had previously received tutoring.

One-minute increments – The learning content of a time telling skill suggested in this study that includes 720 possible time items, which involve

---

1 Definitions apply to terms and conditions in this study only.
random combinations of 12 possible hours (i.e. 1 through 12) and 60 possible minutes (i.e., 01, 02, 03 etc., through o’clock). Example: 3:27, 8:41, 5:19.

One quarter hour times – The learning content of a time telling skill practiced in this study that includes the following 12 time items: 1:15, 2:15, 3:15, 4:15, 5:15, 6:15, 7:15, 8:15, 9:15, 10:15, 11:15, 12:15.

Speed – The number of seconds consumed or elapsed during a test that involves 12 time items. That is, the total time a student uses for each test.

The four time-telling skills – The four sets of time telling skills on half hour, one quarter hour, three quarter hour times and times-to-five-minutes.

Three quarter hour times – The learning content of a time telling skill practiced in this study that includes the following 12 time items: 1:45, 2:45, 3:45, 4:45, 5:45, 6:45, 7:45, 8:45, 9:45, 10:45, 11:45, 12:45.
CHAPTER 3

METHOD

This chapter will describe the methods to be used in the study, including the experimental and data collection procedures. The following are detailed descriptions of participants, settings, measurements of dependent variables, independent variables, experimental design, materials, procedures, procedure integrity and interobserver agreement.

Participants

Four first grade students identified with behavior disorders and who also had learning difficulties participated in the study. The participants were selected based on the following criteria: (a) nomination by the classroom teacher as demonstrating poor time telling skills, or students whose IEP had time-telling-related goals; (b) students who demonstrated appropriate language, social, and motor skills that allow them to interact with peers and manipulate computer controls; and (c) students who had consistent school attendance. The school principal sent out an informational letter (see Appendix A) to the parents of all the students in the classroom, which informed the parents of the purpose and procedures of the study, and asked the parents’ permission for their children to
participate. Parents who were willing to have their children participate in the study were able to contact the researcher for further information.

Each student and his or her parents or guardians who agreed to have their child participate in the study, received an informational letter and consent form (see Appendix A). The letter and form described the purpose, procedure, and requirements of the study, and asked the consent of the parents or guardians for their children’s participation in the study.

The four participants were between 7-3 to 8-0 years of age. Student 1 was the only female participant. All four students were diagnosed with emotional and behavioral disorders. Student 2 and 3 were also diagnosed with bipolar disorder. It was known that some participants took medicine that were disorder related, however the information was not available to the experimenter. All participants demonstrated appropriate fine motor skill and language abilities, which allowed them to use the mouse easily and to verbally communicate with their partner. The tutoring pairs were arranged by the classroom teacher based on the participants’ social compatibilities.

Settings

The study was conducted in a special education resource classroom of a public suburban elementary school located in a Midwestern city. All sessions were conducted in this classroom. The tutoring sessions were carried out in a corner inside the classroom where activities within the corner were visually separated from the class to reduce possible distractions. The corner was well lit, surrounded on three sides with five foot high shelves of curriculum materials. The activity area inside the corner measured
approximately 10ft × 6ft. A table (4 ft × 2.5 ft) and two chairs were placed in the middle of the area. When the study was not conducted, the corner was typically used by the teacher aid to provide individualized instructions. During tutoring, one pair of participants was sitting next to each other in front of a laptop computer sitting on the table. The position of the computer allowed both participants to see the screen clearly and the tutor to have unimpeded ability to manipulate the mouse. Observers sat at a distance behind the pair to be able to observe both tutor behaviors and cursor movements on the computer screen. The study sessions were conducted daily for 30- to 45-minute in the morning between 11 AM and noon.

Definition and Measurement of Dependent Variables

The time-telling skills in this study were defined as the child’s ability to accurately and quickly tell time to the half hour, one quarter hour, three quarter hour times, and to five-minute increments (See Glossary for definitions). Henceforth, the skills will be referred to as the “four time-telling skills”. The dependent variables measured in this study were (a) accuracy of time-telling skills, (b) speed of time-telling skills, (c) maintenance of accuracy on the four time-telling skills, (d) maintenance of speed on the four time-telling skills, (e) generalization of accuracy of the four time-telling skills, (f) generalization of speed of the four time-telling skills, and (g) participant opinions of CATS.

Accuracy of Time-telling Skills

Accuracy of time-telling skills was defined as the number of correct responses a student emitted during an assessment where the student is presented with a clock showing
different times. Accuracy of time-telling skills only applied to tutoring and testing phases, but not relevant to the tutor training phase. During tutor training, no time-related skills were practiced. A correct response for an hour time was defined as when presenting a time on screen, the participant says the correct hour followed by “o’clock”. For example, when a time appears on screen with an hour hand (i.e. shorter hand) pointing to number eight and minute hand (i.e., longer hand) pointing to number 12, the correct response for this time is the participant saying “eight o’clock”. It was not considered a correct response if the participant names other numbers, for example “six o’clock”. Neither was it considered correct if the participant says “eight” without “o’clock”.

A correct response for half hour times was defined as when presenting a clock showing time was presented on screen, the participant says the correct hour followed by “thirty”. For example, when a time appears on screen with hour hand pointing between the numbers three and four, and minute hand pointing to number six, the correct response for this time is the participant saying “three thirty”. It was not considered a correct response if the participant names other numbers as hour, for example, saying “four thirty”. Neither was it considered correct if the participant says “three” without “thirty”, or, “half past three”.

A correct response for quarter hour times was defined as when a clock showing time was presented on screen, the participant says the correct hour followed by “fifteen” for one quarter past the hour and “forty-five” for three quarters past the hour. For example, when a time appears on screen with hour hand pointing between number nine and ten, and minute hand pointing to number three, the correct response for this time was
the participant saying “nine fifteen”. Any variation was not considered as incorrect, including saying “fifteen past nine”. A correct response for times-to-five-minutes was defined as when presented with a time on screen, the participant says the correct hour followed by the correct minute number. For example, when a clock appears on screen with hour hand pointing between number nine and ten, and minute hand pointing to number five, the correct response for this time is the participant saying “nine twenty-five”. It was not considered a correct response if the participant names other numbers as hour or minute, for example, saying “ten five”. Neither was it considered correct if the participant says only “nine” or only “twenty-five”, or, with slight possibility, the participant utters “twenty-five past nine”. All other verbal expressions different from the response heard from the earphones were treated as incorrect and counted accordingly.

**Speed of Time-telling Skills**

Speed of time-telling skills was defined as the number of seconds consumed or elapsed during a test, that is, the time from the beginning to the end of a test. The session time begins when the “Start” button is clicked which initiates computer presentation of the first test item. Beginning with the first test item a clock is presented after the tutee responses the tutor clicks the on screen “Next” icon which prompts the presentation of the next test item. Timing stops contingent upon clicking the “Next” button for the last clock or test item in the session. Since the computer program for testing records time consumed during a test and shows accordingly on the result page and printed sheet, the speed of a time-telling test could be copied from the screen or found on the printed sheet under “Speed” (see Appendix B for a sample of printed result sheet).
Maintenance of Time-telling Accuracy

Maintenance was defined by Cooper and colleagues (1987) as “the extent to which the learner continues to perform the target behavior after a portion or all of the intervention has been terminated” (p. 558). Maintenance of time-telling accuracy was defined as the number of correct responses a student emitted in an assessment given three, six, nine, and in a few occasions, 12 days after the tutoring practice terminated for each of the four time-telling skills. The content of a maintenance assessment was the same as that in the last assessment for which the students had previously received tutoring. For example, three days after the last tutoring session of telling half hour times, if a student correctly responded to eight out of 12 half hour times in the maintenance test, the student’s maintenance accuracy on half hour times was recorded “8”.

Maintenance of Time-telling Speed

Maintenance of time-telling speed was obtained simultaneously with maintenance of time-telling accuracy during a maintenance test. Maintenance of time-telling speed was defined as the number of seconds consumed or elapsed during a an assessment given three, six, nine, and in a few occasions, 12 days after the tutoring practice terminated for each time-telling skill. The content of a maintenance assessment was the same as that in the last assessment for which the students had previously received tutoring. For example, six days after the last tutoring session of telling one quarter hour times, if it took a student 35 seconds to complete all 12 time items in the assessment, the student’s maintenance speed on one quarter hour times was recorded as “35 seconds”.

53
Generality of the Accuracy of the Four Time-telling Skills

Each student was assessed on stimulus generality of the time-telling skills practiced and tested in previous sessions. Stimulus generality is defined as “the extent to which the learner improves her performance of the target behavior in environments different from the original training environment” (Cooper et al., 1987, p. 556). The environmental difference in current study was presenting times by showing pictures of a variety of clock and watch faces, instead of presenting unvarying clock faces on a computer screen.

Generality data on all four skills were collected on the same days when maintenance assessments were given. Generality of time-telling accuracy was defined as the number of correct responses a student emitted when presented 12 pictures of clock faces during a generality assessment. The content of a generality assessment was the same as the content in the last assessment following a tutoring session. In other words if the last tutoring session involved telling time to the half hour then the generality session contained clock times to the half hour. The participant’s responses were judged as correct or incorrect based on the same standards applied during computer tutoring and testing. Any deviations were considered incorrect and recorded accordingly. The scores of accuracy obtained in generality assessments were recorded in the same manner as that in maintenance assessments.

Generality of the Speed of the Four Time-telling Skills

Generality of time-telling speed was obtained simultaneously with generality of time-telling accuracy during a generality assessment. Generality of time-telling speed
was defined as the number of seconds consumed or elapsed during a generality assessment that involved presenting pictures of 12 clock faces to the student. The speed score was obtained by using a stopwatch. The scores of speed obtained in generality assessments were recorded in the same manner as that in maintenance assessments.

**Participant Opinions**

Participant opinions were gathered upon completion of the study to examine the social validity of applying CATS. A questionnaire was used to collect participants’ opinions of the program (see Appendix C). Questions inquired about participants’ views of the peer tutoring format, computer software, learning progress during the study, and the participant’s other experiences with peer tutoring and computer. The participants were asked about their overall satisfaction level with the project. The questionnaire was completed at home while participants were instructed to obtain parents’ assistance for spelling and expression. Addressed and stamped envelopes were provided along with the questionnaire.

**Independent Variable**

The independent variable measured in this study was the tutoring program package consisting of: a) a peer tutoring system, and b) a set of software that is capable of presenting clock faces or images one at a time and recording user’s performances in terms of mouse-clicking and time elapse.

**Peer Tutoring System**

The peer tutoring system was adopted from the system developed at The Ohio State University by Cooke, Heron, and Heward (1983). Materials and procedure in this
form of a peer tutoring program is carefully designed and arranged by the teacher. Students of similar age or performance level are paired as tutoring dyads. Tutors are trained to praise or provide corrective feedback to tutee following a standard feedback pattern. Tutors also pre- and post- test a tutee’s progress. The major difference in current peer tutoring study comparing to previous tutoring programs was the application of a computer-assisted program that include training, tutoring, and testing.

*Computer-assisted Tutoring Program*

Five sets of computer software were developed in Macromedia Flash MX® by the researcher. The purpose of each set was for: a) training how to provide tutoring, b) training how to give test, c) pretest, d) tutoring, and e) testing (See Table 3.1 for the summary of the software’s components and features. Also see Appendix D for sample illustrations of the software window). The functions and features of the software are as following:

1) All possess a loudspeaker button that when clicked repeats the time or name of the item appearing on the screen. All possess an initial page with a “Start” button; clicking it launches the program and timing. All have 12 items, time or non-time-related images;

2) All software with training purpose presents images of objects (e.g., a rubber duck, school bus) that are not time-related. The objects were selected through interviews with the students. The images were selected to guarantee that the participants were familiar with the objects and could name them instantly. The images were not for entertaining or rewarding purpose, and might have

56
appeared childish for the participants. The names of the objects were all consisted of two words (e.g. Christmas tree, school bus). All other software presents series of clock faces one at a time. Times to be tutored were following the sequences of hour, half hour, quarter hour times, and times-to-five-minutes (See Table 3.2 for the schedule of the times tutored in each set);

3) Software for tutoring purpose, training or real sessions, has a timing bar at the bottom of the window. Before the tutoring starts, the length of tutoring practice can be set by typing in a desired number of minutes. Upon the start of tutoring, the timing bar shrinks proportionately with time passage. Tutoring is automatically terminated when the pre-set number of minutes passes. All tutoring software has “Next” button to proceed to the next time or object image. All tutoring software do not produce scores, thus do not have printing option;

4) Software for testing (daily assessment and maintenance) purposes, training or real sessions, has a timing bar of the criterion time (goal) set for that session and participant. The program, however, will not terminate automatically when the goal time is up and timing bar shrinks to “0”. Testing software also has an item counter bar of the same length of timing bar. The bar shrinks proportionately with items completed. The program terminates when the last item is completed; that is, when the item counter bar reaches “0”. Testing software proceeds to the next item by clicking one of the two buttons, a green
“Correct” button and a red “Incorrect” button. The results can be printed by clicking a “Print” button on the result page when testing is completed.

5) Pretest software simulates the appearance and functions of testing software, except that timing bar and item counter bar are not included.
<table>
<thead>
<tr>
<th>Title</th>
<th>Content</th>
<th>Speaker</th>
<th>Buttons</th>
<th>Timing Bar</th>
<th>Question Counter</th>
<th>Print</th>
<th># of Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training-tutoring</td>
<td>Images</td>
<td>Yes</td>
<td>‘Next’</td>
<td>Set as desired</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Training – assessment</td>
<td>Images</td>
<td>Yes</td>
<td>‘Correct/Incorrect’</td>
<td>Criterion</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pretest</td>
<td>Time</td>
<td>Yes</td>
<td>‘Correct/Incorrect’</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Tutoring</td>
<td>Time</td>
<td>Yes</td>
<td>‘Next’</td>
<td>Set as desired</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Assessment</td>
<td>Time</td>
<td>Yes</td>
<td>‘Correct/Incorrect’</td>
<td>Criterion</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Time</td>
<td>Yes</td>
<td>‘Correct/Incorrect’</td>
<td>Criterion</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Generality</td>
<td>Time</td>
<td></td>
<td>Use pictures in stead of software program to assess</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. 1. Summary of the components and features of the computer software.

*Participants’ Tutoring Behaviors*

Tutors’ behaviors included prompting, providing feedback (except during testing) for each question, during testing clicking the “Correct” and “Incorrect” button for each question according to the tutee’s answer, typing in names, and clicking the “Start” button to initiate program. Tutors’ behaviors during tutoring and testing were observed and
scored by the researcher and a second observer in approximately 30% of the tutoring sessions (See Appendix E for scoring form used for observation). The tutoring behaviors occurred successively with no overlapping, which allowed observers to observe one behavior at a time.

**Experimental Design**

A combination of changing criterion (Cooper, Heron, & Heward, in press; Hartmann & Hall, 1976) and a delayed multiple baseline (Cooper et al., in press) across skill design was used to assess the effects of the Computer-Assisted Tutoring System (CATS) on the accuracy and speed of the participants’ acquisition of time-telling skills on half hours, quarter hour times, and time-to-five-minute.

The operation of changing criterion design requires a baseline observation of a single target behavior, which is followed by a series of treatment phases. Each phase is one step-wise change in criterion towards the desired level of the target behavior (Hartmann & Hall, 1976). In this study, for each participant, a changing criterion design was applied to each time telling skill. In order to obtain access to the presumed reinforcer (i.e., beat the timing bar), a participant had to meet the assigned criterion (i.e., goal). When participants’ performance met or exceeded the criterion and established stability, the performance demands were increased to the next criterion (i.e., a higher goal). This process was repeated until the mastery was reached. Through a systematic introduction of a higher criterion performance level by raising the accuracy and speed, the target behavior of telling half hour time, for example, was gradually shaped towards the mastery of the skill. In this case, experimental control was demonstrated when the
accuracy and speed of telling time occurred simultaneously and repeatedly with the change of new criterion, and had improved towards the direction of mastery.

A multiple baseline design was used across the time-telling skills. Data collections on the time telling skills sequenced, with three to 14 days between each skill, depending on the stability of data and time allowance in school. Changes of experimental condition were made for only one skill at a time while conditions for the other skills were held constant. Interventions on the other skills were not started until stability was observed in data for the first time-telling skill. Experimental control was presented when changes in accuracy and speed occurred when and only when the intervention was manipulated by the researcher.

Materials

Materials included electronic and manual equipment. Electronic equipment consisted of a laptop computer with a Windows XP® operating system, a pair of earphones, one out-connected mouse, and a set of software in executive format (i.e., with document type ending “.exe”) for Windows. The set of software was developed by the researcher using Macromedia® Flash MX®, and required no plug-ins to operate on the laptop computer to be used (See Appendix D for sample illustrations of the interactive windows of the software). Also see Table 3.1 for the summary of the set of software’s components and features. Manual equipment included a) Forty-eight laminated color pictures of various clock faces of half hour, one quarter hour, three quarter hour times, and times-to-five-minutes, b) interobserver agreement forms (see Appendix E and F), c) a clock, and d) a stop watch. A printed image (Figure 3.1) was placed at the left side of the
computer for Student 3 who repeatedly misnamed number six, eight and nine. On the image, number “8” was spelled “ei-t” since Student 3 could not read the word “eight.” Student 3 read words “nine” and “six.”

Figure 3. 1: Image used for Student 3.
Procedures

*Comparing Group Level Measurement*

Before the initial session of this study, eight typically developing first grade students attending the same school with the participants were assessed on their ability to tell time to the hour, half hour, quarter hour, and times-to-five-minutes using the pretest software sets. The averaged results of their accuracies and speeds on these time-telling skills were calculated to decide the comparing group level on each skill. Comparing group levels were then referenced when the researcher assigned final goals for each participant to achieve during the tutoring intervention.

The comparing group level of a particular time telling skill was defined in this study as the averages of accuracy and speed by the eight typically developing first graders (i.e., children without disabilities). A comparing group level was recorded in the form of the average correct responses students emitted and the average number of seconds needed to complete the test. Average of accuracy was calculated by adding the accuracy scores of the students without disabilities and then dividing the sum by eight. Average of speed was calculated in the same way. For example, if on average the eight students without disabilities scored 10 correct items on a half hour time pretest, and on average it took these students 35 seconds to complete the test, then the Comparing group level of half hour time skill is recorded as “10 corrects in 35 seconds”.

The eight first grade students without disabilities were selected randomly based on the teacher’s recommendations as students with above average academic performance
in the general education classroom. Each student was tested individually by the researcher using pretest software on hour, half hour, quarter hour times, and times-to-five-minutes. The same procedures were used subsequently with the tutees (i.e., participants with behavior disorders). Accuracy and speed scores of each participant during each test were recorded.

*Tutor Training*

All the participants were trained to provide tutoring and testing using the model-lead-test procedure (Heron, Heward, Cooke, & McGill, 2000) and the computer-assisted software. The model-lead-test procedure was developed based on behavioral principles and multiple effective teaching practices. Using scripted lessons, training through the model-lead-test procedure clarified goals or expectations on the trainees, presented examples and non-examples of sequenced skills, and provided students with extensive practice; also, the trainer monitored the training and practicing process with corrections and feedback.

Tutoring skills trained in this study include typing in names of tutor dyad, starting tutoring or test, prompting tutee’s response, clicking the loud speaker button to hear an answer repeated when necessary, providing feedback, clicking the “Correct” or “Incorrect” button based on tutee’s responses. Of these skills, providing feedback is one important element. The tutors were trained to provide feedback contingent upon tutee’s response. Upon correct responses, they were trained to praise their tutee (e.g., “good job”, “correct”); for incorrect response, tutors were instructed to prompt a second try from the tutee, and to provide the answer when a second mistake was made.
The tutoring skills were taught using the tutor training software through skill demonstration, modeling, group role-playing, and student practice and testing. Tutor training was conducted in a series of four sessions with approximately 25 minutes a session, depending on students’ comprehension process. The content and scripts used in each training session were modified based on McKain’s study (2004).

Training Sessions 1. In this session, training was primarily on how to tutor. Training software for tutoring was used. This session included 1) a brief orientation to the computer tutoring program, 2) modeling of using training-tutoring software by the researcher and a participant, and 3) each participant’s role-playing as tutor using the software while the researcher as tutee. (See Sessions 1 in Appendix G for the script used in this training session.)

Training Sessions 2. In this session, training was focused on giving the test. Training software for testing was used. The speed bar was set at a criterion of 36 seconds, with an average of 3 seconds allowance for each question (i.e., $12 \times 3 = 36$). This session included 1) reviewing content learned in a previous session, 2) modeling of using training-testing software by researcher and a student volunteer, and 3) role-playing by each tutor giving a assessment to the researcher. (see Session 2 in Appendix G for the script used in this training session.)

Training Sessions 3. This session was a comprehensive practice using the training-tutoring and testing software by each pair of participants. The researcher started with a review of both tutoring and testing sessions. Participants then practiced in pairs to take turns as tutor and tutee. When participants practiced, the researcher gave corrective
feedback when mistakes occurred on tutoring behaviors. The researcher recorded tutors’
behaviors in one tutoring and testing trial using the IOA form (Appendix E). Only the
“feedback providing” behaviors were observed in a tutoring session, and only “button
clicking” was observed in testing session. As McKain (2004) suggested that lack of
implementing tutoring procedure at a not so high accuracy (mean of 70% to 79%) did not
appear to affect skill acquisition, a participant in current study was considered having
mastered the tutoring and test administration skills when the accuracy of the observed
behaviors reached or exceeded 80% (See Session 3 in Appendix G for the script used in
this training session).

Training Session 4 procedures were the same as Session 3. Participants continued
practicing tutoring and testing skills using the training software.

*Pretest*

Following the completion of the tutor training session, the participants were
assessed on all five groups of time telling skills (i.e., hour, half hour, one quarter, three
quarter hour times, and times-to-five-minute). For each participant, an assessment was
conducted by the researcher using the pretest software for hour times. The accuracy and
speed scores collected in the pretests on each time group were recorded as the baseline
skill level of the participant in corresponding time group. These data, combining with
Comparing group level data, were used to determine the student's criterion levels in the
following tutoring sessions. See section *Go setting (criterion determination)* for the
criterion calculating method.
Before a pretest, the researcher instructed the participants about the purpose of the test: to find out how well their tutee could tell time, instead of naming images (e.g., Christmas tree, school bus) as in the tutoring sessions. Except for the instruction, the procedure remained the same as in training. In a pretest, the paired dyad sat in front of the computer, with the tutor on the right and the tutee on the left. On the screen of the computer, the pretest window had been opened. Different from the tutoring or daily assessment software, there was no timing bar or item counter in a pretest. That is, no speed goal (criterion) was set, and the participant were not aware of the time elapsing and the number of items completed during a pretest. Both tutor and tutee typed in their own names. The tutor then put on the headphones and clicked the “Start” button. The first clock screen appeared. The tutor then prompted the tutee, “What time is it?” The tutee responded by saying a time. With no verbal feedback, the tutor clicked the “Correct” or “Incorrect” button as determined by whether the tutee’s response was correct or incorrect. If the tutor could not decide because he or she had failed to hear the tutee’s response, the tutor re-prompted the tutee by saying “Say it again”; if the tutor had missed the time voiced by the software, he or she could click the loud-speaker button to hear the time repeated in his or her headphones. A click of either the “Correct” or “Incorrect” button would automatically transition the program to the next clock face. On the completion of the last clock face in the test, the scores of both accuracy and speed the tutee achieved in the test appeared on screen for the researcher to record. The five sets of pretest software were used to obtain each participant’s baseline performance in each time group. The pretests were given in an order following the schedule in Table 3.2. Baseline data
collection using the pretests began one after another, with three to 14 days in between, depending on the stability of data and time allowance in school.

*Goal Setting (Criterion Determination)*

For each student, his or her accuracy and speed scores in pretests were the baseline criterion; the Comparing group level of the time telling skills measured with students without disabilities (i.e., typically developing first graders) were the presumed and targeted ceiling. For each time group, four criterion changes were assigned and individualized for each participant. Thus the magnitudes of the criterion change were calculated as the baseline criterion subtracted from the ceiling criterion then divided by 4. The result was rounded down to an integer (i.e., whole number). For example, if a student’s speed baseline and the final are 60 seconds and 35 seconds respectively. The magnitude of the speed criterion change will be \((60-35) / 4 = 6.25\) rounded down to 6 seconds. Thus, the first criterion level in treatment for the student is 54 seconds (i.e., 60-6 = 54), and the second is 48 seconds (i.e., 54-6 = 48), and so forth. The calculating method for accuracy criteria was the same, except the ceilings were all assigned as 12. That is, the ceiling goal for practicing each time telling skill was to recognize all 12 items in one test. For a student whose accuracy baseline was 0, the criterion was therefore 3 (i.e., \((12-0) / 4 = 3\)). Thus, the first criterion was 3 (i.e., 0+3 = 3), the second criterion will be 6 (i.e., 3+3 = 6), and so forth. Participants’ daily assessment results were used to compare to the criteria. Participants worked on one criterion of accuracy and speed for one set of time skill and would not start another criterion until they met the criterion of both accuracy and speed for consecutive three days.
The criterion level assignment was not carried out absolutely strictly as designed above. Unexpected factors adjusted the criterion assignment in such a way that some participants completed one time telling skill through three or five criteria, and most participants were added with one more criterion as ceiling, since their data showed potential to use less time to complete a test (i.e. higher speed) at the completion of each time skill comparing to the Comparing group level ceilings assigned originally. Accuracy criteria were also adjusted according to each participant’s progress. More or less criteria were assigned on an individual basis. Trials were added using pretests to adapt to certain situations (e.g., student demonstrated emotional stress when being timed). Criterion adjustments are discussed in detail in Chapter 5.

*Instruction on Time-telling*

Upon the completion of the pretest, the researcher provided a 20-minute group instruction on a) recognizing and counting number from 1 to 12, b) the direction to which the clock hands go, c) discrimination of hour hands and minute hands, and d) reading an hour when the hour hand was not strictly pointing to a numeral. A real clock was used when presenting (See Appendix H for scripts to be used in this section).

*Tutoring Session*

At the completion of the group instruction session, students were paired as two dyads working in pairs on daily tutoring practice. Different from pretest and daily assessment software, the software for practice allows the researcher to set a desired length of time for practicing. The daily practice time for each time skill was set at one minute. When practice started, a timing bar appeared at the bottom of the program
window indicating time elapsing. When one minute passed, the program terminated automatically. Another feature in tutoring software different from that in Pretest and Daily assessment was the use of a “Next” button to proceed instead of “Correct” or “Incorrect” buttons. That is, during tutoring practices, tutors clicked the “Next” button to go to the next time item No score was recorded. Twelve different times in random order was presented. If the dyad completed all 12 times before one minute ended, the program repeats presenting the same 12 times in different random orders until the one minute session ended.

Each tutoring session began with one dyad sitting on the chairs in front of the computer. The researcher reminded the dyad to say the times as many times as they could before the orange bar (i.e. timer) shrank to zero. The procedure was the same with that used in Pretest sessions except that tutor provided feedback to the tutee. That is, after the tutee responded to a time item, the tutor provided feedback by saying “Good job” to a correct answer, and “Try again” to an incorrect answer. If the tutee uttered a wrong answer for a second time, the tutor would tell the answer to the tutee, and prompt the tutee to repeat it, “Say (time)”. The tutor could hear the answer from the earphones that he or she wore. When software terminated after the one minute practice, the tutor tested the tutee on the time skill practiced using the assessment software (See the following section Daily Assessment for testing procedure). After the tutee completed all the practice and testing required for him or her for that day, the dyad switched roles and seats, as the tutor would always sit on the right hand. The dyad repeated the tutoring practice switching roles.
The tutoring practice session allowed one dyad to work on the computer. In the meantime, the other participants and students who did not participate in the session worked on their regular school work with their teacher. The whole session was monitored by the researcher, who provided corrective feedback to the tutors’ behaviors and responses.

<table>
<thead>
<tr>
<th>Set # (in order)</th>
<th>Quantity</th>
<th>Contents</th>
<th>Example (Description)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>Hour times</td>
<td>Two o’clock (2:00)</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>Half hour times</td>
<td>Two thirty (2:30)</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>One quarter times</td>
<td>Two fifteen (2:15)</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>Three quarter times</td>
<td>Two forty-five (2:45)</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>Times-to-5-minutes</td>
<td>Two forty (2:40)</td>
</tr>
</tbody>
</table>

Table 3. 2. Schedule of time-telling skills.

*Daily Assessment*

Each time a tutee completed practice on one set of time skill, the researcher opened the assessment software for that time telling skill. The test consisted of the same time telling skills practiced in tutoring practice session, only the items were presented in a different order. The features of test software are the same with those in the Pretest session, except that a timing bar and an item counter appeared at the bottom of the
window (See Appendix D). The timing bar, appearing in red color, shrank evenly with the elapsing of time to indicate the remaining time. The item counter, appearing in blue color, shrank one twelfth (1/12) of its length for each item to indicate the number of questions yet to be answered. The time assigned for timing bar in Daily assessment was the goal (or criterion level) set for that phase. That is, if 48 seconds to complete all items was the goal for a phase, the red timing bar disappeared 48 seconds after the test started. Different from the orange timing bar used in tutoring practice, the testing did not automatically terminate when time ran out, that is the test continued even if time elapsed. That is, untested times continued to appear until all 12 time items were tested.

The tutee was instructed to 1) answer all the time questions as fast and correctly as possible, and 2) meet the timing goal by referring to the two bars – to keep the blue bar runs at least on the same pace with the red timing bar to “beat the timing goal”. At the completion of test, the program showed the result page, which included names of the tutor and tutee, date, number of correct answers, number of seconds used. The researcher would print the result sheet or record the scores.

Trial

With all four participants there was a one session trial inserted between baseline conditions and tutoring intervention throughout the skills tutored. This unique condition was different from baseline in that tutoring was initiated. This trial condition was also different from the forthcoming first criterion level sessions in that the criteria (i.e., goals) assigned for accuracy or speed were different. To say in another way, trial condition was the first tutoring session in which a participant was tutored to reach untested goals – goals
assigned based on their baseline performance. However, those goals were immediately adjusted in the next tutoring session if the goals applied in the trial were not appropriate. The new criteria were calculated in the same methods as the original criteria, except that the baseline averages were replaced by the trial results. For example, if a participant’s average baseline speed on half hour times is 50 seconds, the participant’s speed criterion in the trial is then assigned to 45 seconds. However, during the trial the participant spent 70 seconds in the daily assessment, a time length that is not comparable to the original criterion (i.e., 45 seconds) based on baseline performance. In the next tutoring session, the speed criterion is then adjusted to 62 seconds instead to adapt to the changed condition. This difference too large to ignore between baseline and first trial of tutoring became apparent when during baseline pretests the participants tended to respond “I don’t know” with shorter delay when presented with unknown time items, comparing to a longer delay when the tutoring intervention started.

Maintenance Assessment

For each time tell skill, a maintenance assessment was given three, six, nine, and in a few cases, 12 school days after the participant reached the comparing group level or upon the completion of the study due to limited time allowance. That is, three school days after a participant met the ceiling criterion of both accuracy and speed for consecutive three days, a maintenance test was administered. Another three school days after the first maintenance test, a second maintenance test would be given, and so forth. If by the time such a test could not be provided (e.g., student absence), the test would then be given the first day when the student was available. The assessments were given by the
researcher using the pretest software. The procedure of maintenance tests was the same as
pretests.

Generality Assessment

Generality assessments were given on the same days when maintenance tests were given. The assessments’ contents were the same with the maintenance tests. The procedure used in a generality assessment was the same as that in a pretest, except that instead of computer software, the researcher worked with each participant individually using laminated color pictures of clock faces of times. The researcher started a stopwatch to time the assessment, then presented to the students with the pictures of time items one after another. Pictures to which the participant responded correctly were laid in one pile, while pictures to which the participant answered incorrectly were placed in another pile in front of the researcher. Upon the completion of the last picture, the researcher stopped the stopwatch and recorded results of the number of seconds used and counted the number of pictures the student answered correctly and incorrectly.

Procedural Integrity and Interobserver Agreement (IOA)

Interobserver agreement on tutoring procedure and assessment results was recorded by the researcher and one observer. Agreement scores were calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Integrity and IOA on Computer’s Program

The researcher served as one observer to record the procedure of the computer program and compared the record with the results recorded by the computer.
Approximately 30% of all sessions were randomly selected and observed. Sessions under observation covered tutor training, and daily assessments. In this process, the computer was regarded as the first observer of its own procedure. The computer showed the scores of speed and accuracy on the result page when a test was completed. Those scores could be printed out if desired and were compared to results recorded by the researcher. With tutoring or tutoring training software that do not produce a score, the researcher recorded the length of the practicing time and compared it with the pre-set practicing time (usually one minute).

During a test, the researcher sat behind the tutoring dyad, held a stop watch, recorded and tallied on a blank recording sheet (See Appendix F for recording sheet). Components to be recorded included: number of correct answers as recorded by the tutor, number of seconds consumed (i.e., speed) between the tutor clicked “Start” and clicked the last “Next” button, whether the program terminated upon the disappearance of the timing bar, and the speed that the timing bar (if there was any) had run. The records were compared using the interobserver agreement sheet for computer procedure (Appendix F). An agreement on one component was scored when both the researcher's record and the results recorded by the computer had exactly the same content. For example, after one assessment session, if the researcher recorded that the "Correct" button had been clicked by the tutor seven times, while according to the result by the computer the "Correct" button had been clicked six times, the component of "number of correct responses as recorded by the tutor" would be considered not being agreed on. Due to unavoidable errors that could be made by human (i.e., the researcher), any component that concerns
time (e.g., speed of a timer during tutoring practice, speed criterion level set for particular assessment session) allowed a difference within 1 second between the observers. For example, if the speed criterion for one participant in one assessment session was 45 seconds, agreement was still considered if the result of the computer record was 45 seconds while the record by the researcher showed any number within the range of 44 to 46 seconds.

*Integrity and IOA on Tutor’s Behavior*

Researcher and a trained second observer collected data on participants’ tutoring behaviors to determine the validity of the treatment. The second observer was a first year doctoral student enrolled in College of Education at The Ohio State University. The second observer was trained by the researcher, who served as the primary observer, of the definitions of each component to be observed. Tutors’ behaviors during tutoring and testing were observed and scored by both observers in approximately 30% of the tutoring sessions (See Appendix E for scoring form used for observation). The components observed included: providing feedback during tutoring and no feedback during testing, and clicking “Correct” or “Incorrect” button based on judgment. Prior to the study, the researcher worked with the second observer on the definitions of the observed components until both could correctly identify and record tutor behaviors with 100% accuracy. The observers recorded data using the Interobserver Agreement (IOA) Data Collection Form (Appendix E) during training and real-time scoring.

During tutoring sessions, the researcher and the second observer independently recorded on the IOA Form (Appendix E) whether or not each component of the tutoring
and testing was implemented by the tutor as he or she was trained. An incorrect tutor response was marked as an “X” in the corresponding cell in the form, while correct response would leave the cell blank. A component was scored correct when an observer decided that each behavior within the component had been implemented as described in training. For example, during a tutoring session, after tutee responded to the tutor’s prompt, tutor’s praise of “Good job” or “Right” to an correct answer was scored as correct and leave the cell blank. However under the same situation during a testing session, if the tutor still provided feedback, the cell would be marked “X” as incorrect tutor response. An agreement on one component cell would be scored when both observers left the cell blank or a mark of “X”.

77
CHAPTER 4
RESULTS

This chapter reports the results from the study. Interobserver agreement (IOA) results are presented. Second, treatment integrity data are presented with regard to measures of tutoring and testing behaviors of the tutor, and of the functions of the computerized program. Third, data collected during baseline, treatment, maintenance, and generalization conditions are presented for each participant. Finally, the results of the social validity questionnaire given to the participants are summarized.

Interobserver Agreement

Interobserver agreement was obtained for daily tutoring and assessment results, and for tutoring procedures. Agreement scores were calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Interobserver Agreement on Tutor’s Behaviors

Two graduate students and the researcher collected interobserver agreement (IOA) data on 30% of the tutoring sessions. All sessions with IOA included agreement data on tutor’s tutoring and testing behaviors. Total IOA ranged from 80% to 100% with a mean of 96.4%.
Interobserver Agreement on Program Procedure

The researcher and the computer program collected interobserver agreement (IOA) data on 33% of the tutoring sessions on the computer program’s procedure. All sessions with IOA included agreement data on 1) recording the results of the accuracy (i.e., the number of correct answers as recorded by the tutor), 2) recording the speed, and 3) functions of the bars (i.e., timing bar disappears when goal speed ends, and item bar shrinks one twelfth in length after completion of each item). IOA data on computer program’s accuracy and functions of the bars maintained 100% throughout all observed sessions. IOA data on the program’s time spent were calculated by dividing the number of agreed seconds (between computer program and the researcher’s observation, whichever is smaller) by the total number of agreed plus disagreed seconds and multiplying by 100. Total IOA ranged from 95.8% to 100% with a mean of 99.9% (See Table 4.1).
Table 4.1. Percent interobserver agreement on procedure integrity across participants. Numbers in parentheses indicate number of sessions scored for IOA.

### Procedure Integrity

Analysis of procedure integrity data shows that the computer program recorded the scores of accuracy with 100% accurate rate. The program also functioned on 100% accuracy to present all 12 time items of each time category in each tutoring and testing session, as well as presented a timing bar shrinking proportionally. There were a few discrepancies on recording the scores of speed between the observations by the researcher and by the computer program. In total, these differences were less than 1% of all recorded speeds (See Table 4.1).
Procedure integrity data of the participants’ behaviors when providing tutoring to their partners are on average all above 90% in total across participants, however, with various ranges. Student 1 and 2 each had four sessions when their tutoring procedure accuracy fell below 80%; Student 3 and 4 each had three sessions. Student 1’s occurrences of low procedure accuracy were observed in early days of the study, while such low procedure accuracy of Student 2, 3 and 4 occurred at different times throughout the study. Student 1 and 3’s lowest tutoring accuracy was 57.1%, Student 2 and 4’s lowest tutoring accuracy was 66.7%, while all participants could provide tutoring and testing with 100% accuracy in some sessions. See Table 4.2 for the participants’ tutoring procedure integrity data.
<table>
<thead>
<tr>
<th></th>
<th>Tutoring Behaviors</th>
<th>Testing Behaviors</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Feedback Providing</td>
<td>Button Clicking</td>
<td></td>
</tr>
<tr>
<td>Student 1</td>
<td>93.9 (57.1-100)</td>
<td>98.1 (83.3-100)</td>
<td>95.9 (57.1-100)</td>
</tr>
<tr>
<td>Student 2</td>
<td>92.8 (66.7-100)</td>
<td>99.3 (91.7-100)</td>
<td>97.7 (75-100)</td>
</tr>
<tr>
<td>Student 3</td>
<td>93.5 (57.1-100)</td>
<td>99.7 (96.4-100)</td>
<td>95.5 (66.7-100)</td>
</tr>
<tr>
<td>Student 4</td>
<td>93.7 (66.7-100)</td>
<td>98.5 (91.7-100)</td>
<td>92.4 (66.7-100)</td>
</tr>
<tr>
<td>Total</td>
<td>93.4 (57.1-100)</td>
<td>98.9 (83.3-100)</td>
<td>95.4 (66.7-100)</td>
</tr>
</tbody>
</table>

Table 4.2. Percent tutors’ procedure integrity across participants. Numbers in parentheses indicate the range.

**Participants’ Accuracy and Speed**

*Comparing Group Level Measurement*

The comparing group level measurements of each time telling skill were calculated based on the average accuracy and speed scores of eight first grade typically developing students attending the same school with the participants.

82
The comparing group level of half hour time skill was 10.8 correct responses in 35 seconds. The accuracy ranged between five to 12 correct responses, and speed ranged between 21 seconds to 44 seconds.

The comparing group levels of quarter hour times and five-minute increments were based on only one student, since the other seven students expressed that they “do not know all these” when tested on the above skills. The tests were terminated without scores. The one general education student who took the comparing group tests scored 12 correct responses in 35 seconds on one-quarter times, 10 correct responses in 49 seconds on three-quarter times, and nine correct responses in 80 seconds on five-minute increments.

All four participants completed the study.

**Student 1**

In the three pretests on hour times, Student 1 responded correctly to all items. The skill of telling hour times was then excluded from Student 1’s subsequent tutoring sessions.

**Baseline**

Figure 4.1 shows Student 1’s accuracy scores for each session across all four time-telling skills (i.e., half hours, one-quarter hours, three-quarter hours, and five-minute increments). Before intervention Student 1 did not respond correctly to any time item in half hour or either quarter hour times. Her accuracy scores remained zero out of 12 time questions across the above three skills during baseline. In telling time in five-minute increments Student 1’s baseline accuracy scores averaged 3.8 correct responses out of 12
time questions. During the five baseline sessions on five-minute increments, Student 1 correctly named four items in four sessions and three items in one session. The four items correctly named were hour, half hour, one quarter and three quarter hour times.

Figure 4.2 shows Student 1’s speed scores for each session across all four skills. During baseline, on average, Student 1 spent 42.3 seconds to complete the 12 half-hour time questions, 45.4 seconds on one-quarter-hour times, 46.5 seconds on three-quarter-hour times, and 56 seconds on five-minute increment time telling.

Trial

In the first session when tutoring started, that is, the Trial session, Student 1 responded correctly to three half hour times, four one-quarter times, one three-quarter times, and six five-minute increments (See Table 4.3). Based on the scores in the Trial session, all originally selected accuracy criteria were retained except that of the five-minute increments, which was adjusted to eight correct responses.

Student 1’s speed scores in the Trial, however, showed larger discrepancies compared to baseline averages. Three of the speed criteria to be applied in Treatment Level 1 were adjusted accordingly. In the Trials, Student 1 spent 57 seconds on half hours, 42 seconds on one-quarter hours, 81 seconds on three-quarter hours, and 87 seconds on five-minute increments. Except for the criterion for one-quarter hour times, the other three were adjusted to 55 seconds, 70 seconds, and 85 seconds instead of the original 43 seconds, 44 seconds, and 56 seconds, respectively (See Table 4.3).
Figure 4.1: Accuracy scores across half hour, quarter hour times, and times-to-five minutes by Student 1.
Figure 4.2: Speed scores across half hour, quarter hour times, and times-to-five minutes by Student 1.
<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.8</td>
</tr>
<tr>
<td>Original Criterion 1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Trial</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Adjusted Criterion 1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td><strong>Speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Average</td>
<td>42.3</td>
<td>45.4</td>
<td>46.5</td>
<td>56</td>
</tr>
<tr>
<td>Original Criterion 1</td>
<td>43</td>
<td>44</td>
<td>44</td>
<td>56</td>
</tr>
<tr>
<td>Trial</td>
<td>57</td>
<td>42</td>
<td>81</td>
<td>87</td>
</tr>
<tr>
<td>Adjusted Criterion 1</td>
<td><strong>55</strong></td>
<td><strong>44</strong></td>
<td><strong>70</strong></td>
<td><strong>85</strong></td>
</tr>
</tbody>
</table>

Table 4.3. Student 1’s baseline averages of accuracy and speed scores across skills, and the original Criterion level-1s calculated based on the baseline averages, comparing to the Trial results and adjusted Criterion level-1s made accordingly. (A bold number indicates that there was adjustment made in that skill measurement).

**Peer-tutoring**

Peer-tutoring sessions started with the adjusted criteria after the Trial session. Multiple criteria on both accuracy and speed were assigned to each skill.

*Half hour times.* When tutoring was first initiated in the Trial session, Student 1 correctly responded to three half hour time items, when the criterion was set at four correct responses. Since the result of the Trial was comparable to the originally assigned criterion, this criterion was thus retained as the first criterion level of CATS. Student 1’s average accuracy scores on half hour times steadily increased along with the criteria.
assigned to each level or goal. Her average accuracy started at 5.5 correct responses during Level 1 with the criterion set at four, increased to 8.5 during Level 2 (criterion set at seven), 11.7 during Level 3 (criterion set at 12), 12 during Level 4 (criterion set at 12), and 11.8 during Level 5 (criterion set at 12). Student 1 incorrectly responded to only one item in the four sessions of Level 5.

Student 1 used 57 seconds in the Trial session on half hour times, 33% more than the average speed level (42.3 seconds) during baseline. The first criterion level was then adjusted accordingly to 55 seconds. In the following tutoring on five criterion levels, the average speed scores of Student 1 on half hour times show a downward trend in length of time used to complete each session. On average, Student 1 spent 52.5 seconds to complete all twelve half hour times during Level 1 (criterion set at 55 seconds), then the length of time decreased to 44 seconds on Level 2 (criterion set at 49 seconds), 38.7 seconds on Level 3 (criterion set at 42 seconds), 30 seconds on Level 4 (criterion set at 35 seconds), and 24 seconds on Level 5 (criterion set at 28 seconds). See Table 4.4 and 4.5 for complete scores of Student 1 during baseline and the peer tutoring intervention.
### Table 4.4

Ranges, means, and criteria of accuracies (in number out of 12 time items) across four skills during baseline and intervention by Student 1.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0-0</td>
<td>0-0</td>
<td>0-0</td>
<td>49-64</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.8</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>5-6</td>
<td>8-8</td>
<td>4-10</td>
<td>9-12</td>
</tr>
<tr>
<td>Mean</td>
<td>5.5</td>
<td>8</td>
<td>7.7</td>
<td>10.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>6-11</td>
<td>10-12</td>
<td>7-12</td>
<td>9-11</td>
</tr>
<tr>
<td>Mean</td>
<td>8.5</td>
<td>11.3</td>
<td>10.4</td>
<td>10.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>11-12</td>
<td>11-12</td>
<td>11-12</td>
</tr>
<tr>
<td>Mean</td>
<td>11.7</td>
<td>11.8</td>
<td>11.7</td>
<td>11.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12-12</td>
<td>12-12</td>
<td>10-12</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>12</td>
<td>11.5</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Criterion 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>12-12</td>
<td>11-12</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11.8</td>
<td>12</td>
<td>11.8</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4.5. Ranges, means, and criteria of speed (in seconds) across four skills during baseline and intervention by Student 1.

<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>36-47</td>
<td>39-59</td>
<td>38-59</td>
<td>49-64</td>
</tr>
<tr>
<td>Mean</td>
<td>42.3</td>
<td>45.4</td>
<td>46.5</td>
<td>56</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>51-54</td>
<td>36-42</td>
<td>55-67</td>
<td>67-77</td>
</tr>
<tr>
<td>Mean</td>
<td>52.5</td>
<td>38.7</td>
<td>61.7</td>
<td>73</td>
</tr>
<tr>
<td>Criteria</td>
<td>55</td>
<td>44</td>
<td>70</td>
<td>85</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>39-49</td>
<td>34-38</td>
<td>46-57</td>
<td>59-72</td>
</tr>
<tr>
<td>Mean</td>
<td>44</td>
<td>35.3</td>
<td>50.6</td>
<td>63.5</td>
</tr>
<tr>
<td>Criteria</td>
<td>49</td>
<td>41</td>
<td>57</td>
<td>75</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>36-41</td>
<td>29-40</td>
<td>34-41</td>
<td>47-54</td>
</tr>
<tr>
<td>Mean</td>
<td>38.7</td>
<td>32</td>
<td>37.3</td>
<td>50</td>
</tr>
<tr>
<td>Criteria</td>
<td>42</td>
<td>32</td>
<td>46</td>
<td>55</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>27-34</td>
<td>26-29</td>
<td>25-29</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>30</td>
<td>27.7</td>
<td>26.8</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>35</td>
<td>30</td>
<td>35</td>
<td>-</td>
</tr>
<tr>
<td>Criterion 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>23-25</td>
<td>21-23</td>
<td>21-24</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>24</td>
<td>22.3</td>
<td>23</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>28</td>
<td>25</td>
<td>25</td>
<td>-</td>
</tr>
</tbody>
</table>

*One-quarter hour times.* When tutoring started in the Trial session of one-quarter hour times, Student 1 correctly responded to four time items, which exactly matched the
criterion assigned. This criterion was then retained as the first criterion level in the following intervention. However, despite of the criterion, Student 1’s accuracy scores on one-quarter hour times increased rapidly during criterion Level 1 with an average of eight correct responses, compared to zero during baseline. The second criterion level was then adjusted accordingly to nine correct responses. Instead of a gradual increment along with the criteria, Student 1’s accuracy scores maintained 12 correct responses in the following levels except two occasions. Student 1 missed two items once in Level 2 and missed one item in Level 3. Student 1’s average accuracy scores during Level 2 through Level 5 demonstrates a stable pace with 11.3 correct responses on Level 2 (criterion set at nine), 11.8 on Level 3 (criterion set at 11), and 12 on both Level 4 and 5 (criterion set at 12).

Student 1 spent 42 seconds in the Trial session on one-quarter hour times, which is comparable to the average speed level (44 seconds) in baseline. This criterion was then retained as the first criterion level. In the following tutoring on five criteria, the average speed scores of Student 1 on one-quarter hour times gradually decreased along with the criteria assigned to each level. On average, Student 1 used 38.7 seconds to complete all twelve one-quarter hour times during Level 1 (criterion set at 44 seconds), then decreased to 35.3 seconds on Level 2 (criterion set at 41 seconds), 32 seconds on Level 3 (criterion set at 32 seconds), 27.7 seconds on Level 4 (criterion set at 30 seconds), and 22.3 seconds on Level 5 (criterion set at 25 seconds).

Three-quarter hour times. When tutoring started in the Trial session of three-quarter hour times, Student 1 correctly responded to only one time items, while the criterion was assigned to four. However this criterion was retained as the first criterion
level in the following intervention. In the first tutoring session immediately after the Trial, Student 1 reached this criterion with four correct responses. The following two sessions under criterion Level 1, however, were similar to one-quarter hour times, when Student 1’s accuracy scores quickly increased to nine and 10 respectively. The second criterion was then adjusted to nine. Student 1’s accuracy scores fluctuated during Level 2 (i.e., scores ranged seven to 12). In the following three levels, Student 1 kept a score of 12 with a few variations of 10 or 11. Student 1’s average accuracy scores during Level 2 through Level 5 were 10.4 on Level 2 (criterion set at nine), 11.7 on Level 3 (criterion set at 11), 11.5 and 11.8 on Level 4 and 5 respectively (criterion set at 12).

Student 1 used 81 seconds in the Trial session on three-quarter hour times, almost doubled the average speed (44 seconds) during baseline. The first criterion level was then adjusted accordingly to 70 seconds. In the following tutoring on five criteria, the average speed scores of Student 1 on three-quarter hour times gradually decreased along with the criteria assigned to each level. On average, Student 1 used 61.7 seconds during Level 1 (criterion set at 70 seconds), 50.6 seconds on Level 2 (criterion set at 57 seconds), 37.3 seconds on Level 3 (criterion set at 46 seconds), 26.8 seconds on Level 4 (criterion set at 35 seconds), and 23 seconds on Level 5 (criterion 25 seconds).

*Five-minute increments.* When tutoring started in the Trial session of five-minute increments, Student 1 correctly responded to six time items, which was comparable to the criterion assigned (i.e., 7 correct responses). The first criterion level was set at eight, however Student 1’s accuracy scores made the most increment during this level from an average of 3.8 correct responses during baseline to 10.7. In the second session within
level one, Student 1 responded correctly on all 12 items. Student 1’s accuracy scores show no progression during Level 2 with an average of 10.3 (criterion set at 10). Student 1 had an average 11.7 correct responses during the last level (criterion set at 12).

Student 1’s speed score (87 seconds) on five-minute increments Trial was 50% more than the average speed (56 seconds) during baseline. The first criterion level was then adjusted accordingly to 85 seconds. Like other speed score trend, Student 1 had decreased average scores on this skill. Over time during the first level, Student 1 used an average of 73 seconds to complete all twelve items (criterion set at 85 seconds), then 63.5 seconds on Level 2 (criterion set at 75 seconds), and finally 50 seconds on Level 3 (criterion set at 55 seconds).

**Maintenance**

Maintenance data on each skill were collected every three days after CATS were completed, with the exception of weekends or participants’ absences. When these events occurred, maintenance tests were postponed to the first available weekday. Student 1 maintained 100% accuracy (i.e., 12 correct responses out of 12 time items) on all four skills except in one session on one-quarter hours, when she missed one item. Student 1 achieved an average 24 seconds per test on half hour times, comparing to the final goal on this skill of 28 seconds. For the other three skills, Student 1’s average maintenance speeds were 25.5 seconds on one-quarter hours (final goal of 25 seconds), 24 seconds on three-quarter hours (final goal of 25 seconds), and 51 seconds on five-minute increments (final goal of 55 seconds) respectively. See Table 4.2c for Student 1’s complete maintenance scores.
Table 4. 6. Ranges and means of speed (in seconds) by Student 1 across four skills during maintenance and generality assessments comparing to the final goals assigned to each skill.

### Generality

Generality data on all four skills were collected on the same days when maintenance assessments were given. When presented with pictures of clock faces different from those used during CATS, Student 1 correctly responded to all but two time items. Both mistakes were made on five-minute increments. The lengths of time used to complete generality assessments were comparable to the final goals of each skill. On average, Student 1 used 28.3 seconds on half-hour times (final goal of 28 seconds), 23.5 seconds on one-quarter hours (final goal of 25 seconds), 26 seconds on three-quarter hours (final goal of 25 seconds), and 52.3 seconds on five-minute increments (final goal of 55 seconds) respectively. See Table 4.6 for Student 1’s complete generalization assessment scores.
In the three pretests on hour and half hour times, Student 2 responded correctly to all items. The skills of telling hour and half hour times were then excluded from Student 2’s subsequent tutoring intervention.

**Baseline**

Figure 4.3 shows Student 2’s accuracy scores for each session across all three time-telling skills (i.e., one quarter hours, three quarter hours, and five-minute increments). Before intervention Student 2 did not respond correctly to any time item in either quarter hour times. His accuracy scores remained zero out of 12 time items across the above two skills during baseline. In telling time in five-minute increments Student 2’s baseline accuracy scores averaged 4.3 correct responses. During the nine baseline sessions on five-minute increments, Student 2 correctly named three time items in the first three sessions. In the fourth session, he correctly named six items. In the following five baseline sessions, Student 2’s accuracy scores were five correct responses in four sessions and four correct responses in one session. The time items correctly named varied.

Figure 4.4 shows Student 2’s speed scores for each session across all three skills. During baseline, on average, Student 2 spent 65.2 seconds to complete the 12 one-quarter-hour times, 81.3 seconds on three-quarter-hour times, and 96 seconds on five-minute increments.
Figure 4.3: Accuracy scores across one quarter hour, three quarter hour times, and five-minute increments by Student 2.
Figure 4.4: Speed scores across one quarter hour, three quarter hour times, and times-to-five minutes by Student 2.
In the first session when tutoring started, that is, the Trial session, Student 2 responded correctly to all but one of one-quarter times, all 12 three-quarter times, and nine five-minute increments (See Table 4.7). Based on the scores in the Trial session, all originally selected accuracy criterion for Treatment Level-1 were adjusted to 10, 10, and six respectively.

For the two quarter hour skills Student 2’s speed scores in the Trial (51 seconds and 79 seconds) were comparable to the original criteria (55 seconds and 80 seconds), thus the original criteria for these two skills were retained. For five-minute increments, Student 2 spent 78 seconds more in the Trial (173 seconds) than the originally assigned criterion (95 seconds). The speed criterion Level-1 for five-minute increments was adjusted accordingly to 155 seconds. (See Table 4.7).
One-quarter Hours  Three-quarter Hours  Five-minute increments

<table>
<thead>
<tr>
<th></th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Average</td>
<td>0</td>
<td>0</td>
<td>4.3</td>
</tr>
<tr>
<td>Original Criterion 1</td>
<td>4</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Trial</td>
<td>11</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Adjusted Criterion 1</td>
<td><strong>10</strong></td>
<td>10</td>
<td><strong>6</strong></td>
</tr>
<tr>
<td><strong>Speed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Average</td>
<td>65.2</td>
<td>81.3</td>
<td>96</td>
</tr>
<tr>
<td>Original Criterion 1</td>
<td>55</td>
<td>80</td>
<td>95</td>
</tr>
<tr>
<td>Trial</td>
<td>51</td>
<td>79</td>
<td>173</td>
</tr>
<tr>
<td>Adjusted Criterion 1</td>
<td>55</td>
<td>80</td>
<td><strong>155</strong></td>
</tr>
</tbody>
</table>

Table 4. 7. Student 2’s baseline averages of accuracy and speed scores across skills, and the original Criterion level-1s calculated based on the baseline averages, comparing to the Trial results and adjusted Criterion level-1s made accordingly. (A bold number indicates that there was adjustment made in that skill measurement).

**Peer-tutoring**

Peer-tutoring sessions started with the adjusted criteria after the Trial session.

Multiple criteria on both accuracy and speed were assigned to each skill.

*One quarter hour times.* Except the second session in Level 1 when Student 2 incorrectly responded to one item, he scored 12 correct responses out of 12 items throughout the tutoring sessions on one-quarter-hour times.

During tutoring on five criterion levels, the average speed scores of Student 2 on one-quarter-hour times show a downward trend in length of time used to complete each
session. On average, Student 2 spent 54.5 seconds to complete all twelve one-quarter-hour times during Level 1 (criterion set at 55 seconds), then the length of time decreased to 40 seconds on Level 2 (criterion set at 49 seconds), 37.7 seconds on Level 3 (criterion set at 42 seconds), 31.7 seconds on Level 4 (criterion set at 35 seconds), and 29.6 seconds on Level 5 (criterion set at 30 seconds). See Table 4.8 and 4.9 for complete scores of Student 2 during baseline and the peer tutoring intervention.
<table>
<thead>
<tr>
<th></th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0-0</td>
<td>0-0</td>
<td>3-6</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
<td>0</td>
<td>4.3</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>10-12</td>
<td>7-10</td>
</tr>
<tr>
<td>Mean</td>
<td>11.8</td>
<td>10.8</td>
<td>9</td>
</tr>
<tr>
<td>Criteria</td>
<td>10</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12-12</td>
<td>11-12</td>
<td>11-12</td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>11.7</td>
<td>11.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12-12</td>
<td>11-12</td>
<td>11-12</td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>11.7</td>
<td>11.8</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12-12</td>
<td>12-12</td>
<td>12-12</td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>N/A</td>
</tr>
<tr>
<td>Criterion 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12-12</td>
<td>12-12</td>
<td>11-12</td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>12</td>
<td>11.8</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 4.8. Ranges, means, and criteria of accuracies (in number out of 12 time items) across three skills during baseline and intervention by Student 2.
<table>
<thead>
<tr>
<th>Criterion</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>50-94</td>
<td>58-98</td>
<td>79-123</td>
</tr>
<tr>
<td>Mean</td>
<td>65.2</td>
<td>81.3</td>
<td>96</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>53-61</td>
<td>69-97</td>
<td>136-148</td>
</tr>
<tr>
<td>Mean</td>
<td>54.5</td>
<td>78</td>
<td>140.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>55</td>
<td>80</td>
<td>155</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>38-42</td>
<td>57-65</td>
<td>118-120</td>
</tr>
<tr>
<td>Mean</td>
<td>40</td>
<td>60</td>
<td>119.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>49</td>
<td>65</td>
<td>130</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>36-40</td>
<td>34-55</td>
<td>78-155</td>
</tr>
<tr>
<td>Mean</td>
<td>37.7</td>
<td>44.8</td>
<td>117.8</td>
</tr>
<tr>
<td>Criteria</td>
<td>42</td>
<td>50</td>
<td>110</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>29-34</td>
<td>31-44</td>
<td>73-92</td>
</tr>
<tr>
<td>Mean</td>
<td>31.7</td>
<td>36</td>
<td>82</td>
</tr>
<tr>
<td>Criteria</td>
<td>35</td>
<td>35</td>
<td>N/A</td>
</tr>
<tr>
<td>Criterion 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>26-35</td>
<td>27-39</td>
<td>75-82</td>
</tr>
<tr>
<td>Mean</td>
<td>29.6</td>
<td>30.3</td>
<td>77.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>30</td>
<td>30</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 4.9. Ranges, means, and criteria of speed (in seconds) across three skills during baseline and intervention by Student 2.

Three-quarter hour times. On three-quarter hour times, Student 2’s accuracy scores ranged 10 to 12 correct responses during Level 1 with an average score of 10.8.
During Level 2 and 3, his accuracy scores ranged 11 to 12 with an average score of 11.7 on both levels. On the final two levels, Student 2’s accuracy scores maintained 12 correct responses out of all 12 time items in each session.

During tutoring on five criterion levels, the average speed scores of Student 2 on three-quarter-hour times show a downward trend in length of time used to complete each session. On average, Student 2 spent 78 seconds to complete all twelve items during Level 1 (criterion set at 80 seconds), then the length of time decreased to 60 seconds on Level 2 (criterion set at 65 seconds), 44.8 seconds on Level 3 (criterion set at 50 seconds), 36 seconds on Level 4 (criterion set at 35 seconds), and 30.3 seconds on Level 5 (criterion set at 30 seconds). See Table 4.8 and 4.9 for complete scores of Student 2 during baseline and the peer tutoring intervention.

Five-minute increments. On five-minute increments, Student 2’s accuracy scores ranged seven to 10 correct responses during Level 1 with an average score of nine. During Level 2 and 3, his accuracy scores ranged 11 to 12 with average scores of 11.7 and 11.8 respectively on the two levels. On Level 4, Student 2 correctly responded to all items in each session. On Level 5, except for one session when Student 2 incorrectly responded to one item, he maintained 12 correct responses out of 12 items.

On the first two criterion levels during tutoring, the average speed scores of Student 2 on five-minute increments show a downward trend in length of time used to complete each session. On average, Student 2 spent 140.3 seconds to complete all twelve items during Level 1 (criterion set at 155 seconds), then the length of time decreased to 119.3 seconds on Level 2 (criterion set at 130 seconds). This gradual decreasing trend
was interrupted on Level 3, when a rapid increment in length of time used to complete a session was observed. During this level, Student 2’s speed scores increased from 78 seconds to 155 seconds, with an average of 117.8 seconds, despite of the criterion assigned (110 seconds). The subsequent Level 4 was a special level in which Student 2 was timed but could not see the timing bar. During Level 4 Student 2’s speed scores decreased to an average of 82 seconds with a range of 73-92 seconds. Accordingly the following criterion of Level 5 was assigned to 80 seconds. Student 2 scored 77.3 seconds on average during this level with a range of 75 to 82 seconds.

**Maintenance**

On five-minute increments, Student 2 maintained 100% accuracy (i.e., 12 corrects out of 12 time items) on all three skills. Student 2 achieved an average 23.8 second speed on one-quarter hour times, comparing to the final goal of this skill of 30 seconds, 26.2 seconds on three-quarter hours (final goal set at 30 seconds), and 58 seconds on five-minute increments (final goal set at 80 seconds) respectively. See Table 4.10 for Student 2’s complete maintenance scores.

**Generality**

Except for one mistake made in the third generalization test on three-quarter times, Student 2 responded correctly to all the time items in generalization tests across the three skills he previously practiced on during intervention. The lengths of time used to complete generality assessments were comparable to the final goals of each skill. On average, Student 2 used 27.8 seconds on one-quarter hour times (final goal set at 30 seconds), 28 seconds on three-quarter hours (final goal set at 30 seconds), and 65 seconds
on five-minute increments (final goal set at 80 seconds) respectively. See Table 4.10 for
Student 2’s complete generalization assessment scores.

<table>
<thead>
<tr>
<th></th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maintenance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>21-26</td>
<td>21-33</td>
<td>56-62</td>
</tr>
<tr>
<td>Mean</td>
<td>23.8</td>
<td>26.2</td>
<td>58</td>
</tr>
<tr>
<td>Final goals</td>
<td>30</td>
<td>30</td>
<td>80</td>
</tr>
<tr>
<td><strong>Generality</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>26-31</td>
<td>24-32</td>
<td>62-67</td>
</tr>
<tr>
<td>Mean</td>
<td>27.8</td>
<td>28</td>
<td>65</td>
</tr>
<tr>
<td>Final goals</td>
<td>30</td>
<td>30</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 4.10. Ranges and means of speeds (in seconds) by Student 2 across three skills
during maintenance and generality assessments comparing to the final goals assigned to
each skill.

**Student 3**

In the three pretests on hour times, Student 3 responded correctly to more than
90% of all items. The skills of telling hour times were then excluded from Student 3’s
forthcoming tutoring sessions.

**Baseline**

Figure 4.5 shows Student 3’s accuracy scores for each session across all three
time-telling skills (i.e., half hours, one quarter hours, and three quarter hours). Before
intervention Student 3 did not respond correctly to any time item of the above
three time telling skills. His accuracy scores remained zero out of 12 time questions across the above three skills during baseline.

Figure 4.6 shows Student 3’s speed scores for each session across all the three skills. During baseline, on average, Student 3 spent 39.2 seconds to complete all 12 half hour times, 40.2 seconds on one-quarter-hour times, and 37.4 seconds on three-quarter-hour times.

Trial

In the first session when tutoring started, that is, the Trial session, Student 3 responded correctly to five of 12 half hour times, 11 one-quarter hour times, and seven three-quarter hour times. (See Table 4.11). Based on the scores in the Trial session, Student 3’s accuracy criterion for Treatment Level-1 were retained or adjusted to four, 10, and nine respectively.

Student 3 spent 67 seconds, 42 seconds, and 50 seconds to complete half, one-quarter, and three-quarter hour times during the Trial. The speed criterion Level-1 for the three skills were retained or adjusted accordingly to 65, 40, and 50 seconds respectively (See Table 4.11).
Figure 4.5: Accuracy scores across half hour, one quarter hour, three quarter hour times, and five-minute increments by Student 3.
Figure 4.6: Speed scores across one half hour, one quarter hour, three quarter hour times, and times-to-five minutes by Student 3.
Table 4. 11. Student 3’s baseline averages of accuracy and speed scores across skills, and the original Criterion level-1s calculated based on the baseline averages, comparing to the Trial results and adjusted Criterion level-1s made accordingly. (A bold number indicates that there was adjustment made in that skill measurement).

Peer-tutoring

Peer-tutoring sessions started with the adjusted criteria after the Trial session.

Multiple criteria on both accuracy and speed were assigned to each skill.

*Half hour times.* Student 3’s average accuracy scores on half hour times increased from an average of 8.5 during Level 1 to 11.8 during Level 4, with a few unsteady fluctuations in the progress. His average accuracy and corresponding criteria are: 8.5 during Level 1 (criterion set at four), 9.3 during Level 2 (criterion set at nine), 10.1 during Level 3 (criterion set at 10), and 11.8 during the last level (criterion set at 12).

During tutoring on four criterion levels, the average speed scores of Student 3 on...
half hour times show a downward trend in length of time used to complete each session. On average, Student 3 spent 61.8 seconds to complete all twelve half hour times during Level 1 (criterion set at 65 seconds), then the length of time decreased to 54 seconds on Level 2 (criterion set at 55 seconds), 42.9 seconds on Level 3 (criterion set at 45 seconds), and 32.7 seconds on Level 4 (criterion set at 35 seconds). See Table 4.12 and 4.13 for complete scores of Student 3 during baseline and the peer tutoring intervention.

<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0-0</td>
<td>0-0</td>
<td>0-0</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>7-10</td>
<td>9-12</td>
<td>9-11</td>
</tr>
<tr>
<td>Mean</td>
<td>8.5</td>
<td>11</td>
<td>10.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>4</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>9-10</td>
<td>10-12</td>
<td>12-12</td>
</tr>
<tr>
<td>Mean</td>
<td>9.3</td>
<td>11.4</td>
<td>12</td>
</tr>
<tr>
<td>Criteria</td>
<td>9</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>9-12</td>
<td>12-12</td>
<td>12-12</td>
</tr>
<tr>
<td>Mean</td>
<td>10.1</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Criteria</td>
<td>10</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>11.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.12. Ranges, means, and criteria of accuracies (in number out of 12 time items) across three skills during baseline and intervention by Student 3.
<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>30-58</td>
<td>32-62</td>
<td>31-44</td>
</tr>
<tr>
<td>Mean</td>
<td>39.2</td>
<td>40.2</td>
<td>37.4</td>
</tr>
<tr>
<td><strong>Criterion 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>60-64</td>
<td>31-40</td>
<td>41-55</td>
</tr>
<tr>
<td>Mean</td>
<td>61.8</td>
<td>37.3</td>
<td>44.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>65</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td><strong>Criterion 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>53-55</td>
<td>30-35</td>
<td>34-37</td>
</tr>
<tr>
<td>Mean</td>
<td>54</td>
<td>33</td>
<td>35.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>55</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td><strong>Criterion 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>40-47</td>
<td>26-33</td>
<td>26-28</td>
</tr>
<tr>
<td>Mean</td>
<td>42.9</td>
<td>28.5</td>
<td>27.3</td>
</tr>
<tr>
<td>Criteria</td>
<td>45</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td><strong>Criterion 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>28-40</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>32.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>35</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.13. Ranges, means, and criteria of speed (in seconds) across three skills during baseline and intervention by Student 3.

*One-quarter hour times.* On one-quarter hour times, Student 3’s accuracy scores ranged nine to 12 correct responses during Level 1 with an average score of 11. Average of Level 2 was 11.4 with a range of 10 to 12. Student 3 responded correctly to all items during Level 3.
During tutoring on three criterion levels, the average speed scores of Student 3 on one-quarter-hour times show a downward trend as well. On average, Student 3 spent 37.3 seconds to complete all twelve items during Level 1 (criterion set at 40 seconds), then the length of time decreased to 33 seconds on Level 2 (criterion set at 35 seconds), and 28.5 seconds on Level 3 (criterion set at 30 seconds). See Table 4.12 and 4.13 for complete scores of Student 3 during baseline and the peer tutoring intervention.

*Three-quarter hour times.* On three-quarter hour times, Student 3’s accuracy scores ranged nine to 11 correct responses during Level 1 with an average score of 10.3. During Level 2 and 3, his accuracy scores maintained 12 correct responses out of 12 items.

During tutoring on three criterion levels, the average speed scores of Student 3 on three-quarter-hour times show a downward trend with one occurrence of variation. On average, Student 2 spent 44.6 seconds to complete all twelve items during Level 1 (criterion set at 50 seconds), then the length of time decreased to 35.3 seconds on Level 2 (criterion set at 40 seconds), and 27.3 seconds on Level 3 (criterion set at 30 seconds). See Table 4.12 and 4.13 for complete scores of Student 3 during baseline and the peer tutoring intervention. The variation occurred during the third session in Level 1 when his speed score (i.e., 55 seconds) increased 34% from previous session (i.e., 41 seconds). The speed score retrieved to 44 seconds in the following session.

*Maintenance*

Student 3 maintained 100% correctness (i.e., 12 corrects out of 12 time items) on all three skills. Student 3 achieved an average 27.8 seconds on half hour times,
comparing to his final goal on this skill of 35 seconds, 26.5 seconds on one-quarter hours (final goal of 30 seconds), and 27.8 seconds on three-quarter hours (final goal of 30 seconds) respectively. See Table 4.14 for Student 3’s complete maintenance scores.

<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>24-38</td>
<td>24-29</td>
<td>26-29</td>
</tr>
<tr>
<td>Mean</td>
<td>27.8</td>
<td>26.5</td>
<td>27.8</td>
</tr>
<tr>
<td>Final goals</td>
<td>35</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Generality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>29-32</td>
<td>27-31</td>
<td>28-33</td>
</tr>
<tr>
<td>Mean</td>
<td>30.5</td>
<td>29.8</td>
<td>30.5</td>
</tr>
<tr>
<td>Final goals</td>
<td>35</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 4.14. Ranges and means of speed (in seconds) by Student 3 across three skills during maintenance and generality assessments comparing to the final goals assigned to each skill.

**Generality**

Student 3 responded correctly to all the time items in generalization tests across the three skills he previously practiced on during intervention. The lengths of time used to complete generality assessments were comparable to his final goals of each skill. On average, Student 3 used 30.5 seconds on half hour times (final goal of 35 seconds), 29.8 seconds on one-quarter hours (final goal of 30 seconds), and 30.5 seconds on three-
quarter hour times (final goal of 30 seconds) respectively. See Table 4.14 for Student 3’s complete generalization assessment scores.

**Student 4**

In the three pretests on hour times, Student 4 responded correctly to all items. The skill of telling hour times was then excluded from Student 4’s subsequent tutoring sessions.

**Baseline**

Figure 4.7 shows Student 4’s accuracy scores for each session across all four time-telling skills (i.e., half hours, one quarter hours, three quarter hours, and five-minute increments). Before intervention Student 4 correctly responded to an average of six items of the 12 half hour times (ranged five to eight). He did not respond correctly to any time item in either quarter hour times. His accuracy scores remained zero out of 12 time questions across the above two skills during baseline. In telling times in five-minute increments Student 4’s baseline accuracy scores averaged 5.7 (ranged three to eight).

Figure 4.8 shows Student 4’s speed scores for each session across all four skills. During baseline, on average, Student 1 spent 47.3 seconds to complete the 12 half-hour time questions, 62.2 seconds on one-quarter-hour times, 50 seconds on three-quarter-hour times, and 75 seconds on five-minute increments.
Figure 4.7: Accuracy scores across half hour, quarter hour times, and times-to-five minutes by Student 4.
Figure 4.8: Speed scores across half, quarter, and times-to-five minutes by Student 4.
Trial

In the first session when tutoring started, that is, the Trial session, Student 4 responded correctly to seven half hour times, 11 one-quarter hour times, eight three-quarter hour times, and 10 five-minute increments (See Table 4.15). Based on the scores in the Trial session, all originally selected accuracy criteria were changed or retained to eight, 11, eight, and 10 for the four skills respectively.

Student 4’s speed scores in the Trial, however, showed larger discrepancies compared to baseline averages. All four speed criteria for Treatment Level-1s were adjusted accordingly. In the Trials, on average Student 4 spent 74 seconds on half hour times, 45 seconds on one-quarter hours, 45 seconds on three-quarter hours, and 110 seconds on five-minute increments. The Level 1 criteria were adjusted to 70 seconds, 45 seconds, 45 seconds, and 110 seconds instead of the original 47 seconds, 62 seconds, 50 seconds, and 75 seconds, respectively (See Table 4.15).
Table 4. 15. Student 4’s baseline averages of accuracy and speed scores across skills, and the original Criterion level-1s calculated based on the baseline averages, comparing to the Trial results and adjusted Criterion level-1s made accordingly. (A bold number indicates that there was adjustment made in that skill measurement).

Peer-tutoring

Peer-tutoring sessions started with the adjusted criteria after the Trial session. Multiple criteria on both accuracy and speed were assigned to each skill.

Half hour times. Student 4’s average accuracy scores on half hour times increased with slight fluctuations along the criteria assigned to each level. His average accuracy started at 7.8 during Level 1 with the criterion set to eight, then increased to 9.2 during Level 2 (criterion set at nine), 11 during Level 3 (criterion set at 10), 11.6 during Level 4 (criterion set at 12), and 11.8 (with only one session of 11 corrects) during the last level (criterion set at 12).
During tutoring on five criterion levels, the average speed scores of Student 4 on half hour times show a downward trend in length of time used to complete each session. On average, Student 4 spent 63 seconds during Level 1 (criterion set at 70 seconds), then 51.8 seconds on Level 2 (criterion set at 58 seconds), 43.3 seconds on Level 3 (criterion set at 46 seconds), 30.3 seconds on Level 4 (criterion set at 35 seconds), and 23.5 seconds on Level 5 (criterion set at 25 seconds). See Table 4.16 and 4.17 for complete scores of Student 4 during baseline and the peer tutoring intervention.

One-quarter hour times. Student 4’s accuracy scores had a rapid increment during the first criterion level with an average of 11, comparing to zero during baseline. Student 4 missed two time items during the following four criterion levels, with average scores of 11.7, 11.8, 12, and 12 respectively.

The average speed scores of Student 4 on one-quarter hour times gradually decreased along with the criteria assigned to each level. On average, Student 4 used 41.8 seconds on one-quarter hour times during Level 1 (criterion set at 45 seconds), the time was then decreased to 38 seconds on Level 2 (criterion set at 42 seconds), 31.5 seconds on Level 3 (criterion set at 35 seconds), 27.8 seconds on Level 4 (criterion set at 30 seconds), and 22.7 seconds on Level 5 (criterion set at 25 seconds).
<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>5-8</td>
<td>0-0</td>
<td>0-0</td>
<td>3-8</td>
</tr>
<tr>
<td>Mean</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>5.8</td>
</tr>
<tr>
<td><strong>Criterion 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>7-8</td>
<td>10-12</td>
<td>7-12</td>
<td>10-12</td>
</tr>
<tr>
<td>Mean</td>
<td>7.8</td>
<td>11</td>
<td>10.5</td>
<td>11</td>
</tr>
<tr>
<td>Criteria</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td><strong>Criterion 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>8-11</td>
<td>11-12</td>
<td>10-12</td>
<td>9-12</td>
</tr>
<tr>
<td>Mean</td>
<td>9.2</td>
<td>11.7</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Criteria</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td><strong>Criterion 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>11-12</td>
<td>12-12</td>
<td>11-12</td>
</tr>
<tr>
<td>Mean</td>
<td>11</td>
<td>11.8</td>
<td>12</td>
<td>11.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>10</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td><strong>Criterion 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>10-12</td>
<td>12-12</td>
<td>12-12</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>11.6</td>
<td>12</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td><strong>Criterion 5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-12</td>
<td>12-12</td>
<td>12-12</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>11.8</td>
<td>12</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.16. Ranges, means, and criteria of accuracies (in number out of 12 time items) across four skills during baseline and intervention by Student 4.
<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>39-59</td>
<td>53-70</td>
<td>35-68</td>
<td>54-104</td>
</tr>
<tr>
<td>Mean</td>
<td>47.3</td>
<td>62.2</td>
<td>51.6</td>
<td>80.3</td>
</tr>
<tr>
<td>Criterion 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>62-66</td>
<td>36-44</td>
<td>39-48</td>
<td>67-88</td>
</tr>
<tr>
<td>Mean</td>
<td>63</td>
<td>41.8</td>
<td>43.5</td>
<td>74.7</td>
</tr>
<tr>
<td>Criteria</td>
<td>70</td>
<td>45</td>
<td>45</td>
<td>110</td>
</tr>
<tr>
<td>Criterion 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>47-56</td>
<td>35-40</td>
<td>36-48</td>
<td>58-69</td>
</tr>
<tr>
<td>Mean</td>
<td>51.8</td>
<td>38</td>
<td>40</td>
<td>65</td>
</tr>
<tr>
<td>Criteria</td>
<td>58</td>
<td>42</td>
<td>40</td>
<td>70</td>
</tr>
<tr>
<td>Criterion 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>41-46</td>
<td>28-34</td>
<td>31-32</td>
<td>55-62</td>
</tr>
<tr>
<td>Mean</td>
<td>43.3</td>
<td>31.5</td>
<td>31.3</td>
<td>57.6</td>
</tr>
<tr>
<td>Criteria</td>
<td>46</td>
<td>35</td>
<td>35</td>
<td>60</td>
</tr>
<tr>
<td>Criterion 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>27-37</td>
<td>22-33</td>
<td>26-29</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>30.3</td>
<td>27.8</td>
<td>27.7</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>35</td>
<td>30</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>Criterion 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>21-27</td>
<td>20-25</td>
<td>24-24</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>23.5</td>
<td>22.7</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Criteria</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4. 17. Ranges, means, and criteria of speed (in seconds) across four skills during baseline and intervention by Student 4.

*Three-quarter hour times.* On three-quarter hour times, during Level 1, Student 4’s accuracy scores quickly increased within three sessions to 12 corrects from the initial
seven correct responses. This 100% accuracy was maintained for three consecutive sessions until the last session on Level 1 when he missed two items and scored 10. Student 4’s accuracy scores increased from 10 to 12 during Level 2, and maintained 12 in the following three levels.

The average speed scores of Student 4 on three-quarter hour times gradually decreased along the criteria assigned to each level. On average, Student 4 used 43.5 seconds to complete all twelve three-quarter hour times during Level 1 (criterion 45 seconds), the time was then decreased to 40 seconds on Level 2 (criterion 40 seconds), 31.3 seconds on Level 3 (criterion 35 seconds), 27.7 seconds on Level 4 (criterion 30 seconds), and 24 seconds on Level 5 (criterion 25 seconds).

*Five-minute increments.* Three criterion levels were assigned to five-minute increments for Student 4. Fluctuations occurred throughout the three levels. Student 4’s accuracy scores increased to 12 during Level 1 with an average of 11 and a range of 10 to 12. The range was widened during Level 2 when Student 4 missed three items in the third session on this level, although the average accuracy score of Level 2 remained 11. Of the seven sessions in Level 3, Student 4 scored 12 in all except two sessions when he scored 11. The average accuracy of Level 3 was 11.7.

During the first level of tutoring on five-minute increments, Student 4’s speed scores did not change along with the criterion assigned based on the Trial result. On average, Student 4 spent 74.7 seconds on this level, 47% lower than the assigned criterion of 110 seconds for the level. During the next two levels, however, Student 4’s speed scores changed along with the assigned criteria. On Level 2, Student 4’s average speed
score was 65 seconds, comparable to the criterion of 70 seconds. On Level 3, the average score was 57.6, while the criterion was assigned to 60 seconds.

**Maintenance**

Student 1 maintained 100% correctness (i.e., 12 correct responses out of 12 time items) on all four skills except one session on half hour times, when he missed one item. Due to time limit, only one maintenance test was given on five-minute increments. Student 4 achieved an average of 25.8 seconds on half hour times, comparing to his final goal on this skill of 25 seconds. For the other three skills, Student 4’s average maintenance scores were 21.8 seconds on one-quarter hours (final goal of 25 seconds), 28 seconds on three-quarter hours (final goal of 25 seconds), and 55 seconds on five-minute increments (final goal of 60 seconds) respectively. See Table 4.18 for Student 4’s complete maintenance scores.

<table>
<thead>
<tr>
<th></th>
<th>Half Hours</th>
<th>One-quarter Hours</th>
<th>Three-quarter Hours</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maintenance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>19-34</td>
<td>20-25</td>
<td>26-30</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>25.8</td>
<td>21.8</td>
<td>28</td>
<td>55</td>
</tr>
<tr>
<td>Final goals</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td><strong>Generality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>26-35</td>
<td>22-34</td>
<td>27-28</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>29.8</td>
<td>27.3</td>
<td>27.5</td>
<td>66</td>
</tr>
<tr>
<td>Final goals</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 4.18. Ranges and means of speed (in seconds) by Student 4 across four skills during maintenance and generality assessments comparing to the final goals assigned to each skill.
**Generality**

Except one mistake made in the third generalization test on half hour times, and another mistake in the first generalization test on three-quarter hour times, Student 4 responded correctly to all the time items in generalization tests across the four skills that he was previously tutored during intervention. On average, Student 4 used 29.8 seconds on half-hour times (final goal of 25 seconds), 27.3 seconds on one-quarter hours (final goal of 25 seconds), 27.5 seconds on three-quarter hours (final goal of 25 seconds), and 66 seconds on five-minute increments (final goal of 60 seconds) respectively. See Table 4.18 for Student 4’s complete generalization assessment scores.

**Summary of All Participants’ Result**

All participants could respond or self-correct correctly to all items in hour time pretests, thus hour time skill was excluded from the tutoring practices and assessments of all participants. Student 2’s 100% accuracy scores in half hour time pretests also exempted him from tutoring practices on half hour times. Due to time limit and absences, Student 3 did not start being tutored on five-minute increments.

**Baseline**

Except Student 4, no participant responded correctly to any time item in half hour and either quarter hour times during baseline. Their accuracy scores remained zero out of 12 time questions across the above three skills during this phase. Student 4’s exception was on half hour times, when his average baseline accuracy score was six correct responses out of 12 items. In telling times in five-minute increments, all participants could correctly responded to a few (i.e., ranged 3-8, averaged 4.6) time items in this skill.
The items correctly responded varied, but mostly included these four types of time items: hour, half hour, one-quarter, and three-quarter hour times.

The participants’ baseline speed scores on each skill varied. However, the speeds were consistent within each participant. In another word, during baseline participants who used longer (or shorter) time to complete a half hour test tended to use longer (or shorter) time on other skill tests, comparing to his or her peer participants. For example, on one-quarter hour times, the baseline speed scores of Student 1 through Student 4 were 45.4 seconds, 65.2 seconds, 40.2 seconds, and 62.2 seconds. Their baseline speeds on three-quarter hour times were 46.5 seconds, 81.3 seconds, 37.4 seconds, and 51.6 seconds, respectively. Student 2 consistently spent more time on each skill’s baseline tests, comparing to other participants, while Student 3 used less time consistently.

***Trial***

Based on the accuracy results of all 14 Trials conducted across participants, eight Trials required adjustments on the originally assigned accuracy criteria which were calculated based on baseline results. All eight adjusted criteria for Treatment Level 1 demanded more correct responses (i.e., ranged 1 to 7 more) from the participant to complete a tutoring assessment. Each participant had at least one criterion be adjusted. For Trials with comparable accuracy results to the original assigned accuracy criteria (i.e., accuracy difference within two), the criteria were retained.

Based on the speed results of all 14 Trials conducted across participants, 10 Trials required adjustments on the originally assigned speed criteria which were calculated based on baseline results. Eight of the 10 adjusted Level 1 criteria allowed more time
for the participant to complete a tutoring assessment. Two of the adjusted Level 1 criteria allowed less time (i.e., 15 and 17 seconds less). Each participant had at least one criterion been adjusted to allow more time; Student 4 was the only participant who had criteria being adjusted to allow less time. For Trials with comparable results to the original assigned speed criteria (i.e., difference within 5 seconds), the criteria were retained. With the adjustments made on the first criterion level for each skill, the subsequent criteria levels were adjusted accordingly.

**Tutoring**

All participants reached 100% or approximately 100% (i.e., average score of 11.8 of 12 items) accuracy during their last criterion or final goal level on each skill. For most skills, this increasing progress from baseline to the ceiling level was not in a gradual and even pace. Instead, all participants tended to make the most improvement on accuracy during their first criterion level. Performances on accuracy after the first level varied slightly between participants and skills. Student 2’s accuracy scores on one-quarter hour times, for example, increased from zero during baseline to an average of 11.8 during his first criterion level. He incorrectly responded to only one item in six of the tests during this level. In the subsequent four criterion levels, Student 2 maintained 100% accuracy on this skill. Student 1’s accuracy scores on three-quarter hours also increased from zero of baseline to nine and 10 in the last two assessments in her first criterion level. Her accuracy performances on her second level went through certain fluctuation with a range of seven to 12. In each of the subsequent three levels, there was always one test when Student 1 missed one or two items. In general, all participants’ average accuracy scores
were in an increasing trend, except one or two occasions when a participant missed one item in a test.

Except for Student 2’s performance on five-minute increment, all participants’ speed scores showed a relatively gradually decreasing trend. The average total decrease ranged from 8.8 seconds (i.e., Student 3 on one-quarter times) to 63.1 seconds (i.e., Student 2 on five-minute increments); or in terms of decrease percentage, the range was from 22.9% (i.e., Student 4 on five-minute increments) to 62.7% (i.e., Student 1 on three-quarter times, and Student 4 on half hour times). Across all skills, the participants spent an average 28.1 seconds or 44.9% less time to complete an assessment during the ceiling level of each skill than they did during the first level. Although the baseline and initial criteria varied between the participants, their performances during the final criterion level all surpassed the comparing group levels of each skill. Student 2’s speed scores also showed the gradual decreasing trend except for Level 3, when the scores increased rapidly (i.e., 78 seconds or more than 100% increment of speed score in four consecutive sessions). Level 3 was thus terminated and a special level with no speed criterion limit was initiated for Student 2. The gradual decreasing trend resumed when criterion limit was re-assigned in the subsequent level. Student 2 was the only participant who took the special level.

_Maintenance and Generality_

Maintenance data were collected every three school days after CATS were completed on each skill. In cases of weekends or participants’ absences, maintenance tests were postponed to the first available weekday. In general, all participants maintained
a 100% or close to 100% accuracy (i.e., missed one time item in one maintenance test) on all skills which they previously practiced on during tutoring. Similar accuracy results were obtained in the Generality assessments, which were given on the same days when Maintenance assessments were conducted. Student 3 is the only participants who maintained 100% accuracy on all Maintenance and Generality assessments.

Maintenance and Generality results in terms of speed showed differences among the participants. On average, Student 1, 3 and 4 spent a few seconds more (0 - 4 seconds) or less (2 – 4.6 seconds) in Maintenance assessments than the final goal of that skill they had been practiced on during tutoring. These three participants almost exclusively used more time (1.2 – 8.4 seconds) in Generality assessments, except that Student 3 spent an average 2.2 seconds less in half hour skill’s Generality assessments. Student 2 performed differently from the other three participants in that he exclusively used less time to complete Maintenance and Generality assessments. In addition, the time differences on five-minute increments were greater: Student 2 spent an average 19.3 seconds (i.e., almost 25%) less than his final goal on five-minute increments during Maintenance, and 12.5 seconds (i.e., 16%) less during Generality assessments (See Table 4.19 and Table 4.20).
<table>
<thead>
<tr>
<th></th>
<th>Half Hour</th>
<th>One Quarter Hour</th>
<th>Three Quarter Hour</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>0</td>
<td>14.2%</td>
<td>4.4%</td>
<td>2%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>0 s</td>
<td>3.2 s</td>
<td>1 s</td>
<td>1 s</td>
</tr>
<tr>
<td><strong>Student 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>N/A</td>
<td>-19.6%</td>
<td>-13.4%</td>
<td>-24.9%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>N/A</td>
<td>-5.8 s</td>
<td>-4.1 s</td>
<td>-19.3 s</td>
</tr>
<tr>
<td><strong>Student 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>-15.1%</td>
<td>-7.0%</td>
<td>1.5%</td>
<td>N/A</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>-5.0 s</td>
<td>-2 s</td>
<td>0.4 s</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Student 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>9.6%</td>
<td>-4.1%</td>
<td>16.7%</td>
<td>-4.5%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>2.3 s</td>
<td>-0.9 s</td>
<td>4 s</td>
<td>-2.6 s</td>
</tr>
</tbody>
</table>

Table 4. 19. The percentage difference (%) and difference in seconds (s) between the average speeds in Maintenance assessments and the final goals of all participants. A negative number indicates a speed faster than the final goal.
<table>
<thead>
<tr>
<th></th>
<th>Half Hour</th>
<th>One Quarter Hour</th>
<th>Three Quarter Hour</th>
<th>Five-minute increments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>17.7%</td>
<td>5.2%</td>
<td>13%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>4.3 s</td>
<td>1.2 s</td>
<td>3 s</td>
<td>2.3 s</td>
</tr>
<tr>
<td><strong>Student 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>N/A</td>
<td>-6.1%</td>
<td>-7.4%</td>
<td>-16%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>N/A</td>
<td>-1.8 s</td>
<td>-2.3 s</td>
<td>-12.5 s</td>
</tr>
<tr>
<td><strong>Student 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>-6.7%</td>
<td>4.4%</td>
<td>11.6%</td>
<td>N/A</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>-2.2 s</td>
<td>1.3 s</td>
<td>3.2 s</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Student 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Difference</td>
<td>26.6%</td>
<td>20.2%</td>
<td>14.6%</td>
<td>14.6%</td>
</tr>
<tr>
<td>Difference in Seconds</td>
<td>6.3 s</td>
<td>4.6 s</td>
<td>3.5 s</td>
<td>8.4 s</td>
</tr>
</tbody>
</table>

Table 4. 20. The percentage difference (%) and difference in seconds (s) between the average speed in Generality assessments and the final goals of all participants. A negative number indicates a speed faster than the final goal.
Social Validity

At the completion of the maintenance and generalization assessments, each participant was given a one-page questionnaire to complete (see Appendix C). Due to limited time allowance, participants took the questionnaire home to complete. Addressed and stamped envelopes were provided along with the questionnaire. The participants were addressed to complete the questionnaire by themselves and return it using the envelope provided. Participants were strongly encouraged to seek their parents’ help if there were unclear meanings of words or sentences. All participants returned their questionnaires.

According to the results of the questionnaire, all four participants had access to computer at home. In addition, Student 2 and 3 used educational software, while Student 1, 3 and 4 played games on home computers. All participants but Student 2 had other peer tutoring experience before this study. They reported that they liked those prior experiences. They also indicated that they liked working with their partners in this study, and they preferred being the tutor to being the tutee. Student 2 had never participated in other tutoring program; he indicated that he did not like working with his partner (i.e., Student 1), and preferred to be the tutee instead. Except Student 3, all participants liked this time-telling program in general. Student 3 reported that he like it “sometimes”. Student 3 was also the only one who found that this time-telling study did not help him to be a better tutor. The other three participants agreed that the study helped them to become a better tutor. All participants indicated that this program helped them to learn telling time. For the two open-ended questions, Student 1 expressed that the best thing
about this computer program was that she liked “the timing bars to beat my record”, and the least likable thing about this program was seeing her partner (Student 2) be “upset”. Student 2 liked the best to be a tutee and liked the least to be a tutor. From their legible handwriting, it seems that Student 3 and 4 expressed similarly that the best thing about this computer program was that it is playful. They wrote “fun”, “game”, and “(like to) press the buttons”. Student 3 indicated that the worst thing about the program is “do the fives”, while Student 4 found nothing he did not like about the program. When asked if they would like to do computer-assisted tutoring again in the future, Student 1, 2, and 4 expressed positively, while Student 3 reported that “maybe” he would. In addition, Student 2 verbally expressed to the researcher after the questionnaires were collected that he “had fun” from the program, which taught him things about time that he did not know (“… I didn’t know what those are when the short hand is not pointing to a number”).
CHAPTER 5
DISCUSSION

This study was a systematic replication of McKain’s (2004) study with important modifications. The main purpose of this study was to examine the effectiveness of an intervention package on students’ acquisition of time-telling skills and tutoring skills. The package was similar to that used in McKain’s (2004) study, but with modifications in software design, procedure, and data collection to address problems raised in McKain’s study report. The study also examined the students’ retention of knowledge and generalization of time telling skills under conditions different from training.

This chapter discusses the results of the study and compares the findings with those reported in McKain’s study. This chapter also addresses the limitations of this study that might have affected the outcome and implications for possible users of the program. Implications for practice and suggestions for future research are also discussed.
Research Questions

Research Question 1

What are the effects of the Computer-Assisted Tutoring System (CATS) on the accuracy of time-telling skills to the half hour, quarter hour times and five-minute increments with elementary school students with emotional disorders?

Data from the Computer-Assisted Tutoring System (CATS) with all participants suggest that CATS resulted in an increase in accuracy of recognizing half hour, two types of quarter hour times and five-minute increments. That is, before CATS the participants correctly recognized none or only a few time items, while daily assessment data showed that all participants recognized the time items of the above four types with 100% accuracy or an approximate level when CATS terminated. Further, their performance increased only at the point when CATS was introduced and only at the point when the stepwise change in criterion level was introduced. Hence, a functional relation between CATS and accuracy was established. Finally, this accuracy reached or exceeded the average level of typically developing students of same grade in the same school.

The results are consistent with of McKain (2004)’s study, which showed increases in acquisition of similar time telling skills from zero knowledge to comparing group level on participants. McKain’s accuracy results were compromised in that the participants acquired some knowledge of the skills through training before intervention started, since the same computer program was used during training sessions. Current study eliminated this error by implementing training software that retains the functions of the program but contains no time-related items. The baseline data of the participants in
this study reflected that no incidental learning occurred before intervention started. Thus, the functional relation between CATS and the accuracy of the time telling skills among the participants is more convincing in this study.

A phenomenon observed in accuracy data is that all participants tended to make the most improvement during their first criterion level, while performances on accuracy after the first level varied slightly among participants and skills. Most involved slight fluctuations before their accuracy scores became stable at 100%. If speed is disregarded, tutoring could have been terminated after the second criterion level for some participants on certain skills. Similar results were reported by McKain (2004) when participants reached quarter-hour time telling comparing group within as few as four sessions. This phenomenon suggests that the participants demanded more time to practice to be proficient on time telling skills; while given enough time to respond, the participants did not demand much time to learn the time items.

Except for McKain (2004)’s study, the researcher is not aware of any previous study that shared a similar combination of reciprocal peer tutoring and a computer assisted time telling program to teach elementary students with mild disabilities. Thus, the results of this study are not highly comparable to previous tutoring studies, computer-assisted instruction studies, or studies that used other strategies to teach time-telling skills. Still, in terms of accuracy improvement, results of this study confirmed previous experiments on the effectiveness this type of tutoring system on students’ gains on knowledge facts, such as sight words with students with disabilities (e.g., Al-Hassan, 2003; Harrison, 2002). Similar results are found in the few studies that a computerized
program was combined with a tutoring program different from the current one, to
improve school aged students’ spelling abilities (e.g., Abbott et al., 2006; Greenwood et
al., 1993), when students reached average accuracy level of 90% or more.

In this study, however, students’ improvements showed a ceiling effect when their
accuracy scores reached approximating 100%. It was difficult to tell the students’
potentials on the skills had other dimensions of the behaviors not being measured, (i.e.,
speed to complete a 12-item assessment as in this study). The ceiling effect is determined
by the skill contents: for example, there are only 12 items in a half hour skill group. No
more learning items could be added to this skill group, compared to tutoring on sight
words (e.g., Al-Hassan, 2003; Heron et al., 1983), or math facts (Guy et al., 2002), where
more words or math facts cards can be added when needed. The computerized software
was not flexible to accommodate students’ learning in terms of their progress in accuracy
improvements.

It may need further investigation and component analysis to determine how each
element of the computer assisted tutoring system functioned, and in what proportion, to
increase the accuracy of time-telling skills of the participants. However, the combination
of the two elements showed positive effect in this study on increasing the participants’
accuracy of telling the four types of times.
Research Question 2

What are the effects of CATS on the speed of time-telling skills to the half hour, quarter hour times and five-minute increments with elementary school students with emotional disorders?

Data from CATS with all participants suggest that CATS resulted in a decrease in the time spent to recognizing half hour, two types of quarter hour times and five-minute increments. That is, before CATS the participants used an average 45% longer time or 28 seconds more to complete one test than the participants did during the ceiling levels of CATS. The participants’ ceiling level speeds have reached or exceeded the average level of typically developing students of same age and grade in the same school.

Again, it is difficult to decide how each element of CATS functioned, and in what proportion, to decrease the time spent of time-telling skills of the participants. Still, the combination of the two elements showed positive effect in this study on decreasing the participants’ time telling the four types of times.

Although speed was the dimension measured in this study, the results of speed suggested participants’ proficiency level. This is because that each speed measured was the time used to complete a 12 time item test. Speed scores suggested how fast the students responded to a unit number of items.

No speed or proficiency data were measured in McKain (2004)’s study, however, time consuming problem was reported anecdotally as tedious and frustrating. This may be interpreted as participants’ lower proficiency despite of satisfactory accuracy results in
McKain’s study. Also, McKain reported that arguments between tutoring partners occurred more often and on a more severe level than current study. No contingent time limit to restrain students’ time spent on tutoring could be another factor that extended each tutoring session. McKain’s study revealed demands on proficiency requirements in addition to accuracy measurements. Similar peer tutoring programs also weighed more on accuracy building than proficiency development (e.g., Al-Hassan, 2003; Guy et al., 2002; Lo & Cartledge, 2004). Even certain computerized studies that applied a peer tutoring forms different from the current one did not report proficiency data (Abbott et al., 2006; Greenwood, 1993). Emphasis on accuracy over proficiency of reciprocal peer tutoring may have been determined by the nature of the peer tutoring format. That is, involving proficiency measurement would complicate the tutoring procedure and training. For younger children or students with disabilities, proficiency measurements may not be possible at all without additional assistance from the teacher.

In terms of improving accuracy and proficiency on skill acquisition using reciprocal peer tutoring, the current study contributed to this form of tutoring by involving a goal-setting procedure controlled by a computerized program. The program presents constant time elapse and tutee’s performance through the competing timing bar and question item bar. The results of current study showed that more sessions were needed to reach the comparing group speed level than to reach comparing group accuracy level. Or, as mentioned in previous discussion, the participants demanded more tutoring sessions to be proficient on time telling skills, while given enough time to respond, the participants did not demand much time to learn the time items. The results implied that
besides accuracy building, proficiency building is another key factor for students receiving peer tutoring to catch up with their typically developing peers. Computer-assisted instructions can take advantage of the computing ability of computers to simplify proficiency measurements (Greenwood et al., 1993). Involving proficiency measurements in peer tutoring using computers should be a topic worthy of investigation in future studies. Still, more computer programs need to be designed to improve this peer tutoring formats on various skill learning.

**Research Question 3**

*What are the effects of CATS on students’ maintenance of accuracy of time telling skills?*

Very few mistakes occurred in maintenance assessments across all participants. Student 2 and 3 made no mistake in all maintenance tests. Student 1 and 4 had responded incorrectly to one item only. Across participants no mistake occurred in five-minute increments that was the time telling skill with the most variability. The participants’ performances in maintenance tests indicated that on a high accuracy level the participants maintained the time items they practiced in tutoring three, six, nine, and in a few occasions, 12 days after the tutoring practice terminated.

Consistent maintenance on accuracy in this study was different from the maintenance results of McKain (2004)’s study, in which participants showed a drop of upwards of 50% two days after tutoring ceased. Considering the different age and disabilities between participants from the two studies, the lower accuracy maintenance in McKain’s study could be the results of the participants’ lower functional level, their
disabilities, and younger age. Still, additional studies may reveal if lack of proficiency building interfered with participants’ maintenances of skills acquired.

High maintenance results (i.e., 90% or more on previous tutored knowledge items) were obtained in previous studies implementing the same peer tutoring format but on different skills and student populations (e.g., Al-Hassan, 2003; Guy et al., 2002). For the few previous computerized tutoring programs, however, since the tutoring formats and emphases of the studies were different from the current study, no maintenance data on students’ performance over time were reported (Abbott, 2006; Greenwood et al., 1993).

Neither the current study nor McKain’s study contained an active strategy for programming the maintenance of time telling acquisition. Such strategies do not appear to be included in other studies using similar peer tutoring programs either, although most studies measured participants’ maintenance.

Research Question 4

*What are the effects of CATS on students’ maintenance of speed of time telling skills?*

Overall the lengths of time all the participants used to complete maintenance assessments were comparable to their final goals assigned to each skill. Maintenance assessment results were also comparable to the average speed scores the participants achieved during their ceiling level. Except for one case in which Student 2’s average time used in five-minute increments was 19 seconds less than his average speed in the last criterion level, all maintenance speed averages were 0 to 5.8 seconds less or more than average in the last criterion. Most speed averages were within 3 seconds’ difference. The
results indicated that on a high proficient level the participants maintained the skills they practiced during tutoring three, six, nine, and in a few occasions, 12 days after the tutoring practice terminated.

Student 2 was the only participant who achieved less average time than the final goal level in all maintenance tests (See Table 4.7), suggesting that Student 2 had potential to decrease his time use on all the skills. Considering that during maintenance assessments, tutoring was terminate for the skill assessed, while Student 2 continued to make progress on speed scores. This may suggest that the previous tutoring had inertial effect on Student 2’s performance. Or, a hypothesis with research value is that taking assessment itself could function as practice. Still, other unknown factors could result in Student 2’s different performance.

The researcher is not aware of findings similar to Student 2’s speed performance during maintenance tests in other peer tutoring studies. The differences in the formats of the peer tutoring between current study and other studies prevent meaningful comparisons between the maintenance speeds. Time limit for the study prevented further investigation of Student 2’s potentials in improving speed.

Research Question 5

What are the effects of CATS on students’ generalization of time telling skills under untrained conditions in terms of accuracy?

Similar to the maintenance assessment results, very few mistakes occurred in generality assessments across all participants. No more than one mistake was made in one test, if there was any mistake at all. No pattern could be found in the mistakes, which
occurred in all four time skills. The participants’ performances in generality tests indicate
that on a high accuracy level the participants could generalize the time items they
practiced in tutoring to times on various forms of clocks printed on pictures.

Instead of measuring participants’ stimulus generality when learned time items
were presented in a format (i.e., pictures of various clock faces) different from that used
during intervention (i.e., one clock face on computer screen) in this study, McKain
measured the participants’ response generality by testing untrained time items. McKain’s
accuracy results were much lower (i.e., mean of 42%) than current study (i.e., close to
100%). An apparent explanation is the different aspects measured between these two
studies. Students may have less difficulty to recognize the same time item on a clock face
of different color and shape than a time item that they have never practiced on before.
Again, results of students of different ages and disabilities may not produce meaningful
comparisons.

Research Question 6

What are the effects of CATS on students’ generalization of time telling skills under
untrained conditions in terms of speed?

Overall the lengths of time all the participants used to complete generality
assessments were comparable to their final goals assigned to each skill. Generality
assessment results were also comparable to the average speed scores the participants
achieved during their ceiling levels. The results indicate that the participants could
generalize the skills they practiced during tutoring with comparable speeds under a
different condition. In general, participants used longer time to complete a generality test
than a daily assessment. This might be contributed by the differences in the formats of the assessment: instead of clicking one button to move on to the next question, the researcher flipped pictures of time items during a generality test. Time used was measured manually by a stopwatch.

Most of other studies implementing the same peer tutoring program tested participants’ generalization on the skills, usually on practiced items under altered conditions. For example, Al-Hassan examined students’ generalization by having students read sentences that were created using the sight words practiced during tutoring. The results showed a relative high generalization level (e.g., 86% to 98%). Guy et al. (2002) tested students generalization of math facts practiced in a peer tutoring using flashcards by giving students paper-pencil exams that include math problems the same with those used during tutoring. Similar results were obtained.

Again Student 2 was the only participant who used on average less time in all generality tests (See Table 4.8). Specifically on the skill of five-minute increments, he spent on average 12 seconds less in a generality test than he performed in the last criterion level. On the same skill, Student 2 was also the only participant who had a special session in which the speed criterion or goal was temporally removed because of his emotional behaviors observed when being timed. In generality tests, the timing pressure was not visually presented to the participants as the timing bar did in daily assessments. Generality test results indicate that Student 2 might have potentials to achieve a higher speed level if more time allowed for the study.
Research Question 7

What are the participating students’ opinions of the use of CATS?

The results of the questionnaire demonstrated varying opinions about the tutoring format and the computerized program. In general, three of the four participants who had previous peer tutoring experiences expressed their positive perspectives on their previous and current experience, and on working with their partner. They all also preferred to be in the role of a tutor. In contrast, Student 2 who did not have previous tutoring experience expressed opinions opposite to his peer participants. According to anecdotal notes taken during intervention, Student 2 had multiple occasions of tantrums and had to spend a few minutes in a timeout room before he was returned to resume his tutoring session. Usually before a tantrum, Student 2 intended to accuse his partner of disturbance during tutoring and testing, when his partner was performing the tutoring behaviors as taught in training. Student 2’s lack of previous experience, tantrums and their consequences might have affected his opinions. Still Student 2 confirmed that the program helped him to be a better tutor. Behavior problems also occurred in Student 3 on three days when tutoring sessions had to be terminated for the day. These might also affect Student 3’s opinions as he disagreed that the program helped him to be a better tutor. As to the function of the program, all participants agreed that this program helped them to learn telling time.

The general confirmation among the participants on the function of the program in terms of improving learning are in agreement with previous studies that implemented peer tutoring systems similar to the current study (Al-Hassan, 2003; Lo & Cartledge,
2004; McKain, 2004). Despite some participants’ confirmations on their interests in the program, similar to what McKain (2004) reported, there was one participant having disagreement on the program’s helping him to be a better tutor, and another one did not like the program sometimes. A few minor arguments between tutoring partners were also observed in this study as in McKain’s study.

The current study did not collect opinions of the classroom teacher, since the implementers were the researcher alone, and no other assistance was needed during tutoring. Other studies on the same format of peer tutoring (e.g., Al-Hassan, 2003) or computerized program on a different tutoring format (e.g., Abbott, 2006) included information provided by classroom teachers, who were at least partially involved in the studies. Al-Hassan used a questionnaire similar to the one used in this study for participants. The results showed that in general the teachers agreed positively on the effects of the programs to the students’ target behavior improvements and on using the program in the future. Abbott et al. (2006)’s study collected satisfaction data indirectly by recording the percentage of teacher’s task completion. The results varied largely (42%-100%) between schools, which was suggested by the authors as the results of difference school administration efforts, teacher training, and individual ability to handle technique challenges. From anecdotal notes of current study, the classroom teacher found the computer software and the peer tutoring program easy to comprehend and implement. Had the classroom teacher been the program provider alone, it is reasonable to expect similar positive feedback on the effects of the program from the teacher, considering the results of the students’ performance improvements.
Summary of Research Findings

This study contributes to previous research on implementing a tutoring system to teach and practice academic skills on young elementary students. CATS comprised major elements of previous research on peer tutoring, including tutor training, tutor’s corrective feedback, and short, fast-paced daily practice and assessments (e.g., Cooke et al., 1983; Harrison, 2002; McKain, 2004; Miller et al., 1996). The results of the study confirm previous investigations on the effectiveness of tutoring systems in increasing accuracy and fluency of academic skills among elementary students. Particularly, this study confirms and extends previous findings (McKain, 2004), which involved using similar computerized program to teach time telling skills to similar aged students with more profound disabilities.

The current study further extended McKain’s study by involving a more advanced computerized program that allowed goal setting, timing bar presenting, and included fewer and exclusive time type in each assessment.

Different from McKain’s study that measured only accuracy performances among the participants, students in this study were assigned with a speed goal to meet in each assessment. The participants could see their performances comparing instantly to their goals through the two timing bars at the bottom of the computer screen. Each goal was adjusted to a more difficult level when the participants showed speed by meeting the goals for consecutively three sessions.

Each assessment in this study contained 12 items instead of 24 mixed types of time items as in McKain’s study. The results obtained in the current study show a greater
increment in accuracy than results obtained in McKain’s study. This may due to the
functional levels of the participants: participants in this study were first graders with mild
disabilities, while participants in McKain’s study were preschool children with more
severe disabilities. In addition, results of this study also show the participants’ decrease in
time spent to complete assessments. Still, the new features included in this study’s
computer program might also have contributed to the differences in the results, although
further investigations are needed to confirm this suspicion.

Another feature changed in the computer program of this study is in daily tutoring
practice trials, where a “Next” button replaced the previous “Correct/Incorrect” buttons
as used in McKain’s study. This change emulates this program more to a conventional peer tutoring in which flashcards are used: as tutors do not place cards in correct or incorrect piles during tutoring as they do in a test. This modification may seem trivial in improving students’ performances. However, a “Next” button could have saved a tutor’s efforts and time during tutoring. As reported in McKain study, arguments on the button clicking occurred often among the participants during tutoring. Plus, each student spent long time (i.e., 20 minutes) to complete one tutoring session which made the daily tutoring tedious and frustrating to both the classroom teacher and the participants in McKain’s study. These problems did not occur in current study, although certain disturbing behaviors were observed a few times in Student 2 and 3. Changes made in this study, such as the “Next” button in tutoring, the timing bars to limit time spent on tutoring and assessments, and uniform time type and fewer numbers in one skill group, might have assisted to lessen the problems raised in McKain’s study.
Limitations of the Study

Participant Characteristics

One potential limitation of this study concerns the characteristics of the participants. The classroom teacher assisted the study’s integrity by not teaching time related subjects in class before the completion of this study. The participants were also addressed not to practice telling time in occasions other than the tutoring sessions. However, when approaching the completion of the study, Student 1 told the researcher that she liked telling time so much that she often read time on the clock at home in those days. The researcher reminded her that she was not supposed to practice telling time at home. The researcher inquired other participants if they practiced reading time in other occasions, Student 2 and 4 also confirmed that they read clocks at home more often than before. It is not clear to what extent Student 1, 2 and 4 practiced the skills outside of the study. If the amount of time and efforts were considerable, the results of this study may not generalize to students who would learn telling time exclusively in school, since it could be possible that Student 1, 2 and 4’s interests and extra practices influenced their scores.

On the other hand, this limitation may reveal a possible positive aspect of the program: it might have raised students’ interest in the contents they practiced within the program to such an extent that they unintentionally generalize the skill to real-life settings. Indeed, such generalization can hardly be controlled experimentally.
The participants’ characteristics in terms of their particular disabilities in this study should be a limit to generalize the results of this study to typically developed students or students with disabilities and function levels different from the participants.

*Lack of Comparing Instruction*

Another limitation of this study is that no data were collected on the effectiveness of current instructions on time telling skills given by the classroom teacher had this study not interrupted the teacher’s curriculum. Despite the effectiveness demonstrated from data collected during this computerized tutoring intervention, it is not known whether a regular class instruction on time telling would produce the same results under the same conditions (i.e., same amount of time devoted on the subject, participants with same age and academic level, and same teacher efforts). Without such comparison, conclusions on the effectiveness of CATS in this study are not complete and convincing in terms of its practical setting application value.

*Skill Grouping*

The four time-telling skills were arranged in an order with assumed increasing difficulty levels. The order was: half hours, one-quarter hours, three-quarter hours, and five-minute increments. There were no documented findings that indicate three-quarter times are more difficult to learn than one-quarter times. Neither were there reasons to believe that quarter hours were more difficult to learn than half hours. Also, the study did not initiate the four sets of times with one-minute increments due to lack of daily practice time. The five-minute increment set was arranged as the last skill to learn in this study was decided considering the diversity of possible times encountered in a set of skill. For
example, for both quarter-hour time sets, the number of different times for each set was fixed (i.e., twelve possibilities). However, for five-minute increments, the possibility was 144 for each set (i.e., 12 fixed minutes with a possible 12 randomly selected hours). Also, in case of initiating the sets of times of one-minute increments, more skills in number-counting would have been prerequisite skills. Skill grouping becoming a limit of current study is in that some skills may have been overlapping or almost the same. For example, four of the 12 five-minute increments overlap other skills (i.e., hour, half, and quarter hours). Thus the overlapped items might have interfered with experimental control when using multiple baseline design.

It may be more justifiable to sub-group the four skills into three skill categories: 1) twelve random selection from hour (if necessary) and half hour times, 2) twelve random selection from the two sets of quarter times, and 3) five-minute increments without items included in other categories. The grouping and the order applied in this study had been convenient for organizing the items to be learned (i.e., equal number of items in each set of skill). In addition, twelve items in each set was close to what had been applied in other peer tutoring studies (e.g., ten math facts were practiced and tested in each session in Guy et al, 2002). McKain (2004) found that eight item sets were better accepted by the preschool participants than 24 item sets in her time learning study). Considering children’s attention span, this adjusted three category grouping may be more reasonable in terms of number of items included. Also skill overlapping issue is addressed in determining the program’s effectiveness.
Time of Study

This study was conducted over a three month period, beginning in middle October, and terminated in middle January of the following year, across both Thanksgiving and Christmas breaks. The researcher collected data Monday through Friday, approximately 15 to 20 minutes for each pair of students. The daily period for the study was right before lunch time. Besides the holidays, Student 3 took one week off for an unexpected family event. Student 4’s holidays were extended when he was absent two days before and two days after the school breaks. It is hard to decide if or to what extend the performances of Student 3 and 4 were affected by their interrupted tutoring sessions. When partner was absent, a student would not serve as the tutor, and would be tutored and tested by the researcher. Thus the student would lose a chance to practice tutoring skills. Student 3 expressed in his questionnaire that he found that this program did not help him to become a better tutor. Further more, Student 3 stated in the questionnaire that he liked the program “sometimes” and “maybe” would participate in similar peer tutoring in the future, while all other participants responded that they liked the program and would participate. This might have been affected by the frequent absence of his partner Student 4, although Student 3’s absences were not a concern of his partner. Intervening by the researcher in case of partner absence was a convenient solution, but would be a concern when partner’s absence was frequent.

Due to limited daily time allowance for the study, none of the participants was able to initiate the sets of one-minute increments. Because of the length of the study,
primarily only three criterion levels were assigned to the skill of five-minute increments, comparing to five criterion levels for other skills which were initiated earlier. The students’ potential speed in the skill of five-minute increments could be more than the final goal assigned if more time were allowed.

Suggestions for Future Research

Accuracy Control

In this study, time telling behavior accuracy was less directly and less immediately reinforced than speed, which might have led to weaker control of the criteria on response accuracy. Future research may investigate the effectiveness of the contingencies on both accuracy and speed in a similar Computer-Assisted Tutoring System.

All assessments in this study presented a timing bar along with an item counting bar at the bottom of the screen. The timing bar shrank proportionately with the criterion time (i.e., the goal) while the item counting bar shrank 1/12 of its length with a completion of each item. In this way, the participant could view his or her process in terms of meeting the speed criterion. However, there was no such visual cue indicating the accuracy criterion during a test. The participant could read the accuracy result immediately after the test was completed. Occasionally the participant also viewed his or her performance charts at the beginning and the end of the daily tutoring. The presentation of charting intended to function as a goal setting for accuracy. Still, viewing charts before and after tutoring was a weaker contingency comparing to the shrinking timing bars during assessments. It had been observed in Student 2 who, in order to “save
time” and “beat the timing bar”, skipped and said “I don’t know” to items which he could tell correctly in the previous tutoring session. Student 2 tried to receive a more immediate reinforcer by completing a test when the timing bar was still long, that is, using much less time than the criterion allowed. Some disturbing behaviors were also observed simultaneously on Student 2, including panting, hand-twisting, rigid posture, arguing with his partner for “not clicking quickly”. The researcher had to apply special tests (i.e., using pretest software with no timing or item bars) for Student 2 in a special criterion level. Disturbing behaviors were much less intensive and occurred fewer times on the other three participants. Thus no special test was needed.

Anecdotal observations of the participants suggest that except Student 1, all participants showed certain uncomfortable behaviors (i.e., panting, rigid posture) when the timing bar presented, despite the participants’ overall positive comments on the computer program and the study. These behaviors could be manifestations of the disability the participants were diagnosed. This may also indicate that visual time-limit, as direct and immediate as it is, may raise tension and lead to accuracy regression in some students. Possible research may compare multiple contingencies (e.g., visual time-limits, chart review, verbal reminders) on student’s accuracy performance to discover an efficient while less disturbing way with this population.

Generalization

As no strategies had been purposefully applied in this study to promote generalization of telling times on real clocks or watches other than the clock faces in the computer program, future research may address this issue. One way of promoting
generalization of telling time is to present times learned in the study on a variety of clock or watch faces of different design, size, color, at different situations (e.g., required by different people, at different time, for different purposes).

As verbally indicated by the participants in this study, four sets of 12 times to practice daily were “too many”. To simplify daily practice and to promote generalization of time telling at the same time, a more efficient way could be instead of teaching all possible time items, researcher can teach a few examples from each sets. For example, students could practice on six randomly selected times from each of the five time sets: hours, half-hours, quarter-hours, five-minute increments, and one-minute increments. In this way, tutoring load is lowered from 108 items (i.e., $12 \times 9$) to 30 items (i.e., $6 \times 5$). When the student’s performance reaches certain criterion and remains stable, the rest 78 unlearned time items could be used to promote response generality.

Comparing to Conventional Instruction

To investigate the relative effectiveness of CATS examined in this study, it is worthy of further experiments to compare the results to conventional classroom instructions on these time telling skills. A future study could compare the total number of trials needed for students to reach comparing group level of telling time by implementing a computerized tutoring program and by regular classroom instructions alone. The experiment should keep other elements equal between the two interventions, such as daily time and teacher efforts devoted to CATS, participants’ age, academic levels, and their disabilities. Teachers’ opinions on using both instructions should also be gathered to decide the program’s social significance.
Involving External Reinforcers

No external reinforcer was purposefully provided to the participants in this study, either on academic performances in daily assessments or on tutoring behaviors. However, in previous peer tutoring studies, reinforcers other than tutor’s instant feedback (e.g., “good job”, “correct”) were provided. For example, in Cooke et al.’s study, as part of the external reinforcing system, students charted their own learning process by coloring blocks of the number of correct items they responded right after the daily assessment (Cooke et al., 1983). In Guy et al.’s (2002) study, students were verbally praised and rewarded with stamp marks for their appropriate tutoring behaviors.

A future study could examine the effects of additional reinforcers on this tutoring package by including the students’ performance scores during the tutoring to a reinforcing system that the students were familiar with, such as an established token economy. The scores could be results of assessments on the target skills, or data of students’ tutoring behaviors. The reinforcing system connected to this tutoring package could be in a variety of forms. Since a computer has been involved in this tutoring package, it might be a possible reinforcer to have the students fill their assessment scores in a spreadsheet each time after tutoring, and view their program on a linear chart. An Excel® spreadsheet prepared by a researcher would be all needed to be the potential reinforcer. Because of its simplicity, a spreadsheet reinforcer would be worthy of investigation on its effectiveness. Findings of future studies on additional reinforcers may contribute to discovering optimal ways to implement this tutoring package.
Exploring Efficient Implementation of the Program

The performance of Student 2 on his maintenance assessments suggests possibilities that the system could be implemented in other ways to be more efficient.

Student 2 was the only participant who, on average, used less time than the final goal to complete maintenance tests on all skills. During maintenance assessments, tutoring was terminated for the skill assessed, while Student 2 continued to make progress on speed scores. Despite of other possibilities, this may suggest that the previous tutoring had inertial effect on students’ performance, and taking assessment itself could then function as practice. Also, the results of this study imply that the participants demanded more time to practice to be proficient on time telling skills, while given enough time to respond, the participants did not demand many trials to learn the time items accurately. A future research could examine the effects of the same tutoring package with a simplified procedure, in which tutoring practice is terminated when accuracy reached certain level (e.g., on average 11 out of 12). Data on daily assessments on each skill would be collected to observe if students’ speed levels continue to improve through assessments only.

Exploring the Efficiency of Tutoring Behaviors

Another suggestion for future research concerns the effects and efficiency of some tutoring behaviors. The peer tutoring protocol applied in this study was adopted from Cooke et al.’s peer tutoring study, in which major tutoring behaviors include verbal prompts for answer, praise to correct response, or corrective feedback to incorrect
response, and laying correctly answered and incorrectly answered items in different piles for recording. In the current study, however, only the tutors’ feedback-providing behaviors were recorded in tutoring sessions, and only button-clicking (i.e., equivalent to laying cards in two piles) was recorded in testing sessions when deciding if tutoring skills were mastered. Prompting behaviors were trained, but were not insisted during tutoring. Neither was the form of praise. The only tutoring behaviors that were insisted were corrective feedback to mistakes, and clicking corresponding buttons to record scores. This modification is because during training the researcher observed that tutors’ prompting and praise behaviors had drifted heavily, while tutoring went through with no misunderstanding among the participants. It seems that the appearance of a new time item on screen is itself a prompt salient enough to raise the attention of the tutee for a response. Tutor’s verbal prompt of “What time is it?” became redundant. There are occasions observed when tutors voluntarily prompted their partners when delay was made. Similar behavior drifting was observed in tutors’ praise to correct answers in tutoring trials (since no feedback was supposed to be given in testing trials). It occurred with a variety of degrees in all participants that their praise faded from the initial verbal “Good job” to a slight and fast lip moving “Right” or a nod. McKain (2004)’s study reported similar tutor behavior drifts and suggested that lack of using proper tutoring procedures did not seem to adversely affect the skill acquisition.

A future study may investigate the efficiency of prompting and praise in peer tutoring by comparing the results of peer tutoring programs that insist full scale tutor prompt and praise with programs that apply the least intrusive tutor behaviors (e.g., nods,
interrupt with corrective feedback only when mistake occurs, keep the question flow if no mistake occurs). Other elements in the comparing programs need to be strictly controlled, such as participants’ academic level or age, tutoring content and procedure.

Simultaneously, the researcher should record the number of trials or items a tutee can practice on during tutoring sessions in the comparing two programs. This is to find out if less intrusive tutor praise will provide tutee with more practice opportunities.

*Teaching Functions of Telling Time (Estimation and Calculation)*

Teaching times of one-minute increments was not included in this study. It would have been a more complete study to include all possible time items. To read times of one-minute increments may be beneficial for students to build number concepts. Still it is often not the function of telling time. For example, under most cases “3:30” is functional enough comparing to “3:31” or “3:29”. Future research may explore the potentials of this computerized tutoring program in teaching time-estimation and calculation. After being fluent in reading times of one-minute increments, students can be tutored to estimate a time when minute hand is not strictly pointing to a calibrating dot on the clock face. With students who have established certain calculation abilities, a research could also examine the effectiveness of tutoring time calculation using such programs. This may require revisions of the software program. For example, a program is revised to be able to present two clock times and prompt for lapse of time in between (e.g., “How many minutes between now (4:00pm) and 4:40pm?”). Still, tutor is able to provide corrective feedback by hearing from the earphones in such a program.
Summary

This study examined the effects of the Computer-Assisted Tutoring System (CATS) on the accuracy, speed, generalization, and maintenance of time-telling skills with four first grade students with behavior disorder and learning disabilities. The tutoring program was adopted from the peer-tutoring model created by Cooke et al. (1983). Some procedures from the Cooke et al. model were altered. For example, peer tutoring folders were replaced by the software created by the researcher.

When CATS was completed, all four participants reached or surpassed the comparing group levels of accuracy and speed in all the time categories which they were tutored. All participants maintained their accuracy and speed levels two weeks after CATS was terminated. The participants also achieved comparable scores when presented with time item pictures of clock faces different from that used in tutoring.

The results demonstrated that the CATS package of a computerized program and peer tutoring is an effective alternative to classroom instructions to teach elementary students with mild disabilities to tell time.
LIST OF REFERENCES


Harrison, T. (2002). *The development of a peer tutoring program to teach sight words to deaf elementary students*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.


McKain, K. N. (2004). *Effects of computer-assisted peer tutoring on acquisition, maintenance, and generalization of time telling with primary students with developmental delays*. Unpublished master thesis, The Ohio State University, Columbus, OH.


APPENDIX A

INFORMATIONAL LETTER AND CONSENT FORM
September 24th, 2005

Dear Parent or Guardian,

My name is Ma Yao and I am currently a graduate student in Special Education at The Ohio State University. An important requirement for completing my course of study is to conduct a research project. I am in the process of preparing a research that I will be carrying out at your child’s school. I will be conducting the research under the supervision of Dr. Ralph Gardner, a professor in the College of Education. I am writing to you to explain my research to you and to ask your permission for your child to participate. The following is a description of the study I am planning to conduct.

My study will examine how fast and how well students can learn to tell time using a computer-assisted peer tutoring program. The study will follow the procedure of the peer tutoring system developed by experts at The Ohio State University. The system has shown to be effective in many other researches and will be examined in my study as well. In such a peer tutoring system, your child will participate as one in a pair of students taking turns acting as learner and teacher. What I am going to do differently is that instead of using traditional materials such as flashcards, your child will learn using a set of computer software I developed to practice telling time. The software will train your child to be a teacher (tutor), show times on a clock face in practice, and keep track of times your child have learned and how fast your child can tell in daily test. Your child will participate in a daily 20-minute session in the school library.

No session will be video- or audio-taped. Your child’s name will not be revealed in any kind of publication, document, or any other form of report or presentation developed from this study. I am excited about this study project, and I would be grateful if you would grant permission to your child’s participating in this study.

Attached are two copies of the consent form for Participation in Educational Research. By signing this consent form you agree to allow your child to participate in this study as described in this letter. Please sign and return one signed copy of the consent form using the attached stamped envelope to me, and keep the second copy for your records. If you have any questions about details of this study, feel free to contact me at (614) 421-2048 or Dr. Gardner at (614) 292-3308, or by e-mailing me to ma.125@osu.edu.

Sincerely,

Ma, Yao
Ph.D. Student

Barbara XXXXX
Principal

XXXXX Elementary

Ralph Gardner
Professor and Faculty Advisor

Enclosures: 2 copies of Consent Form for Participation in Educational Research
I consent for my child’s participation in a research study evaluating the effects of a Computer-Assisted Tutoring System (CATS). Ms. Ma Yao will conduct this study under the direction of Dr. Ralph Gardner. Ms. Ma Yao has explained to me the nature, purpose, procedure to be followed, and the expected proficiency of my child participation. I understand that my child's identity will not be revealed to anyone other than people directly involved in the study, or by means of publication, documentation, or any other form of report and presentation developed from this research. Additionally I understand that I have right to withdraw my consent for my child’s participation at any time during the study.

I acknowledge that I have had opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction.

I acknowledge that I have read and fully understand this participant consent form. I sign it freely and voluntarily. An additional copy has been given to me.

_______________________________________  __________________
Signature of Parent or Guardian (Person authorized to consent for participant )  Date

_______________________________________  __________________
Signature of Participant  Date

_______________________________________  __________________
Ma Yao  Date
Ph.D. Student Researcher

_______________________________________  __________________
Dr. Ralph Gardner  Date
Professor and Faculty Advisor
APPENDIX B

SAMPLE ILLUSTRATION OF PRINTED DATASHEET IN A DAILY ASSESSMENT
<table>
<thead>
<tr>
<th><strong>Tutee Name</strong></th>
<th>Johnny <em>(student 1)</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tutor Name</strong></td>
<td>Lisa <em>(student 2)</em></td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>10/28/2005</td>
</tr>
<tr>
<td><strong>Skill category</strong></td>
<td>Half hour times</td>
</tr>
<tr>
<td><strong>(A)</strong> No. of Correct Answers</td>
<td>4</td>
</tr>
<tr>
<td><strong>(D)</strong> Speed (in seconds)</td>
<td>66</td>
</tr>
<tr>
<td><strong>(AC)</strong> Accuracy Criterion</td>
<td>5</td>
</tr>
<tr>
<td><strong>(FC)</strong> Speed Criterion</td>
<td>48</td>
</tr>
</tbody>
</table>
APPENDIX C

SOCIAL VALIDITY SURVEY TO PARTICIPANTS
Social Validity Survey to Participants

Name:_________________________  Student No. ___

Please answer the following questions:

1. Do you like this time-telling program in general?
   Yes  No  Sometimes

2. Have you had other peer tutoring experience before this program?
   Yes  No

   If “yes”, did you like those peer tutoring programs in general?
   Yes  No  Sometimes

3. Did you like working with your partner?
   Yes  No  Sometimes

4. Did you like using the computer-assisted time-telling programs used in this study?
   Yes  No  Sometimes

5. Do you have access to computer at home?
   Yes  No

6. If “yes” to question number 5, do you use educational software or play games on
   your computer?
   Yes (educational software; game)  No

7. Which one did you prefer to be, tutor or student?
   Tutor  Student  Same

8. Did the computer help you to be a better tutor?
   Yes  No  A little

9. Did using the computer help you learn to tell time?
   Yes  No  A little

10. What did you like best about computer-assisted tutoring?

11. What did you like least about computer-assisted tutoring?

12. Would you like to do computer-assisted tutoring again?  Yes  No  Maybe

178
APPENDIX D

SOFTWARE INTERFACE SAMPLES
Sample Interface of Tutor Training Software for Tutoring Procedure Practice
Sample Interface of Tutor Training Software for Assessment Procedure Practice
Sample Interface of Software for Pretest
Sample Interface of Tutoring Software for Practice
Sample Interface of Tutoring Software for Daily and Maintenance Assessment
APPENDIX E

INTEROBSERVER AGREEMENT (IOA) DATA COLLECTION FORM
### Interobserver Agreement (IOA) Data Collection Form

<table>
<thead>
<tr>
<th>Session #</th>
<th>Tutor</th>
<th>Date &amp; Time</th>
<th>Testing Button Click</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**A**

<table>
<thead>
<tr>
<th>Name</th>
<th>Start</th>
<th>Print</th>
<th>Tutoring</th>
<th>Prompting</th>
<th>Feedback</th>
<th>IOA = A/(A+D) × 100%</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Name</th>
<th>Start</th>
<th>Print</th>
<th>Tutoring</th>
<th>Prompting</th>
<th>Feedback</th>
<th>IOA = A/(A+D) × 100%</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**C**

<table>
<thead>
<tr>
<th>Name</th>
<th>Start</th>
<th>Print</th>
<th>Tutoring</th>
<th>Prompting</th>
<th>Feedback</th>
<th>IOA = A/(A+D) × 100%</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX F

INTEROBSERVER AGREEMENT (IOA) DATA COLLECTION FORM
ON SOFTWARE PROCEDURE
Interobserver Agreement (IOA) Data Collection Form on Software Procedure

<table>
<thead>
<tr>
<th>Component</th>
<th>PC</th>
<th>Observer</th>
<th>Agree / Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of correct answers as recorded by tutor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testing/Tutoring speed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing bar disappears when goal speed ends (Y/N)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item bar shrinks 1/12 in length after completion of each item (Y/N)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
A = \text{________} \quad \text{IOA} = \frac{A}{A+D} \times 100\% \\
D = \text{________} \quad = \text{______}\% \\
\text{Total} = \text{________}
\]

Form No.: 1

\[\square \text{ Check if no printed result}\]

Observer __________ Form No. 1, 2 __________

<table>
<thead>
<tr>
<th>Component</th>
<th>PC</th>
<th>Observer</th>
<th>Agree / Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of correct answers as recorded by tutor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testing/Tutoring speed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing bar disappears when goal speed ends (Y/N)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item bar shrinks 1/12 in length after completion of each item (Y/N)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
A = \text{________} \quad \text{IOA} = \frac{A}{A+D} \times 100\% \\
D = \text{________} \quad = \text{______}\% \\
\text{Total} = \text{________}
\]

Form No.: 2 __________

\[\square \text{ Check if no printed result}\]
APPENDIX G

SCRIPTS FOR TRAINING
Scripts for Training

Session 1

1) An orientation to the computer tutoring program (2 minutes):

Today we are going to play with a computer program. Each of you is going to learn to teach each other using this computer program. We will work in pairs and take turns to be the tutor and the student. The tutor is the one to teach. What we are going to learn is time. But for these a few days we only learn how to use this computer program to name pictures. You will see 12 pictures one after another in the computer. A tutor will ask “what is it?” and a student will answer. When we are good at using this computer, we will use a same program to learn time. Are we clear? Let’s see what we need to play: this is the computer, a pair of earphones, and you all know where the printer is. So let’s play.

2) Modeling of using training-tutoring software by experimenter and a student (5 minutes):

(Have the training-tutoring program opened on screen). Now I will play teacher, or a tutor. Anyone wants to play the student? (Wait for a volunteer and pick one; if none, ask one participant “#####, how about you?”) I, the tutor, will sit on the right, and the student will sit on the left. This is the first screen on the program. It has words and two boxes. Today I’m the tutor, so I will type my name in the first box here (Type in “Mayao” after “tutor”). ##### (the volunteer) is the student, he (she) will type his (her) name in the second box (the student type in name after “tutee”). When both our names are typed in, I, the tutor, will put on the earphones. Are you ready? Let’s play. I will click
this START button (Move cursor and click; the first picture question appears). “What is it?” (Wait for the volunteer’s response). “Good job” (if the answer is correct). I can hear the answer in here from the earphone, ####’s answer is right, so I said “good job” and then I click this NEXT button to move on (do so). “What is it?” (Wait for response). If I am not sure if #### is right, I can click this button to listen in my earphones the answer again. (Go through the program without description but procedural feedback only). (When the volunteer makes a mistake, or initiate the volunteer to make a mistake), “Try again”. This is what the tutor say when your student makes a mistake. (Volunteer corrects the answer). “Good job”. (Initiate the volunteer to make a mistake and not able to correct the answer for a second time). “Say ‘(the answer)’”. You may all have noticed this orange bar down here. It becomes shorter and shorter with time passes. It is a timer. The bar will completely disappear after 5 minutes. What we need to do is to say as many and correct answers as we can in these 5 minutes before the bar disappears. What will happen when it disappears? (Continues the program until time is up). See, the program stops. Time is up, and practice is over. So, remember that, your will have only 5 minutes to practice, when the bar disappears, your practice is over for the day. Let’s see I do a complete tutoring with #### (the second observer’s name).

3) Modeling of using training-tutoring software by experimenter and the second observer (5 minutes): Go through the program following the procedure. The second observer, as the tutee, will be wrong on 7 questions randomly and deliberately. Of the 7 questions, 5 will be wrong for a second time which requires tutor’s corrective feedback. No descriptive comments are interrupted.
4) Each participant’s role-playing as tutor using the software while the experimenter as tutee (25 minutes):

Now I will play the student and each of you will be the tutor one by one. (Go through the program following the procedure with each 4 participants. The experimenter serves as the tutee will purposefully make mistakes of the same kinds and quantity as in step 3. The second observer will correct the tutor’s mistakes when occurs).

Session 2

1) Reviewing previous session and introducing testing and charting (5 minutes):

Can we still remember what we learned yesterday? (Go through the tutoring procedure using training-tutoring program for a few images. The experimenter serves as the tutee and the participants serve as tutors).

2) Modeling of using training-testing software by experimenter and the second observer (10 minutes):

Today we are going to learn how to give a test. We are using another program that looks just like the one we used before, but have some differences. (Have the training-testing program opened on screen). We will type in our names as before and I, the tutor, will start testing by clicking START (do so). See what we do. (Giving a test to the tutee, second observer, following the procedure in daily assessment. The tutee makes 5 mistakes purposefully and randomly). When we finish, see there is an orange button with a letter “P” on it. As soon as I see the button, I click it. Then I go to the printer and get the printed scores. (Re-start the testing program). Can anyone tell me what looks different
here comparing to what we used yesterday? (Wait for answer. Cue for answer). Right, there are two bars down here, one is red and another is blue. We can see that the red bar becomes shorter and shorter like the orange bar we have seen; while the blue bar becomes shorter each time I click a button to the next question. You are wondering what they are for. Let me tell you. As you may have found out, the blue bar is counting how many questions left for the test. Each time you finish a question and move on, the blue bar becomes one question shorter. When it disappears, your test is done. The red bar is a timer like in tutoring. How many minutes is it? Five minutes? No. The time is the goal for ####. That is to say, #### has to answer the questions as correct and fast as she can to meet the goal. You will find out that unlike the orange bar we see in tutoring, when the red bar disappears, it won’t stop testing. You still need to finish all 12 questions. And you are right again that we also have two round buttons, one is red with an “X” and another is green with a check mark. Each time when #### (second observer’s name) was right, I click the green button and each time when she was wrong, I click the red button. Another thing you have found out is that I do not say “good job” or tell #### the answer when she is wrong again. And I have to click the PRINT button at the end as quickly as I can. See how we do it (present testing one more time).

So a tutor should do what in a test? One, zip your mouth, don’t tell anything. Two, click red button when your student is wrong and green when he (she) is right. Three, do it quickly. The blue bar should shrink faster than the red bar to beat the goal. And four, click PRINT quickly once you see it.

3) Keying-in results to the Excel® spreadsheet and reviewing result (10 minutes):
Now let’s have a look of the printed paper. It has my name as tutor, and ####’s name as student. See down here there are two words: “Accuracy” and “Speed”. Behind “Accuracy” is a number “7”. That is #### answered 7 questions correctly. Behind “Speed” it says “30 seconds”. That is how long it took #### to answer all the twelve questions. (Open a sample Excel spreadsheet). Now is the time for #### to type in her scores into this spreadsheet. Find today, and type in the scores under “Accuracy” and “Speed”. (#### does so). Now look at the chart next to it. See the blue and green dots? Those are ####’s test scores. See the dotted lines? Those are the goals set for #### today. ####’s goal for Accuracy is 5 today. #### got 7 correct. The dot is above her goal line on this chart (pointing to the dot and the goal line), so #### beat her goal. Congratulations! ####’s goal for Speed today is 28 seconds. Her score is 30 seconds. She needs to do it more quickly and use less seconds to beat her goal. She didn’t meet her goal for speed. But she will practice more and make it!

Each of you will have a chart of your own that looks just like ####’s. Each day after test, you will do the same: to key-in your scores and we can have a look of how fast you have learned. So what we want in our chart? We want to be right on more questions that our Accuracy dot is on or above our Accuracy line; but for Speed, we want to complete the questions as fast as we can, so our Speed dot is below our Speed line.

4) Each participant’s role-playing as tutor giving daily assessment to the experimenter (25 minutes):

Now it’s your turn to give a test. I will play the student. (Go through a test following the procedure with each 4 participants. The experimenter serves as tutee will
purposefully make 4 to 7 mistakes. The second observer will corrects tutor’s mistakes when occurs).

*Session 3&4*

(Review how to use the tutoring and testing programs by having the participants play as tutor and tutee in turns using both programs. The experimenter and the second observer will score on each tutor’s behaviors).
Scripts for Time-telling Instruction

1. Recognizing and counting number from 1 to 12 (approximately 5 minutes)
   Today we are going to play with this clock (show the Judy clock). As you see there are numbers go in a circle on the clock face. So first, let’s count these numbers until you can count them fast and well. You are going to count with me, and then I am going to ask each of you to count. (Pointing to each number when counting. Repeat until fluent. If process does not go smoothly, split the 12 numbers to 4 groups and focus on one group at a time).

2. Direction to which the clock hands go (approximately 3 minutes)
   You all see these two hands here (pointing to the hands on the clock). One is a shorter fat guy and the other is a longer skinny guy. Now let me show you how they walk on this clock face. (Turn the button at the back of the clock for a few circles and stop). Now who can tell me which way the hands go, this way (show in clock-wind) or that way (counter-clock-wind)? (Wait for students’ answer. Cue and prompt if necessary). Remember how we count the numbers? So the hands run the same way as we count the number, right? (Re-count the numbers with the students if necessary).

3. Discrimination of hour hands and minute hands (approximately 15 minutes)
   a. Hands go at different speed
      Can anyone tell me: are the shorter fat hand and the longer skinny hand running together? (Turn the hands again and wait for answer). No they are not.
Which of them runs faster? The longer skinny one. In fact, when the longer skinny hand runs for a whole circle, the shorter fat hand goes only one number further. (Show by slowly turning the minute hand for a circle starting from 12 and finger pointing to the hour hand’s track). Let’s see more (Continue another circle).

b. Hand names

The shorter fat hand is the hour hand. One hour is long, so it goes slowly. The longer skinny hand is the minute hand. One minute goes fast, so the minute hand, the longer one, goes faster.

c. Tell hour time

When we tell time from a clock face, we want to tell the hour first. Can anybody tell me which the hour hand is? (Wait for answer). Right, the shorter fat hand. We need to see where he is. If he is pointing to one, that is hour one, we say “one”. If he is pointing to two, that is hour two, we say “two”. So, whichever number the short fat hand points to, we say that number first. Let’s try. (Practice more hours with the minute hand pointing to 12).

d. O’clock, fifteen, thirty, and fifteen

After we say the hour, we say the minute. For the minute hand – which is the minute hand? (Wait for answer). Right, the longer skinny hand – we are not going to read the number it points to. Reading the number where it points to is only for the hour hand, remember? So how we read minutes? We are going to learn only a few of them. When this longer skinny hand points to 3, we say “fifteen”. Repeat after me, “fifteen”. When it points to 6, we say “thirty”. Repeat after me, “thirty”. When
it points to 9, we say “forty-five”. When it points to 12, we say “o’clock”. (Turn the
minute hand to the 4 positions in the order of o’clock, fifteen, thirty, and forty-five,
and repeat with the students until fluent. Then ask students to name when randomly
point the minute hand to one of the 4 positions).

e. Tell time

   When we tell time, we read the hour hand (the shorter hand) first, and then
read the minute hand (the longer hand). For example (turn the hands to 8:00), we
find the hour hand first, which number is it pointing to? (Wait for answer). Right,
eight. Then we find the minute hand. It points to 12, so we say “o’clock”. So the
time is, “eight o’clock”. (Present a few more examples; use all the four minute-
hand positions to be learned).

4. How to read an hour when the hour hand is not strictly pointing to a numeral
   (approximately 5 minutes)

   Remember we only read the number when the shorter fat hour hand is pointing to
it. You may have found that there are times when the shorter fat hour hand is not directly
pointing to a number. (Turn the hands to 2:30). Like this, it is not pointing to 2 or 3,
which number should we pick as the hour? We always pick the number that is smaller,
except for 12 and 1. When the hand is between 12 and 1, we pick 12. If the hour hand is
between 4 and 5, which one do we pick? (Wait for answer). We pick the smaller one.
(Give a few more examples using the Judy Clock).
(In this section, if the students have difficulty judging between numbers, the instructor can also teach using the following scripts):

Remember the way the hands go? (Show direction by turning the button). When the hour hand is between two numbers that we need to pick one to say the hour, we think of having the hour hand goes backward a little bit. So it’s going to be the number that the hand goes backward. So what hour is this (Showing 2:30 on the clock)? We think of the hand goes backward a little bit. It will reach 2. So we pick 2. (Give a few more examples using the Judy Clock).