COMPUTER ASSISTED INSTRUCTION TO IMPROVE THEORY OF MIND IN CHILDREN WITH AUTISM

Lindsey R. Eason, B.A., M.Ed.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2014

APPROVED:

Smita Mehta, Major Professor
Bertina Combes, Committee Member
Prathiba Natesan, Committee Member
Kevin Callahan, Committee Member
Abbas Tashakkori, Chair of the Department of Educational Psychology
Jerry R. Thomas, Dean of the College of Education
Mark Wardell, Dean of the Toulouse Graduate School
Eason, Lindsey R. *Computer Assisted Instruction to Improve Theory of Mind in Children with Autism.* Doctor of Philosophy (Special Education), December 2014, 205 pp., 13 tables, 8 figures, references, 117 titles.

Children with autism spectrum disorder (ASD) show significant deficits in communication, emotion recognition, perspective taking, and social skills. One intervention gaining increased attention is the use of computer assisted instruction (CAI) to teach social, emotional and perspective-taking skills to individuals with ASD with the purpose of improving theory of mind skills. This study evaluated the effectiveness of CAI for improving theory of mind skills in four children with high functioning autism ages 5 to 12 years. A single-subject multiple baseline research design across participants was utilized to evaluate the effectiveness of CAI. The software contained 22 instructional scenarios that asked participants to identify emotions of characters based on situational cues displayed in line drawn pictures and audio feedback for correct and incorrect responses. Mind-reading skills were assessed using ten randomly selected scenarios for various emotions and no audio feedback. Visual analysis of the data revealed that all four participants increased mind-reading skills during the CAI condition. Additionally, this study evaluated levels of task engagement during experimental conditions. Three of the four participants showed an increase in task engagement during CAI compared to paper-based social stories used during baseline. Generalization of skills was assessed through the use of social scenarios acted out by family members of participants. All four participants were able to correctly identify emotions displayed in generalization scenarios. Results demonstrated that CAI was an effective and socially viable method for improving ToM skills in children with autism and they could generalize their skills to untrained settings.
ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my advisor and committee chair, Dr. Smita Mehta for her unwavering patience, support, and encouragement during my entire doctoral program. You have instilled in me the value of precision and persistence. You have also taught me the process and standards for high quality research. Additional gratitude is extended to the US Department of Education and the authors of the Systematic Training for Autism Researchers and School Personnel Grant (US Department of Education- H325D060017-09), also known as Project STARS, for providing full funding for the coursework that led up to the completion of this study. Without the support of this grant and the faculty members who authored the grant I would not have been able to accomplish these things for children with autism, their families, and the professionals who work with children with autism. I would also like to thank my other committee members Dr. Callahan, Dr. Combes, and Dr. Natesan for the time they dedicated to reviewing my dissertation, providing valuable feedback, and helping me through the dissertation process. Additional thanks to the College of Education for a dissertation award that assisted me in data collection expenses. Next, I offer great thanks to Nicole Caldwell for her time spent reviewing computer logs and coding video data. I could not have completed this study without her huge contribution. I would also like to thank Joseph Hoffmann, Nicholas Daniel and Vishal Malhotra at the UNT CLEAR center for their assistance in creation of the software program used for this study.

Finally, I would like to acknowledge the love and support provided by my family and friends through this long journey. They have been great cheerleaders continually reminding me of where I came from and what my ultimate goals are.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>COMPUTER ASSISTED INSTRUCTION TO IMPROVE THEORY OF MIND IN CHILDREN WITH AUTISM</td>
<td>1</td>
</tr>
</tbody>
</table>

## Introduction                                                                                     | 1   |

Theoretical Foundation of CAI                                                                       | 4   |

CAI Interventions for Children with ASD                                                            | 7   |

Generalization of ToM Skills                                                                       | 15  |

Need for Expanding the Existing Evidence Base on Effective Interventions                        | 18  |

Purpose                                                                                           | 19  |

## Method                                                                                           | 21  |

Participants and Setting                                                                        | 21  |

Measurement of Dependent Variables                                                                | 27  |

Materials                                                                                          | 36  |

Experimental Design and Procedures                                                                | 41  |

Investigator                                                                                      | 50  |

## Results                                                                                         | 50  |

Effectiveness of Computer Assisted Instruction on Mind-Reading Skills                             | 50  |

Effect of CAI on Rate of Task Engagement                                                          | 54  |
Effect Size for Mind-Reading and Task engagement ................................................. 58

Discussion.......................................................................................................................... 58

Overcoming Mind-Blindness............................................................................................. 59

Addressing Issues with Response Generalization....................................................... 62

Impact of Instructional Delivery Method on Task Engagement.................................... 65

Implications for Practice ............................................................................................... 67

Indicators for High Quality Research ........................................................................... 68

Limitations and Directions for Future Research ......................................................... 69

References ......................................................................................................................... 70

THE USE OF COMPUTER ASSISTED INSTRUCTION FOR INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: A SYSTEMATIC LITERATURE REVIEW .......................................................... 82

Introduction ...................................................................................................................... 82

Theoretical Foundation of CAI ..................................................................................... 85

CAI Interventions for Children with ASD ................................................................. 88

Generalization of ToM Skills ....................................................................................... 90

Need for Expanding the Existing Evidence Base on Effective Interventions .......... 92

Purpose of Literature Review ....................................................................................... 93

Method ............................................................................................................................. 94

Locating Studies .......................................................................................................... 94

Inclusion/exclusion criteria ......................................................................................... 94

Results ............................................................................................................................ 96
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant Characteristics</td>
<td>26</td>
</tr>
<tr>
<td>2. Social Validity Results from Parent Questionnaire</td>
<td>35</td>
</tr>
<tr>
<td>3. Social Validity Questionnaire completed by Participants</td>
<td>36</td>
</tr>
<tr>
<td>4. Procedural Fidelity Checklist for Baseline</td>
<td>48</td>
</tr>
<tr>
<td>5. Procedural Fidelity Checklist for CAI</td>
<td>49</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Example of teaching mind-reading skills to children with autism ............................................. 28
2. Sample log of the internal data collection system ........................................................................ 29
3. Direct observation recording form for task engagement ............................................................... 31
4. Sample social narrative for baseline procedures ........................................................................... 37
5. Social situations for role playing during skill generalization ....................................................... 41
6. Graphic organizer for research procedures ................................................................................... 43
7. Percentage of mind-reading skills for all participants ................................................................. 53
8. Rate of task engagement for all participants .............................................................................. 56
COMPUTER ASSISTED INSTRUCTION TO IMPROVE THEORY OF MIND IN CHILDREN WITH AUTISM

Introduction

The *Diagnostic and Statistical Manual of Mental Disorders (DSM 5)* defines autism spectrum disorder as impairment in social communication and interaction across multiple contexts, failure to initiate and respond to social interactions, as well as restricted and repetitive patterns of behavior. Other characteristics associated with ASD are deficits in social-emotional reciprocity and deficits in understanding body language (APA, 2013). These deficits are often exhibited as an inability to understand and process subtle social and emotional cues expressed through facial expression, voice intonation, and context cues (Tager-Flusberg, 2007). These characteristics can contribute to the display of overt and covert forms of problem behavior such as tantrums and/or social withdrawal, which limit an individual’s ability to successfully participate in their social environment (Bauminger, 2002). One specific characteristic of individuals with ASD is a deficit in theory of mind (ToM) skills or what is referred to as mind-blindness (Colle, Baron-Cohen, & Hill, 2007).

Theory of mind (ToM) skills, also known as mentalizing, mind reading, or social intelligence, is defined as an ability to infer other people’s mental states such as thoughts, beliefs, and intentions and use the information to interpret and predict what others will say and do in specific social situations (Hutchins, Bonazinga, Prelock, & Taylor, 2008). ToM plays a critical role in interpreting communicative intent to see beyond spoken words and infer or hypothesize about a speaker’s mental state. Additionally, it allows individuals to self-reflect, realize fallibility in personal beliefs, and use social communicative skills of persuasion to repair breaks in conversation (Howlin, Baron-Cohen & Hadwin, 1999). Inaccurate assessment of the
mental state of others can create social challenges such as inability to interpret sarcasm or figurative language, recognize others intent to deceive, and maintain reciprocal communication (LaCava, Golan, Baron-Cohen, & Myles, 2007). Individuals with ASD usually show significant delays in reaching ToM milestones when compared to individuals without disabilities. To demonstrate this delay, Baron-Cohen (1989) developed a ToM task in which individuals with ASD were asked to evaluate pictures of either an entire face or the eye region and determine the correct facial emotion. Results indicated that individuals with ASD performed significantly lower than age and IQ matched peers at identifying emotions on both types of pictures.

Similarly, the individuals with ASD also performed markedly lower on the eyes-only condition than their age-matched peers in a follow up study by Baron-Cohen and colleagues (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). These results support the notion that emotion recognition is a key to understanding the nature and severity of social deficits in individuals with ASD (Baron-Cohen, Leslie, & Frith, 1985; Howlin et al., 1999).

The ability to recognize emotions is just one component of the complex processes involved in ToM skills (Tager-Flusberg, 2007). There is growing evidence to suggest that ToM can be learned by individuals with ASD through explicit teaching strategies with a specific focus on generalization training (Feng, Lo, Tsai, & Carledge, 2008; Hutchins & Prelock, 2008; Tager-Flusberg, 2007). A variety of interventions have been utilized in an attempt to improve these skills, with mixed results.

Ryan and Charragain (2010) conducted a study with 30 participants (mean age 9.5 years; 20 in the experimental and 10 in the control group) to evaluate the impact of a systematic emotion recognition training program on emotion recognition and emotion-related vocabulary
comprehension tests. Pre- and post-test measures showed that children in the experimental group achieved higher levels of improvement when compared to their peers in the wait-list control group. The latter group showed similar gains after being exposed to the 4-week training program. While the results were promising and supported the role of explicit instruction of emotion recognition, learned behaviors of participants did not generalize to real social situations. Parents of the participants reported that although children attended to faces and accurately identified emotions, they commented and inappropriately laughed after recognizing an angry face. These findings suggest that it may be possible that the correct responses were memorized by participants without cognitively internalizing the concepts.

Similar issues related to lack of transference to daily life situations, presumably because of social cognition deficits, were also displayed in a study by Begeer et al. (2011). They conducted a randomized control trial with 40 participants to evaluate the effectiveness of a 16-week ToM training package with 8-13 year olds with ASD. The intervention package included strategies to explicitly teach ToM skills during 1.5 hours per week in small group sessions. Results indicated that while participants improved their elementary understanding of ToM, they did not show any improvement in more complex skills such as empathy. Additionally, they were not able to transfer learned skills into daily life situations. It is possible that the low density of the intervention or absence of specific generalization training may have resulted in limited effect on participants.

Moving beyond basic emotion recognition to a complex understanding of emotions and empathy is a recurrent challenge faced with ToM interventions. In a study by Feng et al. (2008), ToM training was provided through Macromedia Flash Player to teach a 11-years old male
participant with high-functioning autism to identify emotions in social scenarios. Results indicated that he was able to identify emotions in basic social scenarios (e.g., desire-based emotions and belief-based emotions) but not the more complex skills.

The results of these studies further confirm that ToM is a complex construct which requires a comprehensive training package designed to develop advanced skills beyond rote identification of emotions. These skills must then be generalized to social situations in order to have a greater possibility of impacting life satisfaction and long term social outcomes for individuals with ASD. Howlin and colleagues (1999) have suggested some key elements for developing successful ToM interventions which include: (a) breaking information into small steps; (b) providing instruction in normal developmental sequence; (c) systematic reinforcement of behavior; and (d) errorless learning. These principles are all critical elements of effective instruction that can be easily integrated into a computer assisted instructional (CAI) format for delivery, which could lead to more successful outcomes for teaching ToM skills to individuals with ASD. CAI refers to the use of computers to teach a variety of skills and includes computer modeling and computer tutors (Collet-Klingenberg, 2009).

Theoretical Foundation of CAI

The current literature shows recurrent evidence that learners differ in numerous ways and effective teaching should regard and address these differences in order to maximize student progress (Molenda, 2012). In 1911, Edward L. Thorndike developed general principles of learning including the law of effect which states that responses are learned depending upon whether or not they are followed by a preferred outcome. He proposed the first theory of
instruction in which readers could only progress through a textbook after they demonstrated mastery of the basic content (Thorndike, 1911).

This concept was expanded by Sidney Pressey (1927) who developed the “machine for automatic teaching of drill material.” Pressey’s machine appeared similar to a typewriter carriage which displayed a question with four answer choices in a window. On one side of the machine were four keys, one of which the neurotypical adult user pressed in order to indicate the correct answer. The machine recorded the answer on a counter at the back of the machine and progressed to the next question. Once the user was finished, the person scoring the test placed the test sheet back into the device and noted the score on the counter. The use of this machine was reported to improve efficiency of teaching content which required drill and practice to develop mastery.

In a separate field of study, along the same time period, visual instruction was gaining support (Molenda, 2012). B. F. Skinner further explicated Thorndike’s law of effect to experimentally demonstrate that laboratory animals elicited complex new behavior when the schedule and intensity of consequences were manipulated. Based on personal experience, he then developed a mechanical device for interactive learning which he referred to as a teaching machine (Skinner, 1954). With momentum in the field for promoting visual instruction, it was soon realized by the professional community it was the software, not the hardware, which was responsible for successful learning (Molenda, 2008). This belief shifted the focus from the mechanical machine to the instructional method called programmed instruction.

Programmed instruction allows students to interact individually with the content, record answers, and receive constructive feedback, thus enhancing and improving the rate of learning.
This teaching model is based on the following instructional principles: (a) distinctive units; (b) small steps; (c) active participation; (d) simultaneous correction; (e) processing of learning, and (f) individualized learning rate (Efendioglu & Yelken, 2010). As the technology advanced, the programmed instruction model was enhanced by incorporating audio-visual devices for greater impact and delivered through desktop computers. In schools today, the theory of programmed instruction is incorporated into “e-learning” and computer-assisted instruction is used to teach a wide array of abstract concepts (Clark & Mayer, 2003). With the growing literature base suggesting the use of CAI as a promising practice for providing individualized instruction for a variety of learners (Ramdoss et al., 2011; Wainer & Ingersoll, 2011), there is a need to delineate the most critical components necessary for CAI to be a successful intervention.

Although there is a growing trend in the use of technology to provide individualized instruction to a variety of learners, there are no specific guidelines to distinguish between assistive technologies, computer-based video modeling, technology integrated instruction support, and computer assisted instruction (Anohina, 2005). With specific reference to CAI, Skinner’s definition of teaching machines based on the theory of programmed instruction, allows for a differentiation between CAI and all other technology based interventions (Benjamin, 1988). This definition stipulates that CAI consists of a student completing educational activities through the use of a computer-based software program. There are three criteria that differentiate CAI from either assistive technology or computer-based video modeling: (a) CAI uses an automatic self-controlling device that presents a unit of new information (i.e., antecedent); (b) CAI provides a way for the learner to respond to the information (i.e., behavior); and (c) CAI provides feedback about the accuracy of the learner’s
response (i.e., consequence; Benjamin, 1988). This definition and criteria for CAI was adopted for the current research. The multisensory component along with the controlled and structured learning environment and ability to individualize instruction are some of the features that make CAI a promising intervention tool. Additionally, the delivery of immediate feedback, consistent prompting and cueing to elicit accurate responses, and high levels of reinforcement add to its success.

CAI Interventions for Children with ASD

There is evidence to suggest that children with ASD have difficulty responding to traditional teaching strategies which has led to an increase in the investigation and use of alternative treatments to accommodate for the unique characteristics of children with ASD (Whalen, Liden, Ingersoll, Dallaire, & Liden, 2006). Many students with ASD have responded well to interventions based on visual supports including video modeling (Wainer & Ingersoll, 2011), pictures (Bryan & Gast, 2000; Dettmer, Simpson, Smith-Myles, & Ganz, 2000), and computers (Bernard-Opitz, Sriram, & Nakhoda-Sapuan, 2001). As children with ASD seem to show a strong preference to interventions that involve visual supports, it appears that computers are a promising and logical intervention choice for skills-instruction. Current research has shown that CAI can provide an efficient and effective learning environment for basic skills-instruction (Bosseler & Massaro, 2003; Coleman-Martin, Heller, Cihak, & Irvine, 2005; Hopkins et al., 2011; Moore & Calvert, 2000). The NAC report (2009) described technology-based interventions as an emerging practice for skills instruction on a variety of target behaviors for children with autism (ASD), ages 6-14 years. Additionally, the nineteen studies evaluated in this category included various types of technology-based interventions
compiled in a single category (e.g., CAI, Alpha Program, Delta Messages, the Emotion Trainer Computer Program).

To evaluate the usefulness of CAI in teaching individuals with ASD, Bosselor and Massaro (2003) developed a computer animated tutoring program to teach vocabulary and grammar to eight children with low-functioning autism. Within the software an animated face, “Baldi,” led the participants through vocabulary lessons by visually modeling speech production and providing feedback for student responses. Results demonstrated that CAI was successful in teaching all eight participants a significant amount of new vocabulary words. In a further analysis, six participants were able to generalize these skills by using the vocabulary words to identify novel images (Bosselor & Massaro, 2003). A possible explanation for such a high rate of success in generalization was the training to 100% mastery during the treatment condition and use of additional generalization training if students did not identify two of the three novel images during initial generalization probes.

In another example, Coleman-Martin et al. (2005) compared the use of CAI to teacher led instruction for teaching word identification to three children, one of whom had low-functioning autism. Researchers used the nonverbal reading approach (NRA) to conduct lessons which focused on phonemes and word production. During the CAI condition, the NRA curriculum was converted into a PowerPoint presentation. They used a multiple baseline design across participants with three conditions (i.e., teacher led instruction, CAI + teacher led instruction, CAI alone) to evaluate the relative effectiveness. Results showed that all three participants were able to acquire new vocabulary words during all three conditions but the
acquisition rate for the participant with ASD during the CAI alone condition was significantly higher than the other two conditions.

Recent research has even demonstrated superior outcomes for reducing problem behavior using CAI when compared to paper-based social stories (Mancil, Hayden & Whitby, 2009). Researchers used an alternating treatments design to compare the effectiveness of CAI vs. paper-based social stories for reducing pushing behavior in three children with high-functioning autism. Authors converted a social story into PowerPoint presentation for the CAI condition and participants would alternately read the paper-based or the computer-based social story prior to transition to recess or the classroom. Results indicated that while both methods were effective, participants had slightly fewer instances of pushing during the CAI condition compared to paper-based (i.e., control condition). Additionally, all three participants chose CAI as their preferred intervention method when interviewed at the conclusion of the study suggesting that it is an engaging teaching method. Teachers also reported that the CAI condition was easy to use and they would use it in the future suggesting it is a socially valid teaching method. Establishing social validity is necessary for ensuring that goals, procedures, and/or effects are socially important to participants, parents, teachers, etc. rather than convenient for the study (Gast, 2010).

Recently, the use of CAI to increase ToM skills, specifically emotion recognition, has been evaluated by several researchers, with promising results (Hopkins et al., 2011; LaCava, Golan, Baron-Cohen, S., & Myles, 2007; Silver & Oaks, 2001; Tanaka et al., 2010). Because of the strong visual processing skills in many children with autism and the highly engaging nature of technology-based interventions, there is reason to believe that CAI may be an effective
intervention for teaching the complex ToM skills to this population (Samson, Mottron, Soulieres, & Zeffiro, 2012). This aspect is the focus of the current investigation and is reviewed in detail in the following section.

Existing research on CAI for enhancing ToM skills. Understanding emotions is a foundational skill for social understanding because it is necessary for perspective taking and empathizing (Tager-Flusberg, 2007). Several studies have explored the use of computer assisted instruction to teach the foundational skills of emotion recognition to individuals with ASD. For example, Silver and Oaks (2001) designed the Emotion Trainer which incorporated photographs of people displaying emotions, along with multiple choice cartoon drawings as response options (i.e., happy, sad, angry, and afraid) to the question “How does ________ feel?” They utilized this software as an intervention in a study with 22 participants with high-functioning autism, ages 10-18 years, with 11 each in the experimental and control group respectively. Participants in the experimental group completed an average of 8.5 computer sessions over a 2 to 3 week period. Authors conducted a pre-post assessment to evaluate the impact of the software using the Facial Expression Photographs test, Happe’s Strange Stories test, and emotion recognition cartoons. Results showed significant reduction in the mean number of errors made from the first to the last session but noted that there was a large amount of individual variation related to the number of times a participant used the computer program. However, the limited sample size and lack of procedural fidelity regarding the amount of access individual participants had to the software raises concerns about the validity of these findings.
Similar results were found by LaCava et al. (2007) who taught eight children, ages 8-11 years, to recognize emotions using a software program known as Mind Reading. This software included an emotions library catalogue of over 400 different emotions which were presented in photographs, short movie clips, and audio clips along with contextual examples. Participants used the software for 10 weeks in their home for an average of 10.5 hours over the duration of the intervention condition. Pre- and post- intervention assessment of participants’ scores on a variety of standardized measures including the Cambridge Mindreading Face-Voice Battery for Children (Golan & Baron-Cohen, 2006), revealed a significant improvement in emotion recognition skills for all 8 participants. Similar to the Silver and Oaks (2001) study described above, there was great variability in the way participants used the software and there was little control over the impact of extraneous variables such as parental input and amount of time accessing the software. Additionally, results need to be interpreted with caution because of the small sample size in the study.

In another study conducted by Hopkins et al. (2011), children with high-functioning autism (HFA) and low-functioning autism (LFA), ages 6-10 years, were compared on their responses to Face Say, an emotion recognition training software. Participants were presented with arrays of photographs and schematic drawings of emotions which consisted of six faces. As they were presented with a stimulus item, participants heard a label (i.e., angry, disgusted, scared, happy, sad, or surprised) and were asked to match the corresponding emotion by touching the appropriate photograph or drawing. Interestingly, while both groups of children significantly improved their ability to recognize emotions in a photograph, only the group of children with HFA showed improvement in recognizing emotions in a schematic drawing. These
results indicate that CAI can be effective for emotion recognition training across a range of individuals on the spectrum; however, it appears that a certain level of cognitive functioning may be necessary to improve complex skills that are critical for generalization of behavior to social situations in natural environments. Given that previous research has documented the challenge of teaching generalized emotion recognition and perspective-taking to individuals with ASD (Baron-Cohen 1989; Ryan & Charragain, 2010; Tager-Flusberg, 2007), it may be worthwhile to analyze alternative explanations for lack of skill generalization for this population.

Bolte et al. (2006) evaluated the impact of CAI on emotion recognition at the neurological level in order to help explain why individuals with ASD continue to have difficulty generalizing skills and why there is so much variability in the response to ToM interventions. These authors used the Frankfurt Test and Training of Facial Affect Recognition (FEFA) software to teach 10 individuals with high-functioning autism skills in facial affect identification. Subsequently, they evaluated the software’s cognitive impact by using functional magnetic resonance imagining (fMRI). Results revealed that, even though participants showed significant improvement in emotion recognition on the software program, no significant changes in activation levels were observed in the image of the fusiform gyrus (FG), the area of the brain associated with emotion recognition and affect. This indicates that ToM interventions may be successful at teaching compensatory strategies but training may not necessarily improve cognitive capacity of individuals with ASD as measured using fMRI. The implication of such findings is fundamental for understanding why certain types of ToM skills (e.g., mind-reading or perspective taking) may not generalize to natural settings even for individuals with high
functioning autism. Yet, researchers are compelled to develop training materials to enhance mind reading skills of individuals with ASD (Howlin et al., 1999) given the social relevance of such skills-instruction.

The results of these studies combine to create promising evidence for the ability of CAI to teach emotion recognition skills at the primary level and enhance compensatory strategies for processing emotions. However, the research is limited and has many weaknesses such as small numbers of participants, lack of control within experimental conditions, lack of assessment for generalization, and lack of training at the more complex level of mind reading (i.e., perspective taking). Addressing these weaknesses is the primary focus of the current study.

Existing research on CAI for enhancing task engagement. Individuals with ASD typically display difficulty with social exchanges, which very likely affects their engagement during academic instruction as well. This often results in students with autism being excluded from academic activities which involve whole group instruction, small group activities or even working with a partner (Kluth & Darmody-Latham, 2003).

Whether communicating information about routines, academic content, or social expectations, teachers typically rely on verbal language which, for individuals with ASD, is less effective, attributable to their limited ability to process complex verbal information (Dettmer, Simpson, Smith-Myles & Ganz, 2000). This point is illustrated by Carnahan, Musti-Rao, and Bailey (2009) who evaluated the engagement level of six students with ASD during reading tasks incorporated with visual stimuli through an interactive reader paired with music. It was observed that all participants were more actively engaged in reading tasks during the
interactive reader condition compared to baseline level which used traditional books with no visual supports or interactive materials. This outcome supports previous claims that visually stimulating materials can increase task engagement for students with ASD. Additionally, if the use of technological tools is effective for teaching academic content to students with ASD, it stands to reason that the use of CAI to teach social understanding or perspective taking skills would more likely improve social performance of students.

One possible explanation for an overall lack of task engagement in social situations and live-instruction is aversion to facial gaze (Vivanti, Nadig, Ozonoff, & Rogers, 2008). This theory is supported in a study by Hobson and Hobson (2007) in which researchers evaluated amount of time children with autism spent attending to the face of live models when asked to imitate actions. The results indicated that participants with ASD spent significantly less time attending to the face compared to participants without ASD. This lack of attention to faces and a tendency to avoid eye contact impacts individuals’ ability to imitate actions or engage in social conversations in a meaningful manner.

Strong visual processing skills but aversion to social interaction and reduced ability to encode verbal language may explain preference for computerized instruction, which may be more reinforcing and therefore more engaging than live instruction for individuals with ASD (Kaey-Bright & Howarth, 2012). When individuals are more engaged in the learning environment, their skills are more likely to transfer into untrained environments. This is critical for significant life improvements versus simply improving limited skills in limited settings.
Generalization of ToM Skills

One concern raised by autism researchers is that unless appropriately designed, the use of CAI might reduce opportunities for social interactions and social problem-solving (Bernard-Opitz et al., 2001). Because lack of social interaction skills is a core deficit in autism, it is crucial that interventions used to remediate these deficits focus on generalization training.

Generalization, the ability to transfer a learned behavior acquired during a training activity to another setting, person, or activity with similar but different materials has been identified as a difficult task for children with ASD (Koegel, Egel, & Williams, 1980). To illustrate this point, Froehlich et al. (2012) examined prototype formation in 27 adults with high functioning autism using a dot pattern task and compared them to a control group of 25 neurotypical adults. The task trained individuals to categorize dot patterns into three separate groups. During assessment procedures, participants were shown 45 different dot patterns, the original 9 from training and 36 new patterns which were a distorted version of the training stimuli. They were required to label the pattern as “new” or “old,” and provide the appropriate category to which the dot pattern belonged. The results demonstrated no significant differences between individuals with ASD and the control group in recognizing and categorizing familiar and minimally distorted dot patterns. However, as distortion of dot patterns increased, individuals with ASD identified significantly fewer patterns to the correct category when compared to the control group.

Although dot pattern recognition does not directly relate to generalization of life skills, this study highlights a cognitive processing deficit in ASD which has been characterized by a difficulty processing information from multiple stimuli and difficulty with generalization to
untrained stimuli or situations (Froehlich et al., 2012). Additionally, these findings support previous generalization research which suggests as new stimuli become less similar to previously learned stimuli, individuals with autism have more difficulty transferring skills.

More specifically, research evaluating social skills interventions delivered through CAI has shown mixed results regarding generalization of skills. For example, Sansosti and Powell-Smith (2008) evaluated the use of computer-based social stories with embedded video modeling to increase the social communication in three children with autism, ages 6-10 years, using a multiple baseline design across participants. The results showed that “joining in” behavior increased for all three participants during intervention, which also included adult prompting. However, only one of the three participants generalized these skills to a natural setting (which did not include computer-based social stories, video modeling, or teacher prompting).

The variability in generalization of skills was also demonstrated in the Hopkins et al. (2011) study discussed previously, for teaching emotion recognition skills using the CAI program Face Say. They evaluated the performance of three groups of children, ages 6-10 years (i.e., high functioning autism, low functioning autism, neurotypically developing). Similar to the Sansosti and Powell-Smith (2008) study, all 49 participants increased in their ability to recognize emotions during the CAI intervention phase. However, when authors evaluated how these newly acquired emotion recognition skills impacted social interactions in the natural environment, they found mixed results. Through direct observation of social skills the authors determined that participants with LFA showed a significant decrease in inappropriate and negative interactions but no significant change in total amount of interactions and positive
interactions. In contrast, participants with HFA had a significant decrease in inappropriate interactions and a significant increase in positive interactions but no change in negative interactions or total number of interactions. Additionally, based on parental reports, individuals with HFA generally demonstrated increased cooperation and sharing behaviors whereas individuals with LFA demonstrated reduced negative behaviors and increased self-control. These results must be interpreted with caution however, because of possible parental bias as opposed to direct observation of social skills. Also, there is no obvious link between emotion recognition and cooperation or self-control, which leads to further questions about the validity of these results.

Many published studies that used CAI to teach emotion recognition or increase ToM skills did not include assessment of generalization or exposure to social skills training in the natural setting (Bolte et al., 2006; LaCava et al., 2007; Silver & Oaks, 2001; Tanaka et al., 2010; Whalen et al., 2010). Assessment of generalization of learned behavior is not only critical for social relevance of research, but a critical measure for meeting the criteria for high quality research (Reichow, Doehring, Cicchetti, & Volkmar, 2011).

Existing research has suggested that a lack of generalization may partly be due to stimulus over-selectivity observed in individuals with autism, especially in situations where natural settings do not share common stimuli with simulated instructional conditions (Koegel et al., 1980). Stokes and Baer (1977) stated that generalization occurs only when programmed as an integral component of intervention and not as an outcome one hoped to achieve. This highlights the importance of choosing target behaviors that will more likely meet naturally existing contingencies of reinforcement, such as social interaction with peers for individuals
with HFA, to increase the likelihood that behaviors will generalize to natural settings (Cooper, Heron, & Heward, 2007).

One effective method for programming generalization is to provide sufficient exemplars of stimuli. This strategy involves teaching students to respond correctly to multiple forms of antecedent stimuli and providing opportunities for generalization to novel situations which share some common stimulus properties. To ensure that CAI is an effective intervention, generalization of skills must be incorporated as a critical component of the intervention. Identifying innovative interventions that have empirical evidence of remediating core deficits in autism is critical to the field as this population continues to grow and the prevalence rate increases. This is compounded by a social trend of increasing demands within school and work environments (Adreon & Stella, 2001).

Need for Expanding the Existing Evidence Base on Effective Interventions

Recent statistics published by the Centers for Disease Control and Prevention indicate the prevalence of autism to be one out of fifty children (CDC, 2013), which is an increase of more than 500% in the last decade. With the increasing number of children being diagnosed with ASD, the need for effective educational interventions is becoming increasingly important. While a wide variety of interventions have been designated as being empirically effective by the National Autism Center (NAC, 2009), there are many other interventions that lack empirical support for remediating autism deficits.

The NAC report (2009) described technology-based interventions as an emerging practice for skills instruction. Similarly, interventions for improving theory of mind skills were also designated as having emerging evidence of effectiveness. While interventions in this
category targeted both high and low functioning children with ASD, ages 6-14 years, only four studies met the inclusion criteria for evaluation (Bell & Kirby, 2002; Fisher & Happe, 2005; Gevers, Clifford, Mager, & Boer, 2006; Wellman et al., 2002). The same is true of the recent report published by Wong and colleagues (2014) that lists technology-aided instruction and intervention as an EBP. However, this classification included computer aided instruction, speech generating devices/VOCA, smart phone and tablets, and virtual networks but only 5 of the 20 studies focused on CAI to improve mind-reading skills. In other words, further research in this area is needed.

Purpose

The National Standards Report (NAC, 2009) identified four studies evaluating interventions for increasing ToM skills, all of which utilized a group design. Fisher and Happe (2005) evaluated the relative effectiveness of ToM training and executive functioning (EF) training for 27 participants, ages 6 -15 years. Twenty-one participants with a diagnosis of ASD were randomly assigned to a ToM or EF training group and evaluated after 5-10 days of training on a variety of ToM and EF tasks. Results showed that, generally, all participants increased performance on ToM skills. However there was large variability across participants. When further analyzed, the authors determined that performance on Test for Reception of Grammar was the only significant difference between participants who did and did not improve in ToM skills.

Recently, Golan and colleagues (2010) evaluated the effectiveness of The Transporters, a DVD program which combines animated vehicles with human-like facial expressions, to teach 20 participants with HFA, ages 4 – 7 years, to recognize 15 emotions presented in a PowerPoint
format. Results of their group design study indicated that, following four weeks of watching the program for approximately 15 minutes per day, participants significantly improved in their ability to recognize emotions on the faces of animated vehicles and were also able to generalize these skills to identify emotions in untrained human faces.

Currently, there is no research that has evaluated the effectiveness of CAI to improve ToM skills of individual participants using a single subject research design. This highlights a weakness in the depth and breadth of available empirical evidence because group designs typically mask the individual variance in response to intervention as scores of all participants are pooled and overall performance tends to deviate towards the mean (Gast, 2010). Additionally, because the population of individuals with ASD is relatively small in size yet with a large amount of variability in functioning levels across the spectrum, researchers typically have difficulty recruiting a sufficient amount of participants to enable generalization of results to the population.

The purpose of this study was to evaluate the effectiveness of CAI (utilizing animation software) for teaching specific mind-reading skills (i.e., perceiving things from the perspective of other people) to students with high functioning autism. The specific research questions are as follows:

1. Is there a functional relation between computer-assisted instruction and mind-reading skills for individuals with high functioning autism?

2. Do participants show higher levels of task engagement during computer-assisted instruction when compared to instruction using paper-based social narratives?
3. Is there a functional relation between computer-assisted instruction and generalization of mind-reading skills across settings?

Method

Participants and Setting

The participants for this study included four children who had a determination of autism by their local education agency (LEA) or a licensed psychologist using the DSM-IV criteria (American Psychiatric Association, 2000) (see Table 1). Participants were recruited if they met the following inclusionary criteria: (a) were between the chronological age of 5 and 21 years; (b) displayed verbal skills including expressing wants and needs in complete sentences (5-7 word phrases); (c) passed the pretest measures of basic emotion recognition skills derived from Levels 1 and 2 from the mind-reading curriculum developed by Howlin et al. (1999) with 75% accuracy; (d) scored 35% or lower on selected items on Level 3 of the Howlin et al. (1999) curriculum. No participant was excluded from the study based on race, ethnicity, gender, and/or socio-economic status as long as the inclusion criteria were met.

Recruitment flyers with a description of participation requirements and inclusion criteria were sent to a variety of locations including an autism intervention center, offices of special education directors of three school districts and 6 autism therapy organizations. When parents showed interest in participating in the study, an initial screening was conducted via telephone to establish the pre-requisite skills of their child. Next, I individually met with the candidates and their parents for approximately 30 minutes during which time she interacted with the potential participant in order to gain a better understanding of his/her ability. Additionally I...
conducted a pre-screening assessment using Levels 1, 2 and 3 from the Howlin et al. (1999)
curriculum (i.e., Teaching Children with Autism to Mind-Read) as stated in the inclusion criteria.

The procedures for recruiting participants and conducting the study were approved by
the university’s Institutional Review Board (IRB). The letter for obtaining informed consent was
provided to the parents following recruitment. The informed consent letter included the
purpose and procedures of the study, foreseeable risks and benefits, and participant’s rights.
The procedures for the study commenced after receipt of the signed informed consent letter
from the parents. Four children participated in the study.

Evan was an 11-year old boy of Asian-American descent. He lived in a single-family
household with his biological parents and younger sister. Initial screening and parent report
indicated that he primarily communicated using complete sentences (e.g., “I need to get more
points,” “When can we be done?”) with some repetitive phrases and echolalia. Evan was
evaluated by a licensed psychologist in 2013. The results of the Wechsler Intelligence Scale for
Children – IV (2003) indicated that his full-scale IQ score of 66 was significantly lower than his
academic achievement scores revealed by the Wechsler Individual Achievement Test – III
(2009): math problem solving - 96; word reading - 91; pseudoword decoding - 92; numerical
operations - 98; spelling - 100. Possible explanation for the relatively low IQ score as noted by
the licensed psychologist included inattention and distractibility during test-taking, obsessive
indicated a standard score of 38 which qualified Evan with autism in the mild-moderate range.
During the initial emotion recognition screening (Howlin et al., 1999), Evan scored 87.5% and
100% respectively on Levels 1 and 2 and 32% on Level 3.
Ashley was a 5-year old girl of Asian-American and Caucasian descent. She lived in a single-family household with her biological parents and younger sister. Initial screening and parent report indicated that she primarily communicated using complete sentences (e.g., “I want to open the box,” “Can I have the apple?”) with some repetitive and echolalic phrases. Ashley’s cognitive or intellectual functioning was evaluated by the LEA in 2012 using the Kaufman Assessment Battery for Children – II Nonverbal (2004) which revealed mental processing of 74 (i.e., below average) and learning/long-term retrieval of 97 (i.e., average range). Academic Achievement was assessed using the Developmental Assessment of Young Children (DAYC, 1998). The test revealed an overall standard developmental score of 77 (i.e., below average range) and a standard score of 89 on adaptive behavior (i.e., average range). Additionally she was evaluated for autism eligibility using the Childhood Autism Rating Scale – II (2010). She received a standardized score of 40 suggesting mild to moderate autism. During initial screening process Ashley scored 75% on both Levels 1 and 2 of the emotion recognition tasks and 27% on Level 3, making her eligible to participate in the study.

Trevor was a 9-year old boy of Caucasian-American descent. He lived in a single-family household with his biological parents and younger sister. Initial screening and parent report indicated that he primarily communicated using complete sentences (e.g., “Let’s take a picture,” “Can I have more goldfish crackers, please?”) with some echolalic and scripted speech. Trevor’s cognitive and intellectual functioning was evaluated by the LEA in 2013 using the Leiter International Performance Scales – Revised (LIPS-R, 1997). His full scale IQ was a 73 which is in the below average range with a strength in visual processing score at 81 and performance in the average range on the subtest “Classification.” The examiner cautioned that scores may be
influenced by his level of activity and inattention during testing. His academic achievement was assessed using the Kaufman Test of Educational Achievement – II (2004), which indicated varied scores, for example, a standard score of 105 on letter and word recognition (i.e., average range) and 71 on math concepts and applications (i.e., low range). His adaptive behavior was assessed using the Vineland Adaptive Behavior Scale – II (2005), which showed a composite score of 65 (i.e., low range). Additionally, Trevor was assessed using the Gilliam’s Autism Rating Scale (GARS, 2006) and had an overall Autism Index of 94 which suggests a very likely probability of autism. During the initial screening on the emotion recognition tasks, he scored 87.5% and 100% respectively on Levels 1 and 2 and 23% on Level 3, qualifying him for the study.

Pranav was a 5-year old boy of Indian-American descent. He lived in a single-family household with his biological parents and older sister. Initial screening and parent report indicated that he primarily communicated using complete sentences (e.g., “Can I have a candy?” or “Sister wants to play with you,”) indicating some echolalic speech. Pranav’s cognitive and intellectual functioning was evaluated in 2014 by the LEA using the Kaufman Assessment Battery for Children-II Nonverbal Index (2004) which indicated a standard score of 82 (i.e., slightly below average). The Kaufman Test of Educational Achievement – II (2004) was used to evaluate his academic achievement, indicating a standard score of 143 on letter and word recognition (i.e., superior range), 97 on math concepts and applications (i.e., average range), and 52 on written expression (i.e., significantly below average). Additionally, the Gilliam’s Autism Rating Scale (GARS, 2006) was completed and revealed a standardized score of 122 (i.e., very likely probability of autism). During initial screening, Pranav scored 75% on both
Levels 1 and 2 on the emotion recognition tasks but only 5% on Level 3, making him eligible to participate in the study.

The study was conducted in the homes of individual participants. Baseline and intervention procedures were conducted in a specified location at a computer desk or other area designated by the parent as quiet and free from distractions and suitable for the study. A built-in video recording device on the laptop was used to record sessions for baseline, generalization probes, and CAI phases of the study and a handheld recording device was used to record sessions for the generalization phase.
<table>
<thead>
<tr>
<th>Participant</th>
<th>Cognitive/Intellectual</th>
<th>Academic Achievement</th>
<th>Adaptive Behavior</th>
<th>ASD Determination</th>
<th>Language/Communication</th>
<th>Level 1 &amp; 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evan (m) – 11</td>
<td>WISC-IV: 66</td>
<td>WIAT-III: math problem solving – 96</td>
<td>No Adaptive Behavior Scale Completed</td>
<td>CARS-II: 38 (mild-moderate)</td>
<td>Complete Sentences</td>
<td>87.%; 100%</td>
<td>32%</td>
</tr>
<tr>
<td>Asian American</td>
<td></td>
<td>word reading – 91 Spelling - 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation: 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashley (f) – 5</td>
<td>KABC-II: Mental</td>
<td>DAYC: 77</td>
<td>DAYC: 89</td>
<td>CARS-II: 40 (mild-moderate)</td>
<td>Complete Sentences</td>
<td>75%; 75%</td>
<td>27%</td>
</tr>
<tr>
<td>Caucasian Asian</td>
<td>processing – 74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American</td>
<td></td>
<td>Long-term retrieval - 97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation: 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trevor (m) – 9</td>
<td>LIPS-R: 73</td>
<td>KTEA-II: letter/word recognition – 105</td>
<td>VABS-II: 65</td>
<td>GARS: 94 (very likely probability)</td>
<td>Complete Sentences</td>
<td>87.5%; 100%</td>
<td>23%</td>
</tr>
<tr>
<td>Caucasian American</td>
<td></td>
<td>Phonological awareness – 69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation: 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pranav (m) – 5</td>
<td>KABC-II Nonverbal</td>
<td>KTEA-II: letter/word recognition – 143</td>
<td>SIB-R: 72</td>
<td>GARS: 122 (very likely probability)</td>
<td>Complete Sentences</td>
<td>75%; 75%</td>
<td>5%</td>
</tr>
<tr>
<td>Indian American</td>
<td>Index: 82</td>
<td>math concepts – 97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation: 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Measurement of Dependent Variables

Guidelines for single subject research methodology by Horner et al. (2005) were utilized to operationally define and frequently measure the dependent variables throughout the study. Additionally, measurements were assessed for reliability by a secondary observer and behavior change was evaluated for both empirical and social significance to the participants as well as the field. The following dependent variables were measured for the study.

Mind-reading skills. For the purpose of this study, mind-reading was operationally defined as the ability to correctly identify the emotion of a person based on environmental cues in a picture when the facial expression was masked and no voice clues were provided. During baseline and intervention assessment, mind-reading skills were measured as a mutually exclusive correct or incorrect response. A correct response was recorded when a participant used the cursor to click the icon representing the emotion (shown through a line drawn face) that matched the social situation displayed in the picture presented on the computer screen (see Figure 1).
Figure 1. Example of Teaching Mind-Reading Skills to Children with Autism software for CAI instruction. Includes prompting (e.g. red X) and feedback.

During generalization, mind-reading was measured when a participant, at the first attempt, correctly verbalized the emotion that the parent or sibling felt in a specific role-played scenario (described below). Non-examples of mind-reading during baseline and intervention assessment included a participant choosing an emotion that did not correctly match the social situation displayed, displaying or verbally stating an emotion but not clicking on the computer
icon, or responding to distractors (e.g., items in the picture like “his hands are up”). Mind-reading was the primary dependent variable for this study and used to make decisions regarding phase changes.

Data on mind-reading skills for baseline and intervention were recorded using a frequency-based data collection system embedded in the computerized software that indicated the number of correct responses for each assessment session. The software was programmed to randomly select any 10 scenarios from a pool of 22 social problem scenarios at each assessment (see Figure 2). Assessment data were collected through this software program immediately following the instructional session for baseline and intervention.

---

A.BL1.1-28's Results

Question Question 18:
  Answer #1: happy - incorrect : 19 second(s) elapsed

Question Question 22:
  Answer #1: happy - incorrect : 42 second(s) elapsed

Question Question 6:
  Answer #0: happy - correct : 16 second(s) elapsed

Question Question 3:
  Answer #1: sad - incorrect : 16 second(s) elapsed

Question Question 12:
  Answer #1: angry - incorrect : 22 second(s) elapsed

Question Question 5:
  Answer #1: angry - incorrect : 14 second(s) elapsed

Question Question 10:
  Answer #0: happy - correct : 12 second(s) elapsed

Question Question 11:
  Answer #1: angry - incorrect : 14 second(s) elapsed

Question Question 13:
  Answer #1: happy - incorrect : 12 second(s) elapsed

Question Question 19:
  Answer #1: sad - incorrect : 18 second(s) elapsed

---

Figure 2. Sample log of the internal data collection system recorded by the CAI software.
Task engagement. This variable was also measured to determine whether or not task engagement of participants varied across experimental conditions involving different instructional methods to teach mind-reading skills. Task engagement was defined as the participant (a) sitting with his/her shoulders, head, and eyes oriented towards the source of instruction (i.e., person or the computer monitor); (b) having hands on the lap, table top, or the computer mouse as appropriate; and (c) verbally responding or clicking the icon on the computer screen within 10 seconds of presentation of the prompt or question. Non-examples include (a) making verbal sounds not related to the answer; (b) looking away or orienting body away from the source of instruction; (c) displaying stereotyped and repetitive movements with hands or mouse; or (d) failing to respond to a question or prompt within 10 seconds. Task engagement for baseline and CAI was measured using 30-s partial interval recording system (see sample data sheet in Figure 3).
Direct Observation Data Sheet to Record Student Task Engagement

Data Collector: ________________  Participant Code: ________________

Date: _________________________

Instruction: For each of these time intervals, if the target student is engaged in the task, note a “+” and if not engaged, note a “−”

<table>
<thead>
<tr>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (30s)</td>
<td></td>
<td>11.(30s)</td>
<td></td>
<td>21.(30s)</td>
<td></td>
<td>31.(30s)</td>
<td></td>
</tr>
<tr>
<td>2. (60s)</td>
<td></td>
<td>12.(60s)</td>
<td></td>
<td>22.(60s)</td>
<td></td>
<td>32.(60s)</td>
<td></td>
</tr>
<tr>
<td>3. (30s)</td>
<td></td>
<td>13.(30s)</td>
<td></td>
<td>23.(30s)</td>
<td></td>
<td>33.(30s)</td>
<td></td>
</tr>
<tr>
<td>4.(60s)</td>
<td></td>
<td>14.(60s)</td>
<td></td>
<td>24.(60s)</td>
<td></td>
<td>34. (60s)</td>
<td></td>
</tr>
<tr>
<td>5.(30s)</td>
<td></td>
<td>15.(30s)</td>
<td></td>
<td>25.(30s)</td>
<td></td>
<td>35.(30s)</td>
<td></td>
</tr>
<tr>
<td>6.(60s)</td>
<td></td>
<td>16.(60s)</td>
<td></td>
<td>26.(60s)</td>
<td></td>
<td>36.(60s)</td>
<td></td>
</tr>
<tr>
<td>7.(30s)</td>
<td></td>
<td>17.(30s)</td>
<td></td>
<td>27.(30s)</td>
<td></td>
<td>37.(30s)</td>
<td></td>
</tr>
<tr>
<td>8.(60s)</td>
<td></td>
<td>18.(60s)</td>
<td></td>
<td>28(60s)</td>
<td></td>
<td>38.(60s)</td>
<td></td>
</tr>
<tr>
<td>9.(30s)</td>
<td></td>
<td>19.(30s)</td>
<td></td>
<td>29.(30s)</td>
<td></td>
<td>39. (30s)</td>
<td></td>
</tr>
<tr>
<td>10.(60s)</td>
<td></td>
<td>20.(60s)</td>
<td></td>
<td>30.(60s)</td>
<td></td>
<td>40.(60s)</td>
<td></td>
</tr>
<tr>
<td>5 minutes over</td>
<td>10 minutes over</td>
<td>15 minutes over</td>
<td>20 minutes over</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 3.* Direct observation recording form for task engagement during baseline and CAI phase.
A video camera built into the laptop was used to record each session which I reviewed to obtain percentage of task engagement. These data were subsequently reviewed and coded by the secondary observer for reliability.

Direct observation. I served as the primary observer for this study (given that a software program was used to deliver the intervention). The secondary observer was a doctoral candidate who was trained on the data collection procedures and operational definitions of the target behaviors prior to initiating the study. Both observers watched videos of individuals (prior to the onset of the study) using CAI in order to practice coding task engagement. Observers were trained to record the dependent variables using the operational definitions and procedures noted above. Data collection for the study was initiated only after interobserver agreement reached 90% accuracy over three practice sessions.

Interobserver agreement. Interobserver agreement for mind-reading was established by having the secondary observer review the data logs stored in the CAI software program and record the percentage of correct answers for each assessment session per participant (i.e., 100% of the logs were assessed for reliability). These percentages were then compared to the data recorded by the primary observer (Hutcherson, Langone, Ayers, & Clees, 2004).

IOA for task engagement was established by the primary observer reviewing the videos of baseline and intervention sessions for each participant and recording the duration of task engagement onto a 30 second partial interval recording system. The videos were then reviewed by the secondary observer using the same operational definitions and recording procedures. IOA for task engagement was collected for 34% of sessions across baseline and intervention phases.
IOA data for both dependent variables was calculated by dividing the number of rater agreements plus disagreements and multiplying the results by 100 to get a percentage of agreement for each session (for each dependent variable). An agreement for observance of mind-reading skills was defined as both observers documenting the same answer result (i.e., correct or incorrect) based on computer logs. A disagreement for mind-reading was defined as both observers recording a different answer result from the computer logs. An agreement for task engagement was defined as an occurrence when both observers agreed that on task behavior occurred (+) or did not occur (-) during a specified interval of time. A disagreement for task engagement was defined as one observer indicating the occurrence or non-occurrence of a target behavior but the other observer did not record the same event. The outcomes of IOA for mind-reading skills are as follows: Evan (M = 100); Ashley (M = 98.75; range = 90 - 100); Trevor (M = 100); and Pranav (M = 99.7; range = 90-100). The outcomes for IOA of task engagement remained high as well at: Evan (M = 98.96; range = 93.75 - 100), Ashley (M = 97; range = 85 - 100), Trevor (M = 95.83; range = 75 - 100), and Pranav (M = 98.5; range = 90.9 - 100).

IOA for generalization sessions was established by the primary observer reviewing the videos of generalization sessions and recording the percentage of correct responses to the social scenarios that were acted out by family members. These videos were then reviewed by the secondary observer and the percentage of correct responses was recorded. The secondary observer reviewed 40% of generalization sessions for IOA. IOA for generalization sessions was 100% for all participants and all sessions.
At the conclusion of the study, Cohen’s kappa was used to compute overall data reliability for mind-reading and task engagement. Cohen’s kappa is a necessary step in calculating IOA because it adjusts the observed proportion of agreement by taking into account the amount of agreement that can be expected by chance (Cohen, 1960; Horner & Kratochwill, 2012). This was accomplished by counting the number of agreements between the two observers and the number of expected agreements and applying the following formula:

\[ K = \frac{P_o - P_c}{100 - P_c} \]

where \( P_o \) = the proportion of agreements between observers and \( P_c \) = the proportion of agreements expected by chance (Cohen, 1968). The interobserver agreement across baseline and intervention phases for mind-reading skills was .98 indicating strong internal reliability. Additionally, the IOA for task engagement was .93 which also indicates strong internal reliability and meets the guidelines specified by Landis and Koch (1977) (greater than .80).

Social validity. Social validity was assessed using two methods. The first method was a questionnaire which the parents of participants completed. The questionnaire was a self-report measure regarding parental perspective on CAI and included five questions with a 5-point Likert-type scale indicating a range of response from strongly disagree to strongly agree. Questions assessed parental perceptions regarding effectiveness and acceptability of computer assisted instruction.

Results of social validity assessment are presented in Table 2. The questionnaire completed by the mothers of all four participants indicated agreement to strong agreement that the CAI program was effective for teaching their child to predict the emotions of others and their child enjoyed using the computer. Additionally, three parents felt that the CAI added
a critical level of instruction that their child did not obtain from teacher-led instruction.

Perhaps most importantly, all four parents indicated agreement or strong agreement that their child was better able to predict the emotions of others in daily life as a result of participating in the CAI intervention and that they were satisfied with the outcome of the study.

Table 2

Social Validity Results from Parent Questionnaire

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel like the computer assisted instruction (CAI) program was effective for teaching my child to predict the emotions of others.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>My child enjoyed using the computer to learn mind reading skills.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>I feel like computer assisted instruction adds a critical level of instruction that my child does not obtain from teacher led instruction.</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>50%</td>
<td>25%</td>
</tr>
<tr>
<td>My child is better able to predict the emotions of others in daily life as a result of participating in the CAI intervention.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>I am satisfied with the outcome of the study.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>75%</td>
</tr>
</tbody>
</table>

The second social validity measure was collected by interviewing the participants about their perspective on computer assisted instruction. The participant questionnaire used an open
ended response format to assess their enjoyment level, overall satisfaction with CAI and preference of instructional delivery method (see Table 3).

Table 3

*Social Validity Questionnaire to be completed by Participants*

<table>
<thead>
<tr>
<th>Participant Social Validity Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you enjoy using the computer to learn about emotions?</td>
</tr>
<tr>
<td>What did you learn about different emotions?</td>
</tr>
<tr>
<td>Do you think using the computer to learn about emotions will help you recognize other’s emotions in your daily life?</td>
</tr>
<tr>
<td>Do you think the computer was easy to use?</td>
</tr>
<tr>
<td>Would you like to use the computer to learn about other things?</td>
</tr>
</tbody>
</table>

All four participants completed the interview and indicated that they enjoyed using the computer to learn about emotions. Additionally, they all indicated that they thought the CAI program was easy to use and three of the four participants indicated that they would like to use the computer to learn about additional concepts.

**Materials**

Social narratives. During baseline instruction, social narratives involving social situations and related emotions were read to the participant. Social narratives were created using PowerPoint and printed on white paper to present content in a booklet format. A booklet format was used to control for variation in the method of presentation of the content on mind-reading and to present the information in a format commonly used for ToM instruction (i.e.,
paper based social narratives). In other words, the amount and type of content did not change across the experimental phases, just the method of instructional delivery was varied. Each social narrative included pictures and text of social problem situations and was approximately 5 to 10 pages with 1-2 sentences of description per page (see example in Figure 4).

Figure 4. Sample social narrative for baseline procedures.

Original mind-reading curriculum. The materials for the intervention were adapted from the mind-reading curriculum, Teaching Children with Autism to Mind-Read, originally developed by Howlin and her colleagues (1999). This curriculum comprises five levels of
The goal of Level 3 was to teach individuals to predict how a character felt given the emotional content of the scenario displayed in the picture. These emotions are referred to by Howlin et al. (1999) as “situation-based” emotions. Schematic black and white drawings of a variety of situations are included in this level. Although it is recognized that identifying situation-based emotions (i.e., Level 3) is not the most complex level required for successfully demonstrating ToM skills, it is the next level of complexity, which has not been separately evaluated in the CAI research at this point in time, and which is a pre-requisite skill for mastering Levels 4 and 5.

In each of the scenarios in Level 3, the character’s facial features (i.e., eyes, facial expressions, lips) are masked, requiring the participant to determine the emotion based solely
on situational cues rather than facial or auditory cues. Each scenario contains a text description of the scene and has an emotion question, for example: “The big dog is chasing Dan down the road. How will Dan feel when the dog chases him?” Each picture is paired with schematic emotion faces from which to select the correct response: happy, sad, angry, and afraid (frightened).

Computer-based curriculum. The original mind-reading curriculum described above (Howlin et al., 1999 and referred to as the Mind-Reading Curriculum from this point onwards) was slightly modified and converted into an electronic format then saved on a university server to make it accessible from any computer or geographical location. The process of modification and conversion was steered by a collaborative team consisting of me, the dissertation mentor, another doctoral student, and the university’s software programmer and his team. Levels 1 and 2 from the Mind-Reading Curriculum were modified by adding real photographs and clearer line drawings depicting the four emotions (i.e., happy, sad, angry, afraid). These were used for screening the participants and were part of the inclusion criteria.

For the intervention, 22 from the available 48 scenarios from the Level 3 were selected by the collaborative team referenced above based on technical requirements for the software development. A fewer number of scenarios were selected to ensure full attention of participants. The selected 22 scenarios represented each of the four emotions equitably [i.e., happy (5); sad (6); angry (6); afraid or scared (5)].

Content validation of the 22 scenarios incorporated into the software program was conducted prior to the study to determine the extent to which they portrayed the emotion they were intended to represent. This was done because there was no information in the original
curriculum regarding this process. The 22 scenarios were sent electronically to five professionals in the field of ASD including Autism Intervention and/or Behavior Interventionists working in local school districts, one faculty member in the (special education (Autism Intervention), and the Executive Director of the university’s autism center. These individuals reviewed the 22 scenarios used for CAI and rated which emotion was portrayed. The content validation results indicated a mean agreement of 92.5% (range = 77.35 – 100) agreement on the correct response for each scenario based on the original curriculum (i.e., Howlin et al., 1999).

Once the content validation was completed the scenarios were integrated into the computer-assisted instructional (CAI) software in two separate program files. The software utilized for instruction included audio cues and performance feedback for each correct or incorrect response option, whereas the one used for assessment purposes did not provide any audio feedback for the response selected by the participant. The audio clues were created and inserted into the software package by the collaborative team noted above and were not based on any interactive software developed by Baron-Cohen or his colleagues.

Materials for role plays. During the generalization phase of the study, siblings and parents were trained to act out or role play problem situations using a script developed by me based on previous work on teaching perspective-taking skills (Laugeson, Frankel, Gantman, Dillon, & Mogil, 2012) (see Figure 5). Other materials used during this phase were items readily and normally available in a home setting (e.g., board game, craft supplies, food items, etc.).
Experimental Design and Procedures

A multiple baseline design across participants was used to evaluate the effectiveness of computer assisted instruction (CAI) on mind-reading skills and task engagement. The baseline phase was initiated concurrently for all participants but the phase change was staggered for one participant at a time based on their response to intervention. Once a stable rate of

---

**Situation 1**

Setting: Living room

Participants: 1 sibling and 1 parent

Activity: playing on the iPad/computer

Emotion: Angry

Action: The adult will be playing on the iPad or computer. The sibling will come and take the iPad or computer away. Parent will say “Hey!” Sibling will look at adult but not respond or give the iPad/computer back.

Assessment: Interventionist and participant will be sitting in the living room within 10 feet of the action. Interventionist will say “________ just took the iPad from your mom/dad.” “How do you think that made your mom/dad feel?”

---

**Situation 2**

Setting: Kitchen

Participant: 1 sibling and 1 parent

Activity: baking

Emotion: Happy

Action: The parent will be in the kitchen and pull some cookies out of a sack and ask the sibling “Would you like a cookie?” The peer will answer “Yes, please!”

---

*Figure 5.* Social situations script for role play during skill generalization.
responding was documented during baseline across a minimum of five consecutive sessions with no increasing trend, participant one moved from baseline to the intervention (CAI) phase. Decision making for phase change was based on the number of correct responses on mind-reading skills. This section will provide a detailed description of the experimental procedures, background training of me, and procedural fidelity of the intervention. A graphic organizer of experimental procedures is provided in Figure 6.
Baseline
1. Social Narrative (2 per session with feedback)
2. Assessment (10 questions on mind reading software; no feedback)

Intervention: CAI
1. CAI Instruction (22 scenarios per session with feedback)
2. Assessment (10 questions on mind reading software; no feedback)

Generalization Training
Social Scenarios (8 per session with feedback; same scenarios used during probes)

Generalization
Social Role Play with Parent and Sibling (1 per session) (different from scenarios used for probes)

Generalization Probes
Novel Social Scenarios (4 per session – 1 for each emotion; no feedback)

M > 50% on probes move to generalization
3 scenarios at 100% move to generalization
M < 50% on probes move to generalization training

Figure 6. Graphic organizer of research procedures.
Baseline. Social narratives and paper-based products are common for teaching perspective-taking and mind-reading skills. Thus, these were used to collect data for baseline, the control condition that was compared with CAI (Gast, 2010). Baseline involved instructional sessions where I read a social narrative to each participant. The narrative involved a social situation requiring the participant to identify an emotion felt by one of the characters. Care was taken to ensure that the duration of each BL session was the same as session length as the intervention session (approximately 7 minutes) to control for threats to internal validity. At the end of each baseline session, participants were asked to take the computerized assessment which included 10 randomized scenarios on mind-reading skills with no feedback as discussed previously.

Computer assisted instruction (CAI). The intervention condition involved the participant sitting at a desk or table with a laptop computer. I started the computer program using the participant’s specified identification number. For each session participants first viewed 22 instructional trials which included embedded audio prompts (e.g., “look at the picture”), clues (e.g., “could the girl be upset?”), feedback for correct and incorrect responses (i.e., green check mark or red “x” respectively), and embedded reinforcement for correct responses (i.e., verbal praise, “Good Job!”). The 22 scenarios were displayed in systematic succession with scenarios for each emotion being displayed in groups (e.g., all scared scenarios were displayed in a row, then all happy scenarios were displayed, etc.). Once participants completed the instructional session, they were asked to take the computerized assessment on mind-reading skills which included 10 randomized questions with no feedback or prompts (identical to the baseline measurement).
Generalization probes. Generalization probes were conducted during baseline and intervention phases for all participants for two reasons: first, to assess the effectiveness of social narratives and CAI on mind-reading skills of participants based on the respective instructional methods; and second, to determine if participants would need to receive additional generalization training prior to beginning the generalization phase.

In order to complete generalization probes, I developed 16 social scenarios based on common social situations identified by parents’ of participants. Each scenario described a different social situation (e.g., “Your mom dropped her phone and it broke. How does that make your mom feel?”). For each generalization probe, the participant was required to listen to four scenarios and then determine how the character would feel in each situation (i.e., happy, sad, angry, scared). Participants were given 100% if they answered all four scenarios correctly, 75% for answering three of the four questions correctly, 50% for answering two correctly, etc. No feedback or reinforcement was provided for correct or incorrect answers. If participants had a mean performance of 50% or higher of answers correct on generalization probes during CAI, they did not receive any additional training prior to beginning the generalization phase. If a participant did not demonstrate a high level of mastery on generalization probes (M < 50%), additional training was provided prior to the onset of the generalization assessment phase.

Generalization training. Generalization training was needed only for one of the four participants (Pranav). Training involved me asking Pranav 16 social scenario questions (the same used during generalization probes; four scenarios for each emotion) and then providing feedback and an explanation for the correct answer (e.g., “your mom feels sad when she rips
her paper,” “your sister is happy when she gets to play outside with her friends”). This was continued until the participant reached 100% accuracy for three consecutive scenarios. Pranav reached 100% accuracy for three consecutive scenarios after two sessions of training and then began the generalization phase.

Generalization assessment. Generalization of mind-reading skills was conducted to evaluate the extent to which CAI was effective in teaching social problem-solving skills to participants as applied to daily life situations. The family members of each participant identified various social situations that occurred in a variety of settings (e.g., living room, kitchen, and dining room). I developed a short script (i.e., 3-4 sentences) for each of these situations (different from the scenarios used during generalization probes and training) and provided them to the parents and sibling of each of the participants with an explanation of the generalization procedure. They were subsequently trained to act out or role-play their parts (see Figure 4). Additionally, they were instructed on how to provide feedback to participants when they correctly (e.g., verbal praise) or incorrectly (e.g., error correction) identified an emotion based on the social situation displayed. The parents and sibling practiced acting out the social situation until they successfully completed all steps of the various social situation scripts with 90% accuracy (I recorded with a notation of “yes” or “no” for each role-play action).

During generalization, when the actors were ready to role-play, the participant was invited to join the family in the setting (i.e., living room, kitchen, or dining room). I ensured that the participant was in close proximity to the scene of the action (approximately 10 feet). For an example of one social situation used during generalization, the parent was engaged with an electronic device in the living room when the sibling got up and grabbed the electronic device
from the parent and started to use it herself. I asked the participant to describe how the parent felt in this situation and explain the reason for the answer. I recorded data on the accuracy of the participant’s response and assessed for reliability by the secondary observer after reviewing the same video recording.

Procedural fidelity. Procedural fidelity for baseline and the intervention (CAI) phase was conducted to ensure that the procedures were consistently implemented with accuracy and integrity (Gast, 2010). Because the intervention was delivered through a computerized software program, there was a low probability of error given that it was pre-programmed to deliver the same type of stimuli at each trial even if the order of presentation of stimulus items was randomized. Thus, procedural fidelity was assessed to ensure accuracy in using the correct log for each participant upon login, recording the session on the laptop, and saving the data on the software program when the session was completed. I used a checklist during each session to assess procedural fidelity during both baseline and intervention phases (see Table 4 & 5) which was then calculated by dividing the number of items on the checklist I correctly completed by the total number of items on the checklist and multiplying by 100. The outcome showed 100% accuracy for each participant on all the steps of the baseline and intervention. IOA on procedural fidelity was assessed by the parents of participants using the same checklist for an average of 36% of baseline and intervention sessions (Evan = 39%; Ashley = 37.5%; Trevor = 33%; Pranav = 37%). The IOA on procedural fidelity based on the checklist record of myself and the parent showed 100% agreement on the steps completed accurately.
Table 4

*Procedural Fidelity Checklist for Baseline*

<table>
<thead>
<tr>
<th>Date/Session #</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on laptop camera</td>
<td></td>
</tr>
<tr>
<td>State initial, BL or CAI, date</td>
<td></td>
</tr>
<tr>
<td>Show student task list: First, Next, Then</td>
<td></td>
</tr>
<tr>
<td>Provide two choices for social story</td>
<td></td>
</tr>
<tr>
<td>Student sits at computer desk facing researcher</td>
<td></td>
</tr>
<tr>
<td>Read 2 social stories</td>
<td></td>
</tr>
<tr>
<td>Student turns and faces laptop</td>
<td></td>
</tr>
<tr>
<td>Click on shortcut link labeled Mind Reading Software</td>
<td></td>
</tr>
<tr>
<td>Enter EUID and Password</td>
<td></td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled Assessment</td>
<td></td>
</tr>
<tr>
<td>Enter Participant ID code</td>
<td></td>
</tr>
<tr>
<td>Monitor participant (must complete 10 scenarios)</td>
<td></td>
</tr>
<tr>
<td>Log out</td>
<td></td>
</tr>
<tr>
<td>Allow 5 minute break-free choice</td>
<td></td>
</tr>
</tbody>
</table>
Table 5

_Procedural Fidelity for CAI_

<table>
<thead>
<tr>
<th>Date/Session #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on laptop camera</td>
</tr>
<tr>
<td>State initial, BL or CAI, date</td>
</tr>
<tr>
<td>Show student task list: First, Next, Then</td>
</tr>
<tr>
<td>Student sits at computer desk facing laptop</td>
</tr>
<tr>
<td>Click on shortcut link labeled Mind Reading Software</td>
</tr>
<tr>
<td>Enter EUID and Password</td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled Instruction</td>
</tr>
<tr>
<td>Enter Participant ID code</td>
</tr>
<tr>
<td>Monitor participant (must complete 22 scenarios)</td>
</tr>
<tr>
<td>Log out</td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled Assessment</td>
</tr>
<tr>
<td>Enter Participant ID code</td>
</tr>
<tr>
<td>Monitor participant (must complete 10 scenarios)</td>
</tr>
<tr>
<td>Log out</td>
</tr>
<tr>
<td>Allow 5 minute break-free choice</td>
</tr>
</tbody>
</table>
Investigator

At the time of the study, I was a fourth year doctoral candidate in special education with a focus on autism research. I completed all the required course work for the Board Certified Behavior Analyst (BCBA®) program along with 72 doctoral credit hours including various research methods courses. Additionally, I had six years of experience teaching students with autism in the public school setting.

Results

Visual analysis, commonly used in single subject research, was the primary method for analyzing data to determine a functional relation between the independent and dependent variables. Specifically, data were concurrently coded and graphed to analyze the stability, level and trend changes, variability, overlapping data across adjacent phases, and similarity of pattern across similar phases (Horner & Kratochwill, 2012). Additionally, effect size was computed using Cohen’s $d$ (1988) in order to determine the magnitude of effect of the intervention on the dependent variables. Results are described in relation to the specific research questions. Additionally, line graphs of results are provided in Figures 7 and 8.

Effectiveness of Computer Assisted Instruction on Mind-Reading Skills

The first research question assessed the existence of a functional relation between CAI and mind-reading skills. Baseline data for Evan indicated a mean of 31.66% correct answers on mind-reading skills, however, a decreasing trend was observed as the phase progressed. Given the stability in the data pattern, intervention was initiated after six consecutive sessions. Data showed an immediate effect, from an average of 13% in the last three sessions to an average of 73% for the first three sessions of intervention. Evan’s mean level of performance was noted at
82.86% during the intervention marked by ceiling performance (i.e., 100%) on the seventh session. The stability envelope indicated that 86% of the intervention data were within 20% of the median indicating stability in the pattern of learning during intervention. The percentage of non-overlapping data was acceptable at 71.4%. Overall, Evan’s data indicate an increase in mind-reading skills as a function of the intervention.

Following an increasing trend in response to intervention for Evan and a concurrent stable baseline for Ashley, intervention was initiated with her. Ashley’s baseline data show a mean level of 26% correct responses. The stability envelope indicated that 90% of data points were within 20% of the median. She also demonstrated an immediate effect in response to the intervention, moving from an average of 27% during the last 3 sessions of baseline to 83% average for the first 3 data points during intervention. The mean level change during CAI was at 90% with 83% of data within the stability envelope. During the intervention phase she demonstrated an accelerating trend and reached the ceiling of 100% correct answers after 4 sessions. There was no overlap in data across adjacent phases. Overall, data for Ashley indicate a clear functional relationship between CAI and mind-reading skills.

Once Ashley reached a stable level of performance during intervention and Trevor demonstrated a stable-flat baseline with 93.75% of data falling within the stability envelope, intervention was initiated with Trevor. His rate of responding increased from a mean baseline level of 26.25% to 84.55% during intervention. He too demonstrated an immediate response to intervention with an absolute change from 23% to 76.66%. He reached the ceiling level of 100% correct answers on the fourth session, however, his data did show some variability across the 11 intervention sessions with only 63.64% of data falling within the stability envelope. This
was likely caused by a decrease in performance on the sixth session due to illness. The percentage of non-overlapping data was significant at 100%. Additionally, given the increasing trend in the data pattern starting the seventh session, overall results indicate that the CAI had a positive, functional impact on Trevor’s mind-reading skills.

When Trevor’s performance returned to the ceiling level at 100% (on intervention session 10) while Pranav’s data demonstrated a stable pattern of responding during baseline (i.e., stability envelope = 83.3%), intervention was started with him. His data appear to show a variable pattern, however, other than two sessions of data overlap with baseline (possibly caused by the initial slow response to intervention in the first session and difficulty maintaining attention in the seventh session), Pranav’s mean score increased significantly from 27.91% during baseline to 80% during intervention. Additionally, the absolute change during intervention was 40% to 100% indicating an increasing trend. The percentage of non-overlapping data was also within the moderate range at 85.71%. These results appear to indicate a functional relation between CAI and mind-reading skills for Pranav as well.
Figure 7. Percent correct on mind-reading skills for all participants.
Effect of CAI on Rate of Task Engagement

Although task engagement was a secondary dependent variable, the second research question assessed the extent to which it was higher for participants during baseline (i.e., paper-based social narratives) vs. CAI. The results for task engagement showed higher variability across participants than what was observed for mind-reading skills (i.e., the primary dependent variable). Three of four participants showed higher levels of task engagement when CAI was used in comparison to baseline even though data showed overlap across adjacent phases for one of these participants (i.e., Trevor). Evan displayed a variable rate of task engagement during both baseline ($M = 13$; range $= 0 – 29$) and intervention ($M = 50$; range $= 30 – 68$), however, the mean level was higher during intervention with 86% of the data falling within the stability envelope. He displayed a relative change of 68 – 54 indicating a decelerating trend yet 100% non-overlapping data indicating that CAI may have had a positive impact on task engagement compared to baseline.

Ashley did not display a significant change in task engagement between the two experimental conditions based on visual analysis of the data. She showed a mean of 56.9% during baseline which increased slightly to 73.8% during intervention. Her data were variable during both conditions and 87.5% of the data overlapped with CAI demonstrating that task engagement was not impacted by the type of instructional method used with Ashley.

Trevor displayed a very stable (93.75% within the stability envelope) mean level of 2.4% task engagement during baseline. He increased to a mean of 30% during CAI however, data demonstrated a large amount of variability with only 27% of the data falling within the stability
envelope. He also demonstrated a decelerating trend for task engagement with a relative change of 50 - 18 during CAI. In spite of the variability during intervention, visual analysis indicates Trevor’s task engagement increased during CAI as compared to baseline.

Pranav displayed a stable level of task engagement during baseline ($M = 6$; range 0 – 33) with 83% of the data within the stability envelope. The level of task engagement increased during CAI ($M = 57.5$; range = 25 - 84) even though only 36% of the data were within the stability envelope. Data show multiple trend lines with a relative change from 65 – 25. Overall Pranav’s data indicate a mean increase of 51.5% during CAI and a percentage of non-overlapping data of 91.4%.
Figure 8. Rate of task engagement for all participants.
Effect of CAI on Response Generalization

As noted previously, the purpose of generalization assessment was to determine the extent to which mind-readings skills learned through computer-assisted instruction would transfer to daily life social situations that required problem-solving (i.e., Research Question 3). The generalization phase was initiated in a staggered manner for participants after each one showed a steadily increasing trend and an average of 80% or higher for CAI with at least one session at criterion (100%). Each participant was exposed to at least one scenario per emotion for a total of four role-play situations per session. Each social situation required the participant to label the emotion experienced by a family member after the role-play ended.

During generalization, all four participants demonstrated mastery (mean 93% or higher) on mind-reading skills when presented with social scenarios acted out by family members in typical home settings. Evan immediately showed 100% correct responses and maintained this level of mastery for all nine generalization sessions ($M = 100$). Ashley started generalization at 75% correct answers, missing only the scenario indicating a scared emotion. She reached 100% mastery in the second session and maintained this for six of the eight generalization sessions ($M = 93.75; \text{range} = 75 - 100$). Trevor immediately reached 100% mastery during generalization and maintained this level for eight of the 10 generalization sessions ($M = 93.75; \text{range} = 75 - 100$). Generalization assessment for Pranav was initiated after implementing 2 generalization training sessions as noted previously. He immediately responded to the generalization assessment with 100% mastery and maintained this level for nine of the 10 generalization sessions ($M = 97.5; \text{range} = 75 - 100$).
Effect Size for Mind-Reading and Task Engagement

Effect size is now considered a critical component of single-subject research for determining the magnitude of effect of intervention (Grissom & Kim, 2005). Effect size was calculated for both dependent variables (i.e., mind-reading skills and task engagement) for all participants across the baseline and intervention phases using Cohen’s $d$ (1988) statistic. The following formula was used to calculate Cohen’s $d$ for baseline and intervention conditions:

$$D = \frac{(M_I - M_B)}{(SD_p/ \sqrt{2(1-r)})},$$

where $M_I$ represents the mean score for Intervention, $M_B$ represents the mean score for baseline, and $SD_p$ is the pooled standard deviation for both experimental phases, and $r$ is the correlation between the baseline and intervention data (Dunst, Hamby, & Trivetter, 2004). This formula for effect size is recommended when comparing correlations between phases for single subject research where the number of data points across adjacent phases are unequal (Dunst et al., 2004). The effect size for all four participants was large for mind-reading skills [Evan ($d = 3.06$); Ashley ($d = 7.49$); Trevor ($d = 4.7$); Pranav ($d = 3.33$)] and the overall outcome of the intervention demonstrated a large effect for mind-reading skills ($d = 4.15$) which is not surprising for single subject research. Additionally, the effect size was large for all participants for task engagement as well [Evan ($d = 3.49$); Ashley ($d = 1.17$); Trevor ($d = 1.84$); Pranav ($d = 3.43$)]. The overall outcome of the intervention demonstrated a large effect for task engagement ($d = 1.59$).

Discussion

Results of this study indicate a functional relation between computer-assisted instruction and mind-reading skills of children with autism (Research Question 1). Additionally,
task engagement was noted to be higher for three of the four participants when CAI was the instructional delivery method whereas for one participant the method of instructional delivery did not appear to matter (Research Question 2). Finally, results demonstrated a functional relation between computer-assisted instruction and generalization of mind-reading skills to untrained social situations in natural settings (Research Question 3). This section will discuss the findings of the study, implications for practice, limitations, and directions for future research.

Overcoming Mind-Blindness

Mind-blindness refers to the absence of theory of mind (ToM) skills or an individual’s social understanding and ability to empathize, predict responses, and infer the emotional state of others typically based on facial expressions, body language, and situational cues. Individuals with ASD tend to display severe deficits in mind-reading skills (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997) which can be improved only with the use of explicit instructional methods (Feng, Lo, Tsai, & Carledge, 2008; Hutchins & Prelock, 2008). Previous research notes the challenges associated with teaching generalized emotion recognition and perspective taking to individuals with ASD (Baron-Cohen 1989; Ryan & Charragain, 2010; Tager-Flusberg, 2007).

Although a variety of interventions have been explored to alleviate ToM deficits such as social skills group training, social narratives, and CAI, existing research consistently documents a large amount of variability in the individual response to these interventions and a lack of procedural fidelity in implementation. To this point, group design research is the only published methodology used to observe the effect of these interventions which does not allow
for documentation of the effect for individual participants because average scores are reported with limited information on the point in time when intervention produced a change in behavior. Additionally, groups design studies rarely report generalization outcomes for participants. The current study adds to the body of research by utilizing single subject research design and procedures which allow for evaluation of individual participant responses by displaying visual data to demonstrate the behavioral pattern as a function of intervention.

In one example of a group design study, Silver and Oakes (2001) evaluated the use of a computer program, the Emotion Trainer, to teach a variety of ToM skills to children with autism using a randomized control trial design. The Emotion Trainer used in their study taught situation based emotions, belief-based emotions, and desire-based emotion through computer based scenarios similar to the software program used in the current study. They used pre- and post-intervention measurement to compare the values in order to determine if the program made a significant impact on ToM skills. Their findings revealed that all participants improved in all the post intervention measures including Happe’s Strange Stories and emotion recognition cartoons regardless of verbal ability or age. However, the authors noted that there was a large amount of variability in the amount of errors individual participants made on computer tasks and in the amount of time each participant accessed the computer program.

In the current study, similar to Silver and Oakes (2001), all participants regardless of age or verbal ability significantly improved in their ability to predict situation based emotions with the CAI software. However there was individual variation in the rate in which the participants responded. For example, three of the four participants responded to the treatment condition immediately. The fourth participant (i.e., Pranav) did not respond to the intervention
immediately but did show a significant increase in performance by the second intervention session. There was also individual variation in the point at which the participants reached the ceiling performance (100%). Trevor reached ceiling performance at the fourth session, Ashely reached it at the fifth session, Evan reached it at the seventh sessions, and Pranav reached 100% during his tenth session. This individual variation does not appear to be related to level of task engagement or overall cognitive skills but instead highlights the individual acquisition rate for learning new material.

Methodological limitations were reported by LaCava et al. (2007) and Silver and Oakes (2001) in regards to the variation in which the intervention was implemented. In both studies, participants were able to access the software at variable amounts of time which may have impacted their overall progress. These methodological issues were resolved in the current study by employing a multiple baseline design in which participants accessed the software for the same amount of time during each session (i.e., controlled for intervention density per session). The total number of sessions delivered to each participant depended on their individual, documented response to the treatment as per the research design, suggesting that intervention delivery was based on a systematic and consistent schedule of performance assessment.

Another advancement to the field of research was the use of more complex material in the current study. Past research has explored the use of CAI to teach emotion recognition skills at a basic level (Feng et al., 2008; Hopkins et al., 2011; LaCava, 2007). The current research utilized CAI to successfully teach a more complex level of perspective taking or mind-reading skills which is the next sequential level in ToM skills (Howlin et al., 1999).
Additionally, previous studies have failed to assess generalization of skills or have documented a diminished ability for individuals with autism to generalize the complex social skills involved in theory of mind into untrained and natural environments. Thus, it has been difficult to assess whether improvements through computerized instruction actually improved participants’ ability to predict situation based emotions in real life or if the improvements were simply a product of memorization and not a result of increased social skills. The current research expands upon this by assessing generalization of skills. Because we cannot observe internal events, it was necessary to assess overt behaviors to determine if the improvements through CAI transferred to real life situations. Generalization is a critical component of applied research in order to establish if an intervention has the power to produce quality of life outcomes.

Addressing Issues with Response Generalization

Because social deficits are a primary characteristic of autism, it was critical that participants generalize their mind-reading skills into natural settings. The current study expands previous research by Begeer et al. (2011) where they evaluated the use of a 16 week social skills group training program to improve ToM skills in 40 children with autism. They conducted a pre- and post-intervention assessment to determine any significant changes in ToM skills as a result of training. They also obtained data from a self-reported empathy questionnaire and a parent reported social skills questionnaire to assess perspectives on skill generalization. Data indicated that participants improved in first order and false-belief reasoning, mixed emotions, and complex emotions. They did not demonstrate improvement in perception, imitation, emotion recognition, emotional awareness, and self-reported empathy.
or parent reported social skills. These results suggest that the treatment did impact some conceptual abilities but did not transfer to daily life problem-solving skills.

This was also noted in research by Ryan and Charragain (2010) who implemented a systematic emotion recognition training program to enhance emotion recognition from photographs and knowledge of emotion-related vocabulary to 20 children with ASD. They found that even though emotion recognition skills improved, this advancement did not generalize to an improved ability to identify emotions or react appropriately to others emotions in real situations.

In the current study, all four participants were able to transfer learned skills by identifying emotions in untrained social scenarios at a high level of mastery (i.e., higher than 75% accuracy for all four emotions). A possible explanation for the improved generalization of skills in this research compared to previous research was the use of generalization probes throughout the study. Generalization probes were conducted throughout the study to track participants’ level of generalization and guide decisions about initiating the generalization phase of the study. Two of the four participants averaged above 75% on generalization probes during CAI. Trevor averaged 54% however he demonstrated a steady increase in performance on the last three probes. Pranav, the fourth participant, did not demonstrate high levels of mastery on generalization probes (46%) during CAI even though his mind-reading skills showed improvement. Demonstration of variability on generalization probes indicated the need for separate generalization training at the end of the CAI phase and prior to conducting generalization assessment. The two training sessions provided explicit instruction on the emotion attributed to each social scenario drawn from the generalization probes and pointed
out critical features of the scenarios (e.g., “sister feels scared when she sees a spider,” “mom is angry when sister takes her iPad away”).

A possible explanation for Pranav’s lower performance on generalization probes might be his age or his verbal skills. He was the youngest of the participants and demonstrated the least expansive vocabulary of all participants, both expressively and receptively. However, when he was provided with generalization training, he was able to identify all four emotions acted out by his family members with 100% accuracy during the first four session of the generalization assessment phase. It was critical to ensure that all participants generalized these mind-reading skills to assure social viability of CAI as an intervention for improving social skills.

The generalization findings add a critical component to the body of research in the area of ToM interventions. Begeer et al. (2011) noted in their findings that the failure to generalize skills may be an intrinsic part of autism and that treatment programs may need to explicitly train for generalization of behavioral skills rather than conceptual understanding. Froehlich and colleagues (2010) also noted that the individuals with autism who participated in their study had a difficult time generalizing dot formation patterns as the patterns become more novel and less like the original stimuli. In the current study, five to six scenarios for each emotion were displayed in groups (e.g., all happy scenarios were displayed in a row, all angry scenarios were displayed in a row, etc.), for a total of 22 scenarios during instructional sessions. This allowed participants to make connections and most likely form a conceptual understanding about the characteristics in each scenario that relate to each type of emotion. This appears to suggest that by providing multiple exemplars of social scenarios in a systematic and predictable format through CAI, participants were able to improve their conceptual understanding of the emotions
of others, which then allowed them to generalize the skills to show behavior change in real life scenarios even though they were simulated.

Another possible reason for the high rate of success in teaching these complex ToM skills along with the use of multiple exemplars was performance feedback during generalization and programming common stimuli (e.g., similar visual stimuli and type of feedback) (Stokes & Baer, 1977). The generalization assessment utilized social scenarios of common family situations. The participants’ familiarity with the social situation along with their newly learned skills provided through CAI seemed to promote accurate identification of the emotions during role-plays conducted by the family. This is supported by the notion that generalization must programmed as an integral component of an intervention (Stoke and Baer, 1977). Another benefit to the instructional delivery method was the interactive manner in which the mind-reading skills were taught.

Impact of Instructional Delivery Method on Task Engagement

Task engagement tends to improve the rate of learning and reduce inappropriate behavior. The current study demonstrated that the CAI software can not only improve ToM skills in children with autism, but also have a positive impact on the amount of task engagement for some participants. This is supported by the findings of Carnahan, Musti-Rao, and Bailey (2009) who demonstrated an increase in task engagement when the instructional delivery method was interactive. CAI was shown to be a successful tool in this study because it provided a high level of systematic auditory and visual input, which increased task engagement and therefore, likely increased rate of social skill acquisition. This is also supported by the findings of Kaey-Bright and Howarth (2012) who indicated that when individuals are highly
engaged in a software program, the levels of competing behaviors such as echolalic speech, self-stimulatory hand motions, and elopement are reduced. Although the current study did not evaluate levels of problem behavior like those recorded in the study by Kaey-Bright and Howarth (2012), it did evaluate the level of task engagement during CAI. As task engagement increased through CAI, competing behaviors such as looking away or making unrelated comments or verbalizations decreased, which allowed for more focus on instructional material.

All participants in the current study demonstrated a significantly large effect size for task engagement during CAI. For Ashley however, though the effect size of CAI on task engagement indicated a large magnitude of effect at 1.17, visual analysis indicated a high level of overlapping data between baseline and intervention phase, which suggests there was not a clinically significant impact of instructional method on level of task engagement. This may be attributed to her high level of task engagement during the baseline condition compared to the baseline rates of the other three participants. Overall Ashley was a very engaged learner regardless of the instructional method and seemed to enjoy the social narratives and enjoyed reading along with me.

Additionally, while Trevor had a significantly large effect size of 1.84 for task engagement, the visual analysis of the data did not indicate a clear increase in task engagement during the CAI condition due to the variability in data during intervention. This may be due to a declining trend in task engagement during the treatment condition. He seemed to lose interest in the computer program once he reached 100% mastery and also had many occasions where he turned away from the computer screen to receive feedback or approval from me while taking the CAI assessment.
Also, though all participants demonstrated a large effect size for increased task engagement during CAI, visual analysis revealed that all participants had a decreasing trend in task engagement during the treatment condition. This may indicate that, when the novelty of the software program was worn-out, participants were less likely to stay engaged in the task at all times. Future research should continue to expand upon the elements of a computer program that are responsible for maintaining interest and novelty such as sound, animation, and types of programmed reinforcement (Lahm, 1996). These results indicate that, while CAI may lead to increased levels of task engagement for some, we should not make the assumption that instructional technology produces the same effect on all children, given the variability in individual response rates to CAI as it relates to task engagement.

Implications for Practice

The mind-reading software is a unique instructional tool which presents material at an individual pace and provides built in systemization, predictability, and immediate reinforcement for correct answers or error correction for incorrect answers. The high fidelity with which the program presents the instructional material adds to its overall success. The program was also easy to navigate allowing children to access the program successfully regardless of their age. Additionally, this study indicated the flexibility of delivering CAI in that it was completed in the individual participants’ homes. Because the instruction was provided by programmed software, parents’ expertise and training was not a factor in the success of participants. This adds to the usefulness in teaching difficult skills which typically require a high level of training, background knowledge, and fidelity of implementation.
Another valuable factor in the implementation of the current study was the participation of parent and sibling during planning and implementation of generalization sessions. Providing parents with social scenario scripts allowed them to see the structure and language that could be used for everyday scenarios they may encounter with their child. Additionally, because the parents were able to see what their child was learning and the progress s/he was making, they would be more likely to continue implementing the program and practicing social scenarios once the study was completed. These implications are also supported by the social validity data. Anecdotally two parents reported that their child began using “emotion language” (e.g., he feels happy, that makes me sad) outside of the study in unprompted scenarios. For example, Ashley’s mother reported that while playing at the playground another child began crying and Ashley responded without prompting, “She feels sad, she is crying.” This was an illustration of social awareness she had not demonstrated prior to the study.

Indicators for High Quality Research

This study expands and improves upon previous research by including necessary elements of high quality research described by Reichow et al. (2011). The specific primary quality indicators for single subject experimental design which were included in this study are: detailed participant characteristics, independent and dependent variable descriptions provided with replicable precision, operational definitions for baseline with more than three data points demonstrating a stable trend, and visual analysis of data demonstrating a large shift in trend and/or level between phases. Additionally, the study contained three demonstrations of experimental effect resulting from manipulation of the independent variable.
The study also included many of the secondary quality indicators for single subject experimental design including: inter-observer agreement above .80, Cohen’s kappa computed for over 20% of sessions with a score greater than .60, procedural fidelity above 90% for all sessions, and assessment of generalization of skills. The high level of methodological rigor that was used in this research study allows us to make conclusions about the results with more confidence.

Limitations and Directions for Future Research

Although this study meets criteria for the number of participants necessary to demonstrate experimental effect in a multiple baseline design (Reichow et al., 2011) the small sample size does lead to a potential limitation to external validity (Gast, 2012). Additionally, this study evaluated the effects of CAI to teach mind-reading skills from Level 3 of the Howlin et al. (1999) curriculum (i.e., predicting emotions from social situation); however, theory of mind extends much further beyond this skill set. Future research should evaluate CAI for teaching more complex ToM skills such as predicting belief-based emotions and desire-based emotions, which are the focus of instruction in Levels 4 and 5 of the Howlin et al. (1999) Mind-Reading Curriculum to increase social reciprocity. These were not evaluated in the current study because the participants demonstrated a limited ability to attend to lengthy instruction therefore it was necessary to keep the instructional sessions limited to 5-10 minutes. Future research should evaluate the use of CAI to teach these more complex emotions and assess if participants are equally capable of generalizing the more complex skills to untrained social scenarios.
While the scope of the current study was to evaluate participants’ ability to correctly identify emotions in real life social scenarios, future research should also explore the use of CAI for training individuals to respond appropriately in social scenarios once the emotion has been accurately identified. Responding to emotions correctly in social situations is a more complex theory of mind skill and may result in increased participation of individuals with autism in their social environment in a more successful and meaningful way.

Finally, this study demonstrated the flexibility of CAI by implementing the intervention in the home setting but future research should explore the use of CAI to teach additional social skills in the school, community, and other settings in order to further enhance social interactions and success across a variety of settings.

In summary, the current research demonstrates that by implementing the critical components of a ToM intervention suggested by Howlin et al. (1999) such as: (a) breaking information into small steps, (b) providing instruction in a normal developmental sequence, (c) systematically reinforcing behavior, and (d) providing errorless learning through CAI, children with autism were successfully able to master mind-reading skills and then generalize these skills to untrained settings.

References

doi:10.1177/105345120103600502


72


THE USE OF COMPUTER ASSISTED INSTRUCTION FOR INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: A SYSTEMATIC LITERATURE REVIEW

The Diagnostic and Statistical Manual of Mental Disorders (DSM 5) defines autism spectrum disorder as impairment in social communication and interaction across multiple contexts, failure to initiate and respond to social interactions, as well as restricted and repetitive patterns of behavior. Other characteristics associated with ASD are deficits in social-emotional reciprocity and deficits in understanding body language (APA, 2013). These deficits are often exhibited as an inability to understand and process subtle social and emotional cues expressed through facial expression, voice intonation, and context cues (Tager-Flusberg, 2007). These characteristics can contribute to the display of overt and covert forms of problem behavior such as tantrums and/or social withdrawal, which limit an individual’s ability to successfully participate in their social environment (Bauminger, 2002). One specific characteristic of individuals with ASD is a deficit in theory of mind (ToM) skills or what is referred to as mind-blindness (Colle, Baron-Cohen, & Hill, 2007).

Introduction

Theory of mind (ToM) skills, also known as mentalizing, mind reading, or social intelligence, is defined as an ability to infer other people’s mental states such as thoughts, beliefs, and intentions and use the information to interpret and predict what others will say and do in specific social situations (Hutchins, Bonazinga, Prelock, & Taylor, 2008). ToM plays a critical role in interpreting communicative intent to see beyond spoken words and infer or hypothesize about a speaker’s mental state. Additionally, it allows individuals to self-reflect, realize fallibility in personal beliefs, and use social communicative skills of persuasion to repair
breaks in conversation (Howlin, Baron-Cohen & Hadwin, 1999). Inaccurate assessment of the mental state of others can create social challenges such as inability to interpret sarcasm or figurative language, recognize others intent to deceive, and maintain reciprocal communication (LaCava, Golan, Baron-Cohen, & Myles, 2007). Individuals with ASD usually show significant delays in reaching ToM milestones when compared to individuals without disabilities. To demonstrate this delay, Baron-Cohen (1989) developed a ToM task in which individuals with ASD were asked to evaluate pictures of either an entire face or the eye region and determine the correct facial emotion. Results indicated that individuals with ASD performed significantly lower than age and IQ matched peers at identifying emotions on both types of pictures. Similarly, the individuals with ASD also performed markedly lower on the eyes-only condition than their age-matched peers in a follow up study by Baron-Cohen and colleagues (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). These results support the notion that emotion recognition is a key to understanding the nature and severity of social deficits in individuals with ASD (Baron-Cohen, Leslie, & Frith, 1985; Howlin et al., 1999).

The ability to recognize emotions is just one component of the complex processes involved in ToM skills (Tager-Flusberg, 2007). There is growing evidence to suggest that ToM can be learned by individuals with ASD through explicit teaching strategies with a specific focus on generalization training (Feng, Lo, Tsai, & Carledge, 2008; Hutchins & Prelock, 2008; Tager-Flusberg, 2007). A variety of interventions have been utilized in an attempt to improve these skills, with mixed results.
Ryan and Charragain (2010) conducted a study with 30 participants (mean age 9.5 years; 20 in the experimental and 10 in the control group) to evaluate the impact of a systematic emotion recognition training program on emotion recognition and emotion-related vocabulary comprehension tests. Pre- and post-test measures showed that children in the experimental group achieved higher levels of improvement when compared to their peers in the wait-list control group. The latter group showed similar gains after being exposed to the 4-week training program. While the results were promising and supported the role of explicit instruction of emotion recognition, learned behaviors of participants did not generalize to real social situations. Parents of the participants reported that although children attended to faces and accurately identified emotions, they commented and inappropriately laughed after recognizing an angry face. These findings suggest that it may be possible that the correct responses were memorized by participants without cognitively internalizing the concepts.

Similar issues related to lack of transference to daily life situations, presumably because of social cognition deficits, were also displayed in a study by Begeer et al. (2011). They conducted a randomized control trial with 40 participants to evaluate the effectiveness of a 16-week ToM training package with 8-13 year olds with ASD. The intervention package included strategies to explicitly teach ToM skills during 1.5 hours per week in small group sessions. Results indicated that while participants improved their elementary understanding of ToM, they did not show any improvement in more complex skills such as empathy. Additionally, they were not able to transfer learned skills into daily life situations. It is possible that the low density of the intervention or absence of specific generalization training may have resulted in limited effect on participants.
Moving beyond basic emotion recognition to a complex understanding of emotions and empathy is a recurrent challenge faced with ToM interventions. In a study by Feng et al. (2008), ToM training was provided through Macromedia Flash Player to teach a 11-years old male participant with high-functioning autism to identify emotions in social scenarios. Results indicated that he was able to identify emotions in basic social scenarios (e.g., desire-based emotions and belief-based emotions) but not the more complex skills (e.g., first-order false belief).

The results of these studies further confirm that ToM is a complex construct which requires a comprehensive training package designed to develop advanced skills beyond rote identification of emotions. These skills must then be generalized to social situations in order to have a greater possibility of impacting life satisfaction and long term social outcomes for individuals with ASD. Howlin and colleagues (1999) have suggested some key elements for developing successful ToM interventions which include: (a) breaking information into small steps; (b) providing instruction in normal developmental sequence; (c) systematic reinforcement of behavior; and (d) errorless learning. These principles are all critical elements of effective instruction that can be easily integrated into a computer assisted instructional (CAI) format for delivery, which could lead to more successful outcomes for teaching ToM skills to individuals with ASD. CAI refers to the use of computers to teach a variety of skills and includes computer modeling and computer tutors (Collet-Klingenberg, 2009).

Theoretical Foundation of CAI

The current literature shows recurrent evidence that learners differ in numerous ways and effective teaching should regard and address these differences in order to maximize
student progress (Molenda, 2012). In 1911, Edward L. Thorndike developed general principles of learning including the law of effect which states that responses are learned depending upon whether or not they are followed by a preferred outcome. He proposed the first theory of instruction in which readers could only progress through a textbook after they demonstrated mastery of the basic content (Thorndike, 1911).

This concept was expanded by Sidney Pressey (1927) who developed the “machine for automatic teaching of drill material.” Pressey’s machine appeared similar to a typewriter carriage which displayed a question with four answer choices in a window. On one side of the machine were four keys, one of which the neurotypical adult user pressed in order to indicate the correct answer. The machine recorded the answer on a counter at the back of the machine and progressed to the next question. Once the user was finished, the person scoring the test placed the test sheet back into the device and noted the score on the counter. The use of this machine was reported to improve efficiency of teaching content which required drill and practice to develop mastery.

In a separate field of study, along the same time period, visual instruction was gaining support (Molenda, 2012). B. F. Skinner further explicated Thorndike’s law of effect to experimentally demonstrate that laboratory animals elicited complex new behavior when the schedule and intensity of consequences were manipulated. Based on personal experience, he then developed a mechanical device for interactive learning which he referred to as a teaching machine (Skinner, 1954). With momentum in the field for promoting visual instruction, it was soon realized by the professional community it was the software, not the hardware, which was
responsible for successful learning (Molenda, 2008). This belief shifted the focus from the mechanical machine to the instructional method called programmed instruction.

Programmed instruction allows students to interact individually with the content, record answers, and receive constructive feedback, thus enhancing and improving the rate of learning. This teaching model is based on the following instructional principles: (a) distinctive units; (b) small steps; (c) active participation; (d) simultaneous correction; (e) processing of learning, and (f) individualized learning rate (Efendioglu & Yelken, 2010). As the technology advanced, the programmed instruction model was enhanced by incorporating audio-visual devices for greater impact and delivered through desktop computers. In schools today, the theory of programmed instruction is incorporated into “e-learning” and computer-assisted instruction is used to teach a wide array of abstract concepts (Clark & Mayer, 2003). With the growing literature base suggesting the use of CAI as a promising practice for providing individualized instruction for a variety of learners (Ramdoss et al., 2011; Wainer & Ingersoll, 2011), there is a need to delineate the most critical components necessary for CAI to be a successful intervention.

Although there is a growing trend in the use of technology to provide individualized instruction to a variety of learners, there are no specific guidelines to distinguish between assistive technologies, computer-based video modeling, technology integrated instruction support, and computer assisted instruction (Anohina, 2005). With specific reference to CAI, Skinner’s definition of teaching machines based on the theory of programmed instruction, allows for a differentiation between CAI and all other technology based interventions (Benjamin, 1988). This definition stipulates that CAI consists of a student completing educational activities through the use of a computer-based software program. There are three
criteria that differentiate CAI from either assistive technology or computer-based video modeling: (a) CAI uses an automatic self-controlling device that presents a unit of new information (i.e., antecedent); (b) CAI provides a way for the learner to respond to the information (i.e., behavior); and (c) CAI provides feedback about the accuracy of the learner’s response (i.e., consequence; Benjamin, 1988). This definition and criteria for CAI was adopted for the current meta-analysis. The multisensory component along with the controlled and structured learning environment and ability to individualize instruction are some of the features that make CAI a promising intervention tool. Additionally, the delivery of immediate feedback, consistent prompting and cueing to elicit accurate responses, and high levels of reinforcement add to its success.

CAI Interventions for Children with ASD

There is evidence to suggest that children with ASD have difficulty responding to traditional teaching strategies which has led to an increase in the investigation and use of alternative treatments to accommodate for the unique characteristics of children with ASD (Whalen, Liden, Ingersoll, Dallaire, & Liden, 2006). Many students with ASD have responded well to interventions based on visual supports including video modeling (Wainer & Ingersoll, 2011), pictures (Bryan & Gast, 2000; Dettmer, Simpson, Smith-Myles, & Ganz, 2000), and computers (Bernard-Opitz, Sriram, & Nakhoda-Sapuan, 2001). As children with ASD seem to show a strong preference to interventions that involve visual supports, it appears that computers are a promising and logical intervention choice for skills-instruction. Current research has shown that CAI can provide an efficient and effective learning environment for
basic skills-instruction (Bosseler & Massaro, 2003; Coleman-Martin, Heller, Cihak, & Irvine, 2005; Hopkins et al., 2011; Moore & Calvert, 2000).

Recently, the use of CAI to increase ToM skills, specifically emotion recognition, has been evaluated by several researchers, with promising results (Hopkins et al., 2011; LaCava, Golan, Baron-Cohen, S., & Myles, 2007; Silver & Oaks, 2001; Tanaka et al., 2010). Because of the strong visual processing skills in many children with autism and the highly engaging nature of technology-based interventions, there is reason to believe that CAI may be an effective intervention for teaching the complex ToM skills to this population (Samson, Mottron, Soulieres, & Zeffiro, 2012).

Another area explored in the research is the effect of CAI on levels of task engagement. Individuals with ASD typically display difficulty with social exchanges, which very likely affects their engagement during academic instruction as well. This often results in students with autism being excluded from academic activities which involve whole group instruction, small group activities or even working with a partner (Kluth & Darmody-Latham, 2003).

Whether communicating information about routines, academic content, or social expectations, teachers typically rely on verbal language which, for individuals with ASD, is less effective, attributable to their limited ability to process complex verbal information (Dettmer, Simpson, Smith-Myles & Ganz, 2000). This point is illustrated by Carnahan, Musti-Rao, and Bailey (2009) who evaluated the engagement level of six students with ASD during reading tasks incorporated with visual stimuli through an interactive reader paired with music. It was observed that all participants were more actively engaged in reading tasks during the interactive reader condition compared to baseline level which used traditional books with no
visual supports or interactive materials. This outcome supports previous claims that visually stimulating materials can increase task engagement for students with ASD. Additionally, if the use of technological tools is effective for teaching academic content to students with ASD, it stands to reason that the use of CAI to teach social understanding or perspective taking skills would more likely improve social performance of students.

One possible explanation for an overall lack of task engagement in social situations and live-instruction is aversion to facial gaze (Vivanti, Nadig, Ozonoff, & Rogers, 2008). This theory is supported in a study by Hobson and Hobson (2007) in which researchers evaluated amount of time children with autism spent attending to the face of live models when asked to imitate actions. The results indicated that participants with ASD spent significantly less time attending to the face compared to participants without ASD. This lack of attention to faces and a tendency to avoid eye contact impacts individuals’ ability to imitate actions or engage in social conversations in a meaningful manner.

Strong visual processing skills but aversion to social interaction and reduced ability to encode verbal language may explain preference for computerized instruction, which may be more reinforcing and therefore more engaging than live instruction for individuals with ASD (Kaey-Bright & Howarth, 2012). When individuals are more engaged in the learning environment, their skills are more likely to transfer into untrained environments. This is critical for significant life improvements versus simply improving limited skills in limited settings.

Generalization of ToM Skills

One concern raised by autism researchers is that unless appropriately designed, the use of CAI might reduce opportunities for social interactions and social problem-solving (Bernard-
Opitz et al., 2001). Because lack of social interaction skills is a core deficit in autism, it is crucial that interventions used to remediate these deficits focus on generalization training. Generalization, the ability to transfer a learned behavior acquired during a training activity to another setting, person, or activity with similar but different materials has been identified as a difficult task for children with ASD (Koegel, Egel, & Williams, 1980). To illustrate this point, Froehlich et al. (2012) examined prototype formation in 27 adults with high functioning autism using a dot pattern task and compared them to a control group of 25 neurotypical adults. The task trained individuals to categorize dot patterns into three separate groups. During assessment procedures, participants were shown 45 different dot patterns, the original 9 from training and 36 new patterns which were a distorted version of the training stimuli. They were required to label the pattern as “new” or “old,” and provide the appropriate category to which the dot pattern belonged. The results demonstrated no significant differences between individuals with ASD and the control group in recognizing and categorizing familiar and minimally distorted dot patterns. However, as distortion of dot patterns increased, individuals with ASD identified significantly fewer patterns to the correct category when compared to the control group.

Although, dot pattern recognition does not directly relate to generalization of life skills, this study highlights a cognitive processing deficit in ASD which has been characterized by a difficulty processing information from multiple stimuli and difficulty with generalization to untrained stimuli or situations (Froehlich et al., 2012). Additionally, these findings support previous generalization research which suggests as new stimuli become less similar to previously learned stimuli, individuals with autism have more difficulty transferring skills.
One effective method for programming generalization is to provide sufficient exemplars of stimuli. This strategy involves teaching students to respond correctly to multiple forms of antecedent stimuli and providing opportunities for generalization to novel situations which share some common stimulus properties. To ensure that CAI is an effective intervention, generalization of skills must be incorporated as a critical component of the intervention. Identifying innovative interventions that have empirical evidence of remediating core deficits in autism is critical to the field as this population continues to grow and the prevalence rate increases. This is compounded by a social trend of increasing demands within school and work environments (Adreon & Stella, 2001).

Need for Expanding the Existing Evidence Base on Effective Interventions

Recent statistics published by the Centers for Disease Control and Prevention indicate the prevalence of autism to be one out of fifty children (CDC, 2013), which is an increase of more than 500% in the last decade. With the increasing number of children being diagnosed with ASD, the need for effective educational interventions is becoming increasingly important. While a wide variety of interventions have been designated as being empirically effective by the National Autism Center (NAC, 2009), there are many other interventions that lack empirical support for remediating autism deficits.

The NAC report (2009) described technology-based interventions as an emerging practice for skills instruction on a variety of target behaviors for children with autism (ASD), ages 6-14 years. Additionally, the nineteen studies evaluated in this category included various types of technology-based interventions compiled in a single category (e.g., CAI, Alpha Program, Delta Messages, the Emotion Trainer Computer Program). Similarly, interventions for
improving theory of mind skills were also designated as having emerging evidence of
effectiveness. While interventions in this category targeted both high and low functioning
children with ASD, ages 6-14 years, only four studies met the inclusion criteria for evaluation
(Bell & Kirby, 2002; Fisher & Happe, 2005; Gevers, Clifford, Mager, & Boer, 2006; Wellman et
al., 2002). The same is true of the recent report published by Wong and colleagues (2014) that
lists Technology-Aided Instruction and Intervention as an EBP. However, this classification
included computer aided instruction, speech generating devices/VOCA, smart phone and
tables, and virtual networks but only 5 of the 20 studies focused on CAI to improve mind-
reading skills. In other words, further research in this area is needed.

Purpose of Literature Review

The reauthorization of the Individuals with Disabilities Act (2004) mandates the use of
evidenced based practice in the education of individuals with disabilities. While the National
Standards Report (NAC, 2009) states that technology based interventions are an emerging
intervention they have not been established as an EBP. To date, no systematic meta-analysis
has been conducted to establish the efficacy of Computer Assisted Instruction as an EBP
utilizing rigorous evaluation criteria such as those developed by Reichow et al. (2008; 2011).
Pennington (2010) conducted a literature review which analyzed the use of CAI for a variety of
skills including social-emotional, communication, and academic. The results of this review
indicated that CAI may be effective for teaching a limited set of academic skills, specifically
literacy, however there were many methodological issues identified with the included studies
such as no control groups in group design studies and no replication of treatment effect for the
single subject research design studies. The purpose of this systematic literature review is as
follows: (a) to synthesize the literature on experimental research utilizing CAI for students with autism, (b) to extend analysis of this research focusing exclusively on CAI, and (c) determine if CAI meets the evidence-based practice criteria described by Reichow et al. (2008; 2011).

Method

Locating Studies

The focus of the systematic literature review was on computer assisted instruction (CAI). In locating relevant studies, journal articles were searched utilizing a university EBSCO-host database (e.g., Academic Search Complete, ERIC, Education Research Complete, JSTOR, and Behavioral Sciences Collection) without date restriction. Specific key words used in the search were as follows: Autism, ASD, CAI, computer based learning, technology based learning, technology based instruction, and assistive technology. Other articles were found by cross-referencing citations from previously identified peer-reviewed studies. Of the 40-plus studies that were identified, 21 studies, published from 1995-2013, met the criteria for inclusion in this review.

Inclusion/exclusion criteria. The following parameters were used as inclusion/exclusion criteria for research reviewed in this literature review.

1. The intervention (i.e., the independent variable) in the study was required to be a computer-assisted instruction strategy based on B.F. Skinner’s definition. This definition includes the features of an automatically self-controlling device, learner interaction with the device, and the provision of feedback to the learner. The CAI intervention had to be the primary intervention component in the research study. Non-examples include studies in which an assistive or mobile technology device
(e.g., PDA or iPad) or video modeling delivered through a computer (as a tool) was used as the primary intervention.

2. At least one participant in each selected study must have been diagnosed with an autism spectrum disorder. This includes assigned labels like autism, autistic disorder, asperger’s syndrome, or pervasive developmental disorder – not otherwise specified (PDD-NOS). Non-examples included research studies in which computer-assisted instruction was used for participants diagnosed with intellectual or developmental disabilities other than ASD, unless the diagnosis was comorbid with ASD.

3. The research design of included studies must have been a design and methodology that permitted an evaluation of the effects of the independent variable (CAI) on the dependent variable (i.e., participant behavior). These are the designs that meet the criteria for inclusion: experimental design with group comparison, quasi-experimental designs, and single-subject designs (i.e., multiple-baseline or reversal designs). Non-examples included single-subject designs that did not include a baseline to compare with intervention and studies that utilized a non-experimental research design.

4. Included studies needed to be published in the English language in a peer-reviewed journal. These options were selected on the “advanced search” option on the EBSCO-host database. Non-examples included articles published in a language other than English without an English translation or in a publication in which the editor alone is the primary reviewer for published articles.
Each identified study was first assessed for inclusion or exclusion. Then, studies selected for inclusion were summarized in terms of the following features: (a) participant characteristics, (b) skills targeted, (c) details regarding the CAI, (d) intervention outcomes, and (e) certainty of evidence. Various procedural aspects were also noted, including setting, experimental design, and inter-observer agreement (IOA). Outcomes of CAI on a variety of dependent measures were summarized for each study (see Tables B1 & B2).

Results

Study and Participant Characteristics

Twenty-one studies were included in this systematic literature review involving a total of 341 participants with ASD. The participants ranged in age from 3 to 30 years, with the majority being 6 to 12 years. In terms of setting, most of the interventions (13) were conducted in school settings; two were in the participants’ homes, one was in a mixed home and school setting, and two were in a university clinic setting. Three of the studies did not specifically describe the setting where the intervention was delivered. The majority of studies focusing on academic and communication skills included students with low functioning ASD while studies which targeted social-emotional skills included participants with high functioning ASD. This is not a surprising finding as individuals with HFA are typically on grade level and show average to above average cognitive functioning therefore would not be the target for academic interventions. More difficult skills such as emotion recognition and social interactions, however, seem to be more suited as an intervention target for individuals with HFA as evidenced in the literature.
Research Design

Nine of the identified studies utilized a single-subject research design. Of these, the majority (6) targeted academic skills. One study which addressed decreasing problem behavior was also a single-subject study. Two studies addressed communication and one focused on social-emotional skills. Within the single subject designs there were many different experimental designs employed. The majority of studies used a multiple baseline design (7) across participants, setting, and task. Others included alternating treatment (2), multiple-probe (1), and one study utilized a multi-component reversal design.

Conversely, most studies utilizing group design research (12) targeted social skills and emotion recognitions skills (7). The remaining five studies targeted academic skills. Within the group design methodology five studies included description of randomized group assignment and the remaining six either assigned participants to groups based on disability criteria or did not include information about group assignment. Generally, studies evaluated results with a t-test (4) or ANOVA (4) though two used ANCOVA and one study utilized a non-parametric test of significance.

Targeted Skills and Outcomes

Interventions using computer-assisted instruction can be designed to teach a wide variety of skills and behaviors. In the articles selected for this review, various skills and behaviors were targeted for change, with some of the studies targeting multiple skill domains. All studies indicated that CAI was associated with acquisition or improvements in some skills or behavior for at least some of the participants (or groups).
Social-emotional skills. Eight studies evaluated the effect of CAI on social/emotional skills (Bernard-Opitz, Sriram, & Nakhoda-Sapan, 2001; Bolte et al., 2006; Hopkins et al., 2011; LaCava, Golan, Baron-Cohen & Myles, 2007; LaCava et al., 2010; Moore, Cheng, McGrath & Powell, 2005; Silver & Oakes, 2001; Tanaka et al., 2010). Seven of these studies specifically focused on the skill of recognizing emotions through facial expressions (Bolte et al., 2006; Hopkins et al., 2011; LaCava et al., 2007; LaCava et al., 2010; Moore et al., 2005; Silver & Oakes, 2001). For these seven studies, CAI appears to be effective in teaching emotion recognition skills to participants with ASD.

For example, Silver and Oaks (2001) designed the Emotion Trainer which incorporated photographs of people displaying emotions, along with multiple choice cartoon drawings as response options (i.e., happy, sad, angry, and afraid) to the question “How does ________ feel?” They utilized this software as an intervention in a study with 22 participants with high-functioning autism, ages 10-18 years, with 11 each in the experimental and control group respectively. Participants in the experimental group completed an average of 8.5 computer sessions over a 2 to 3 week period. Authors conducted a pre-post assessment to evaluate the impact of the software using the Facial Expression Photographs test, Happe’s Strange Stories test, and emotion recognition cartoons. Results showed significant reduction in the mean number of errors made from the first to the last session but noted that there was a large amount of individual variation related to the number of times a participant used the computer program. However, the limited sample size and lack of procedural fidelity regarding the amount of access individual participants had to the software raises concerns about the validity of these findings.
Similar results were found by LaCava et al. (2007) who taught eight children, ages 8-11 years, to recognize emotions using a software program known as Mind Reading. This software included an emotions library catalogue of over 400 different emotions which were presented in photographs, short movie clips, and audio clips along with contextual examples. Participants used the software for 10 weeks in their home for an average of 10.5 hours over the duration of the intervention condition. Pre- and post- intervention assessment of participants’ scores on a variety of standardized measures including the Cambridge Mindreading Face-Voice Battery for Children (Golan & Baron-Cohen, 2006), revealed a significant improvement in emotion recognition skills for all 8 participants. Similar to the Silver and Oaks (2001) study described above, there was great variability in the way participants used the software and there was little control over the impact of extraneous variables such as parental input and amount of time accessing the software. Additionally, results need to be interpreted with caution because of the small sample size in the study.

In another study conducted by Hopkins et al. (2011), children with high-functioning autism (HFA) and low-functioning autism (LFA), ages 6-10 years, were compared on their responses to Face Say, an emotion recognition training software. Participants were presented with arrays of photographs and schematic drawings of emotions which consisted of six faces. As they were presented with a stimulus item, participants heard a label (i.e., angry, disgusted, scared, happy, sad, or surprised) and were asked to match the corresponding emotion by touching the appropriate photograph or drawing. Interestingly, while both groups of children significantly improved their ability to recognize emotions in a photograph, only the group of children with HFA showed improvement in recognizing emotions in a schematic drawing. These
results indicate that CAI can be effective for emotion recognition training across a range of individuals on the spectrum; however, it appears that a certain level of cognitive functioning may be necessary to improve complex skills that are critical for generalization of behavior to social situations in natural environments. Given that previous research has documented the challenge of teaching generalized emotion recognition and perspective-taking to individuals with ASD (Baron-Cohen 1989; Ryan & Charragain, 2010; Tager-Flusberg, 2007), it may be worthwhile to analyze alternative explanations for lack of skill generalization for this population.

Bolte et al. (2006) evaluated the impact of CAI on emotion recognition at the neurological level in order to help explain why individuals with ASD continue to have difficulty generalizing skills and why there is so much variability in the response to ToM interventions. These authors used the Frankfurt Test and Training of Facial Affect Recognition (FEFA) software to teach 10 individuals with high-functioning autism skills in facial affect identification. Subsequently, they evaluated the software’s cognitive impact by using functional magnetic resonance imagining (fMRI). Results revealed that, even though participants showed significant improvement in emotion recognition on the software program, no significant changes in activation levels were observed in the image of the fusiform gyrus (FG), the area of the brain associated with emotion recognition and affect. This indicates that ToM interventions may be successful at teaching compensatory strategies but training may not necessarily improve cognitive capacity of individuals with ASD as measured using fMRI. The implication of such findings is fundamental for understanding why certain types of ToM skills (e.g., mind-reading or perspective taking) may not generalize to natural settings even for individuals with high
functioning autism. Yet, researchers are compelled to develop training materials to enhance mind reading skills of individuals with ASD (Howlin et al., 1999) given the social relevance of such skills-instruction.

In a study evaluating social problem solving skills by Bernard-Opitz et al. (2001), 5 out of 8 participants with ASD demonstrated an increase in novel social problem solving ideas. However there was a large amount of variability in performance across subjects and when compared to participants who were NTD, they produced significantly fewer novel ideas.

The last study in this review targeting social-emotions skills (i.e., Tanaka et al., 2010) targeted face recognition skills. This differed from emotion recognition in that participants used a program designed to increase their ability to identify facial images across expression changes and attend to information in the eye region of the face, rather than identifying specific emotions based on facial expressions. Seventy-nine individuals with ASD were randomly assigned to a treatment group or a waitlist control group. The treatment group, which was composed of 27 participants, received intervention for an average of 20 hours over 19 weeks using the *Let’s Face It!* computer software program. This program includes a variety of drill activities and games designed to increase face recognition. The authors reported a significant improvement for the treatment group in both analytic and holistic face processing. As a result of the intervention, the participants increased their ability to recognize facial features, including mouth and eyes.

The results of these studies combine to create promising evidence for the ability of CAI to teach emotion recognition skills at the primary level and enhance compensatory strategies for processing emotions. However, the research is limited and has many weaknesses such as
small numbers of participants, lack of control within experimental conditions, lack of
assessment for generalization, and lack of training at the more complex level of mind reading
(i.e., perspective taking).

Academic skills. Ten studies (Bosseler & Massaro, 2003; Coleman-Martin, Heller, Cihak,
& Irvine, 2005; Coleman-Martin et al., 2012; Heimann, Nelson, Tjus, & Gilberg, 1995; Hetzroni &
Shalem, 2005; Travers et al., 2011; Whalen et al., 2010; Whitcomb et al., 2002; Williams et al.,
2002; Yaw et al., 2011;) focused on the acquisition of academic skills. The academic skills
targeted in these studies were similar, with the dependent variables being measures of skills in
literacy, sight word acquisition, and vocabulary. All of these studies reported successful levels
of mastery of target skills (i.e., the reading of printed words or identification of vocabulary
words) and one study reported that CAI was superior to teacher led instruction for efficiently
teaching word identification skills.

Specifically, Coleman-Martin et al. (2005) compared the use of CAI to teacher led
instruction for teaching word identification to three children, one of whom had low-functioning
autism. Researchers used the nonverbal reading approach (NRA) to conduct lessons which
focused on phonemes and word production. During the CAI condition, the NRA curriculum was
converted into a PowerPoint presentation. They used a multiple baseline design across
participants with three conditions (i.e., teacher led instruction, CAI + teacher led instruction,
CAI alone) to evaluate the relative effectiveness. Results showed that all three participants
were able to acquire new vocabulary words during all three conditions but the acquisition rate
for the participant with ASD during the CAI alone condition was significantly higher than the
other two conditions.
In a contrasting example, Travers et al. (2011) compared acquisition of alphabetic skills between TLI and CAI for 17 participants, ages 3 to 6 years, with ASD. The results of revealed that both forms of instruction were successful at teaching alphabet skills and there was no significant difference between groups in performance on standardized measures following intervention. Additionally, they found no significant difference in attention and rate of problem behavior displayed during instruction between the two instructional methods. Both instructional methods led to maintenance of skills at similar levels.

To evaluate the usefulness of CAI in teaching individuals with ASD, Bosselor and Massaro (2003) developed a computer animated tutoring program to teach vocabulary and grammar to eight children with low-functioning autism. Within the software an animated face, “Baldi,” led the participants through vocabulary lessons by visually modeling speech production and providing feedback for student responses. Results demonstrated that CAI was successful in teaching all eight participants a significant amount of new vocabulary words. In a further analysis, six participants were able to generalize these skills by using the vocabulary words to identify novel images (Bosselor & Massaro, 2003). A possible explanation for such a high rate of success in generalization was the training to 100% mastery during the treatment condition and use of additional generalization training if students did not identify two of the three novel images during initial generalization probes.

Behavior. Recent research has even demonstrated superior outcomes for reducing problem behavior using CAI when compared to paper-based social stories (Mancil, Hayden & Whitby, 2009). Researchers used an alternating treatments design to compare the effectiveness of CAI vs. paper-based social stories for reducing pushing behavior in three
children with high-functioning autism. Authors converted a social story into PowerPoint presentation for the CAI condition and participants would alternately read the paper-based or the computer-based social story prior to transition to recess or the classroom. Results indicated that while both methods were effective, participants had slightly fewer instances of pushing during the CAI condition compared to paper-based (i.e., control condition). Additionally, all three participants chose CAI as their preferred intervention method when interviewed at the conclusion of the study suggesting that it is an engaging and socially valid teaching method. These results should be interpreted with caution, however, because it is difficult to assure that there were no carry over effects from the CAI to the paper based condition due to the experimental design: ABABCBC. This study provides promising support for the use of CAI to reduce problem behaviors however further investigation is needed utilizing sound single subject research methodology.

Communication. Two studies targeted increasing communication skills. Hetzroni and Tannous (2004) evaluated the use of CAI to teach function communication to 5 participants, ages 7-12 years, with low functioning autism in a classroom setting. They evaluated the impact CAI on echolalia, relevant speech, and child initiations through direct observation and reported that on average all participants improved on all communication skills measured. It must be noted that several of the participants had near zero levels of communication behavior during baseline therefore there was a very large percentage of overlapping data which made it difficult to determine to overall effect of the treatment.

In another study, Kodak, Fisher, Clements, and Bouxsein (2011) compared the amount of correct prompted and unprompted tacts (i.e., labels) during CAI versus one-to-one
instruction for one participant with ASD. Their results demonstrated high levels of independent responses during CAI and no independent responses during one-to-one instruction suggesting that CAI may be more effective for reducing prompt dependence.

Combination skills. CAI programs have also been developed to provide instruction on a combination of skills which may include any of the following: (a) academic, (b) communication, (c) behavior, (d) social-emotional, and (e) adaptive/functional life skills. Whalen et al. (2010) conducted a study evaluate the effectiveness of a CAI software program: Teachtown: Basic, which provides instruction on a combination of skills including language, academics, and social skills. They employed a group design including 47 students, ages 3 to 6 years, in a preschool and K-1 school setting. Authors randomly assigned 8 classrooms to either a CAI treatment group or a control group and compared the performance on a variety of standardized measures including the Peabody Picture Vocabulary Test (PPVT), Expressive Vocabulary Test (EVT), and Brigance Inventory of Early Development (Brigance).

Results demonstrated varying levels of success using the CAI software. On standardized measures preschool students in the CAI group had significant increases in scores on the PPVT but not on the EVT. The K-1 group did not demonstrate significant improvements on either the PPVT or the EVT. Additionally there were no significant improvements for either group on the Brigance. These results indicate that there may be underlying participant characteristics which affect response to intervention and performance on standardized measures. However, authors did find a significant correlation between the amount of time spent on the computer software and increases in Brigance scores. This is a promising finding and further investigation should
evaluate the relationship between intervention intensity, regarding use of CAI, and
demonstration of improved skills by participants.

The impact of participant characteristics on response to interventions has also been
demonstrated in the literature regarding generalization of skills to natural settings.

Generalization of Skills

One concern within the field of special education and autism research is that, because
CAI is not conducted in a natural setting but rather a more systematic format and contrived
setting, skills may not generalize to untrained settings. Generalization, the ability to transfer a
learned behavior acquired during a training activity to another similar or related activity has
been identified as a difficult task for children with ASD (Koegel, Egel, & Williams, 1980).

Of the twenty studies reviewed, only 5 assessed generalization of skills into a natural or
untrained environment; three evaluated academics and two evaluated social-emotional skills.
Bosseler and Massaro (2003) and Coleman, Hurley, and Cihak (2012) both evaluated the use of
CAI to teach reading skills. The results of both studies showed that participants were able to
genralize their reading skills (targeted words learned through CAI) to untrained settings such
as in the classroom in non-computer based activities.

Hetzroni and Shalem (2005) evaluated the generalization of reading skills taught
through a CAI prompt fading program using a single subject research design. The study
included eight participants, ages 10-13 years, with ASD and limited or no verbal language.
Participants were taught to read orthographic symbols from package logos. The results showed
that, though all participants were able to accurately identify all 8 orthographic during CAI, there
was a large amount of variability in generalization of skills. Six of the eight participants
successfully generalized reading orthographic symbols to food package logos and food items but two participants were not able to generalize the skills. Possible explanations for lack of generalization in these two participants were absenteeism, inconsistent use of CAI program, and emotional distress followed by hospitalization.

More specifically, research evaluating social skills interventions delivered through CAI has shown mixed results regarding generalization of skills. For example, Sansosti and Powell-Smith (2008) evaluated the use of computer-based social stories with embedded video modeling to increase the social communication in three children with autism, ages 6-10 years, using a multiple baseline design across participants. The results showed that “joining in” behavior increased for all three participants during intervention, which also included adult prompting. However, only one of the three participants generalized these skills to a natural setting (which did not include computer-based social stories, video modeling, or teacher prompting).

Difficulty with generalization was also noted in studies by Hopkins et al. (2011) and LaCava, Rankin, Mahlios, Cook, and Simpson (2010). In both studies, researchers evaluated generalization of emotion recognition skills. Both studies revealed inconsistencies and large variability in generalization of skills. The variability in generalization of skills was also demonstrated in the Hopkins et al. (2011), for teaching emotion recognition skills using the CAI program Face Say. They evaluated the performance of three groups of children, ages 6-10 years (i.e., high functioning autism, low functioning autism, neurotypically developing). Similar to the Sansosti and Powell-Smith (2008) study, all 49 participants increased in their ability to recognize emotions during the CAI intervention phase. However, when authors evaluated how
these newly acquired emotion recognition skills impacted social interactions in the natural environment, they found mixed results. Through direct observation of social skills the authors determined that participants with LFA showed a significant decrease in inappropriate and negative interactions but no significant change in total amount of interactions and positive interactions. In contrast, participants with HFA had a significant decrease in inappropriate interactions and a significant increase in positive interactions but no change in negative interactions or total number of interactions. Additionally, based on parental reports, individuals with HFA generally demonstrated increased cooperation and sharing behaviors whereas individuals with LFA demonstrated reduced negative behaviors and increased self-control. These results must be interpreted with caution however, because of possible parental bias as opposed to direct observation of social skills. Also, there is no obvious link between emotion recognition and cooperation or self-control, which leads to further questions about the validity of these results.

Similarly, LaCava et al. (2010) evaluated the CAI program: Mind Reading, on its ability to improve emotion recognition skills and enhance social interactions. This study included four participants, ages 7-11 years, with ASD. They found a large amount of variability in participant responses to intervention and generalization of skills. While all 4 participants increased ability to identify emotions in picture, only 3 were ability to identify emotions in cartoons and black and white schematic drawings, which indicates a lack of generalization when stimuli become more abstract. During direct observation, researchers reported large amounts of overlapping data across participants and conditions, which suggests that the software was not a reliable intervention for generalization of emotion recognition skills.
Combined, these results on generalization of skills indicate that there are external variables that impact participants’ likelihood to apply learned skills in new settings. Academic skills appear to be more likely to generalize, specifically literacy skills, compared to emotion recognition skills. However, authors who evaluated social-emotional skills assessed generalization of emotion recognition skills to improve social interactions which may not necessarily be a related skill-set. It is possible that prerequisite skills in emotion recognition and mind-reading must be mastered before successful social interactions in the natural and untrained environments are expected of individuals with ASD.

Many published studies that used CAI to teach emotion recognition or increase ToM skills did not include assessment of generalization or exposure to social skills training in the natural setting (Bolte et al., 2006; LaCava et al., 2007; Silver & Oaks, 2001; Tanaka et al., 2010; Whalen et al., 2010). Assessment of generalization of learned behavior is not only critical for social relevance of research, but a critical measure for meeting the criteria for high quality research (Reichow, Doehring, Cicchetti, & Volkmar, 2011).

Existing research has suggested that a lack of generalization may partly be due to stimulus over-selectivity observed in individuals with autism, especially in situations where natural settings do not share common stimuli with simulated instructional conditions (Koegel et al., 1980). Stokes and Baer (1977) stated that generalization occurs only when programmed as an integral component of intervention and not as an outcome one hoped to achieve. This highlights the importance of choosing target behaviors that will more likely meet naturally existing contingencies of reinforcement, such as social interaction with peers for individuals...
with HFA, to increase the likelihood that behaviors will generalize to natural settings (Cooper, Heron, & Heward, 2007).

Determining if skills learned through CAI generalize to natural settings is critical for establishing the social significance and impact of a particular intervention. Social significance can be further established by assessing the social validity of an intervention.

Social Validity

Social validity is concerned with the assurance that skills taught during intervention are socially important rather than just convenient for the study (Gast, 2010). Wolf (1978) recommended that social validation be evaluated on three levels: Goals, procedures, and effects. Social validation is typically conducted by gathering data from guardians, participants, teachers, and therapists in order to assess and obtain insight into how the treatment or intervention is viewed by consumers.

Of the twenty studies evaluated in this review, only six included measures to obtain social validity. All participants that were assessed for preference on CAI reported that they enjoyed using CAI and preferred it to traditional teaching strategies. Additionally, teachers reported positive outcomes for CAI and stated that it was easy or convenient to install and use. Some teachers and parents reported a noticeable increase in social interactions and reduction in problem behaviors as a result of CAI, though these reports were anecdotal. In one study (LaCava, Ranking, Mahlios, Cook, & Simpson, 2010) teachers reported that though students seemed to enjoy CAI, it did not result in noticeable gains in social interactions compared to neurotypical peers.
The collaborative results for social validity assessment suggest that CAI appears to be a socially valid intervention. Parents, teachers, and participants value its use for the training of a variety of skills including academic and social. Additionally, it is efficient and easy to use and parents, though this may be the result of placebo effect, report that their children’s behavior or skills changed as a result of using CAI.

Conceptual Overview of Evaluation System

A standardized, empirically-validated, and structured process was used to determine the strength of research and level of evidence for Computer Assisted Instruction for individuals with ASD (Reichow et al., 2008; 2011). These critical evaluation criteria consisted of six primary quality indicators for SSR and group design and 7 secondary quality indicators for single subject research and 8 secondary quality indicators for group design. Each study was independently coded by a reviewer on its primary and secondary quality indicators. Studies were assigned a research strength rating of “strong,” “adequate,” or “weak” according to the number of primary and secondary quality indicators (see Table B3).

Evaluation Criteria

Quality indicators were derived from the Reichow et al., (2008; 2011) criteria and addressed various criteria for group design (see Table B4) and single subject research designs (see Table B5). Criteria were addressed in both primary and secondary to address the various weighted value of certain criteria.

Support for CAI as an Evidence-Based Practice

In terms of the evidence for CAI being considered an evidence-based practice, when the Reichow et al. (2008; 2011) evaluation criteria were applied, the majority of the studies (17)
received an “adequate” rating and one study received a “weak” rating. Three studies received the highest rating of “strong” only one of which used a single subject research design and methodology. The results of the Reichow et al. (2008; 2011) evaluation criteria are presented in Tables B6 and B7.

Reichow and colleagues (2008; 2011) describe an established evidence-based practice as a treatment that has demonstrated effectiveness “across multiple methodologically sound studies” conducted by at least two separate research teams in different geographic locations. These authors provide a formula (see Table B8) for the calculation of a score which may be used to determine the status of a treatment or intervention as an evidence-based practice (EBP). This formula takes into account the number of strong group studies (Group₅), the number of adequate group studies (Groupₐ), the number of participants for whom the treatment was successful from the strong single-subject studies (SSED₅), and the number of participants for whom the treatment was successful from the adequate single-subject studies (SSEDₐ). Studies receiving a “weak” rating are not considered (in this formula) for determining EBP status.

After a score is obtained from the formula, the score is compared to different levels of EBP. According to Reichow et al. (2008; 2011), treatments with scores from 0-30 points are not considered an EBP, treatments with scores from 31-59 are considered a promising EBP and treatments with scores 60 or greater are considered an established EBP. Based on the studies included in this review, CAI appears to meet the criteria for an established evidence-based practice with a score of 215. However, these scores reflect the use of CAI for all the dependent variables (e.g., academic, social-emotional, communication and behavior). Similarly, the scores are also a reflection of the number of participants in a study. Given that most studies utilized
group design research, a higher number of participants is likely to inflate the score as per the formula by Reichow and colleagues (2008; 2011). The scores are likely to change if the effectiveness of CAI is separately evaluated for each of the dependent variables and each of the research designs (i.e., group vs. single subject).

Discussion

The results of this literature review reveal much promise in the development of CAI as an evidence-based practice (EBP). Overall, most of the studies revealed significant improvements for participants in the specified dependent variable. According to the criteria for high quality research described by Reichow et al. (2008; 2011), this intervention should be regarded as an EBP, however based on the limited amount of research available, and the methodological constraints presented by this review, these results should be taken with caution. There were several limitations revealed across these studies.

One limitation was the lack of a specific and consistent definition of computer assisted instruction across all studies. There appears to be extreme variations in the way technology is utilized for instructional purposes and often researchers left critical descriptions of their technology out of the discussion. This reduces the consumers’ ability to make sound comparisons across published research. Additionally, treatment components across the study were often not well described leaving the reviewers unsure of components such as use of prompting, reinforcement, and interventionist’s role.

Another limitation discovered through this analysis was the lack of thorough description of participant characteristics such as standardized scores that would reveal cognitive ability and communication skills; both are critical characteristics of individuals with an autism spectrum
disorder. Due to the wide variability of individuals with (ASD) it is important to include detailed
descriptions of participant characteristics to determine the ability to generalize findings.
Without this information practitioners and researchers may be unable to make decisions about
further implementation of the treatment or intervention.

Additionally, there were many methodological concerns related to the identified
studies. For example, many of the group design research studies had small sample sizes (less
than 10), which does not provide adequate power for statistical analysis. Also, many group
design studies did not include randomized group assignment but rather assigned participants
based on disability categories or convenience.

This adds to the concern about Reichow and colleagues (2008; 2011) evaluation criteria
for establishing EBP as these methodological issues may not have been reflected in the overall
rating of studies. For example, Bosselore and Massaro (2003) conducted a group design study
with only nine participants and did not include randomized assigned of groups. Yet, based on
the criteria the study was rated as adequate which adds fifteen points to the overall score for
determining EBP.

Within the single subject research studies, procedures used did not allow for replication
of effect or control for carryover during comparison designs. One of the biggest concerns,
however, was the large amount of published research that did not report measurement of
effect size. Reporting effect size is critical because it allows consumers to assess the magnitude
of change that occurred as a result of the intervention.

The results of this analysis also reveal an inherent bias in the value placed on group
design studies compared to single subject research design by the evaluation criteria presented
by Reichow et al. (2008; 2011). The group design studies in this review typically had between 10 and 20 participants which is a fairly small sample size. As noted previously, the formula assigns group design studies to be weighted at three times the amount given to single subject research. This inadvertent bias appears to inflate the overall status given to the intervention when Reichow et al. (2008; 2011) standards are applied. It can be argued that, while group design research is very useful, single subject research design may better serve to answer questions of treatment effectiveness for individuals with ASD because this population displays a large amount of variability in academic, communication, cognitive, and social skills. It may be necessary to reevaluate and adjust values that are placed on the results of SSR compared to GD.

Due to the limited amount of research identified across all domains, future research is needed to evaluate the efficacy of CAI. There continues to be a need for research which utilizes sound research designs and includes critical components such as: detailed participant characteristics, operational definitions of the dependent variable and fidelity of implementation, report of inter-observer agreement and effect size, and assessment for maintenance and generalization of skills acquired.

References


doi:10.1177/10883576050200020401


doi:10.1017/S0954579406060305


doi:10.177/1362361310362085

doi:10.1002/(SICI)1099-078x(199805)


doi:10.1177/1088357609347324


doi:10.1007/210803-009-0901-6


*Asterisks indicate studies included in this review.*
APPENDIX A

SUPPLEMENTAL TABLES
Table A1

**Summary of Single Subject Research**

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Research Design</th>
<th>Participants/ Location</th>
<th>DV</th>
<th>IOA for DV</th>
<th>Intervention</th>
<th>FOI</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bosseler &amp; Massaro (2003)</td>
<td>Exp 1: pre &amp; post assessment</td>
<td>Exp 1: 9 total, ages 7 to 12 years, with ASD</td>
<td>Vocabulary</td>
<td>NA</td>
<td>CAI program containing synthesized speech and image of vocabulary items delivered by virtual tutor “Baldi.”</td>
<td>NA</td>
<td>Exp 1: All 49 words mastered at post assessment 42/49 maintained</td>
</tr>
<tr>
<td></td>
<td>Exp 2: multiple-baseline across word sets</td>
<td>Exp 2: 6 of the original participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Setting: Day program for children with ASD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coleman-Martin, Heller, Cihak &amp; Irvine (2005)</td>
<td>Multiple baseline across conditions</td>
<td>3 total, 1 with ASD, age 12</td>
<td>Word identification</td>
<td>97.3% for 25% of sessions</td>
<td>3 conditions (all instruction using NRA): (a) teacher only (b) teacher led + CAI, and (c) CAI only</td>
<td>98.6% across all participants for 25% of sessions</td>
<td>For participant with ASD: during BL accuracy was below 20%; during TL condition criterion was reached in 4 sessions; during TL + CAI criterion was reached in 6;</td>
</tr>
</tbody>
</table>
Table A1 (continued).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Research Design</th>
<th>Participants/Location</th>
<th>DV</th>
<th>IOA for DV</th>
<th>Intervention</th>
<th>FOI</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coleman, Hurley, &amp; Cihak (2012)</td>
<td>Alternating treatments design</td>
<td>3 total, elementary aged, 1 with ASD</td>
<td>Functional sight words</td>
<td>NA</td>
<td>CAI: sight words integrated into PowerPoint</td>
<td>NA</td>
<td>CAI: increase of 77% accuracy; criterion reached in 24 sessions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TL: sight words presented on flashcards</td>
<td></td>
<td>TL: increase of 78.11%; criterion reached in 19 sessions</td>
</tr>
<tr>
<td>Hetzroni &amp; Shalem (2005)</td>
<td>Multiple-probe design across participants</td>
<td>6 total, ages 10-13 years, with ASD, non-verbal or limited language</td>
<td>Reading</td>
<td>97% and 96% for pre and post intervention on 20% of sessions</td>
<td>Package logos displayed via CAI went through a seven stage fading process. Last step was the</td>
<td></td>
<td>98% across all data points</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 participants learned all 8 logos in 5-11 sessions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 participants were able to</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Research Design</td>
<td>Participants/Location</td>
<td>DV</td>
<td>IOA for DV</td>
<td>Intervention</td>
<td>FOI</td>
<td>Outcome</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------</td>
<td>---------------------</td>
<td>---------------------------------------------------</td>
<td>-------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hetzroni &amp; Tannous (2004)</td>
<td>Multiple baseline across 3 settings</td>
<td>5 total, ages 7 to 13 years, with ASD, limited verbal communication</td>
<td>Communication</td>
<td>92%-95% for all DVs collected for 20% of sessions</td>
<td>CAI software: <em>I Can Word It Too!</em> Addressed three forms of language: form, use, and content through pictures of daily life activities paired with familiar phrases</td>
<td>NA</td>
<td>Delayed echolalia during play and food setting decreased for all</td>
</tr>
<tr>
<td></td>
<td>Setting: School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kodak, Fisher, Clements, &amp; Boussein (2011)</td>
<td>Multi-element design</td>
<td>1 total, age 7 years, with ASD, limited verbal skills</td>
<td>Labeling: prompted and unprompted</td>
<td>98.6% for CAI</td>
<td>100% for CAI</td>
<td>Labeling acquired at same rate for CAI and TL</td>
<td>100% for TL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.8% for TL</td>
<td>CAI: pictures presented on computer with 5s time delay</td>
<td></td>
<td>Relevant speech increase during food and play</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Communication intent increase during play</td>
</tr>
</tbody>
</table>

Table A1 (continued).
Setting: private therapy room at university clinic

prompting procedure

TL: pictures presented by teacher with teacher prompts

Fewer prompted responses during CAI

Generalization for CAI and TL

*(table continues)*

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Research Design</th>
<th>Participants/Location</th>
<th>DV</th>
<th>IOA for DV</th>
<th>Intervention</th>
<th>FOI</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>LaCava, Ranking, Mahois, Cook, &amp; Simpson (2010)</td>
<td>Multiple baseline across participants</td>
<td>4 total, ages 7 to 10 years, with ASD</td>
<td>Setting: school</td>
<td>Emotion recognition</td>
<td>Social interactions</td>
<td>94.7% across 34% of sessions</td>
<td>CAI software: Mind Reading</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CAI software: Mind Reading</td>
<td>94.1% on 18% of days</td>
<td>All made improvements in emotion recognition</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Paper based social story</td>
<td>100% for 30% of sessions</td>
<td>All participants reduced pushing behaviors compared to Paper based social story</td>
</tr>
<tr>
<td>Mancil, Hardon, &amp; Whitby (2009)</td>
<td>Alternating treatment design: ABABCA</td>
<td>3 total, ages 6 to 9 years, with ASD</td>
<td>Setting: school</td>
<td>Pushing (i.e., touching, grabbing, shoving)</td>
<td>93-95% across participants for 50-55% of sessions</td>
<td>CAI social story (using PowerPoint)</td>
<td>Alternately read according to condition</td>
</tr>
</tbody>
</table>

Table A1 *(continued)*.
Whitcomb, Bass, & Luiselli (2011) | Multiple baseline across word sets | 1 total, age 9 years, with ASD | Setting: school | Reading phonetically regular words | 99% for 33% of sessions | CAI software: Headsprout | Student clicks pictures, letter combinations and words | 97% | Participant increased reading accuracy on all four word sets

### Table A1 (continued)

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Research Design</th>
<th>Participants/Location</th>
<th>DV</th>
<th>IOA for DV</th>
<th>Intervention</th>
<th>FOI</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yaw, Skinner, Parkhurst, Taylor, Booher, &amp; Chambers (2011)</td>
<td>Multiple baseline across word lists</td>
<td>1 total, age 12 years, with ASD</td>
<td>Reading Dolch sight words</td>
<td>100% across 60% of sessions</td>
<td>CAI software: Sight words presented on PowerPoint</td>
<td>NA</td>
<td>Participant read 25/30 words correctly after 16 sessions.</td>
</tr>
</tbody>
</table>
Table A2

**Summary of Group Design Research**

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Random Assignment</th>
<th>IV</th>
<th>DV</th>
<th>IOA</th>
<th>Statistical Analysis</th>
<th>Effect Size</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bernard-Opitz, Sriram, &amp; Nakhoda-Sapan (2001)</td>
<td>16 total, ages 4 to 8 years, 8 with ASD, 8 NTD</td>
<td>Group assignment based on presence/absence of ASD</td>
<td>CAI software: designed by authors to teach problem solving</td>
<td>Providing novel solutions to social problems</td>
<td>NA</td>
<td>ANOVA</td>
<td>NA</td>
<td>Participants with ASD had significantly fewer novel ideas. Increase in development of novel ideas was variable; 5/8 had significant increases. 7/8 generalized skills.</td>
</tr>
<tr>
<td>Bolte et al. (2006)</td>
<td>10 total, ages 20-40 years, with ASD</td>
<td>Random assignment; 5 in experimental group, 5 in control group</td>
<td>CAI software: FEFA program used to teach emotions</td>
<td>Emotion recognition</td>
<td>NA</td>
<td>Multiple regression ANCOVA</td>
<td>Spearman’s coefficient; Large effect size: Greater than .50</td>
<td>Experimental group increased from 31.6 to 43 on face test which was significant. No associated activation in FG was observed</td>
</tr>
</tbody>
</table>

*(table continues)*
Table A2 (continued).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Random Assignment</th>
<th>IV</th>
<th>DV</th>
<th>IOA</th>
<th>Statistical Analysis</th>
<th>Effect Size</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heimann, Nelson, Tjus, &amp; Gillberg (1995)</td>
<td>30 total, ages 6 to 13 years; 3 groups: 11 with ASD, 9 mixed disabilities, 10 non-disabled</td>
<td>Group assignment based on disability criteria</td>
<td>CAI software: Alpha and interactive language and reading program</td>
<td>Reading and communication skills</td>
<td>NA</td>
<td>Paired t-test</td>
<td>NA</td>
<td>All participants showed significant increase in vocabulary. Non-disabled participants generalized reading skills</td>
</tr>
<tr>
<td>Hopkins et al. (2011)</td>
<td>49 total, ages 6 to 15 years, with ASD, LFA and HFA</td>
<td>Randomly assigned to experimental or control group</td>
<td>CAI software: Face Say an emotion training software with interactive avatars</td>
<td>Emotion recognition</td>
<td>NA</td>
<td>Emotion recognition was analyzed using ANCOVA</td>
<td>Cohen’s d Social Skills Rating System Composite: 1.34 (LFA) 0.29 (HFA)</td>
<td>Social Skills Observation: 0.81 (LFA) 1.34 (HFA)</td>
</tr>
</tbody>
</table>

(table continues)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Random Assignment</th>
<th>IV</th>
<th>DV</th>
<th>IOA</th>
<th>Statistical Analysis</th>
<th>Effect Size</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>LaCava, Golan, Baron-Cohen, &amp; Myles (2007)</td>
<td>8 total, ages 8 to 11 years, with ASD, high functioning</td>
<td>No</td>
<td>CAI software: Mind Reading designed to teach emotion recognition skills</td>
<td>Emotion recognition based on face and voice</td>
<td>NA</td>
<td>Nonparametric: Wilcoxon signed-rank test</td>
<td>NA</td>
<td>Significant improvement on Cambridge mind reading face and voice tests</td>
</tr>
<tr>
<td>Moore, Cheng, McGrath, &amp; Powell (2005)</td>
<td>34 total, ages 7.8 to 16 years, with ASD.</td>
<td>No</td>
<td>CAI software: emotion training program that uses avatars</td>
<td>Emotion recognition</td>
<td>NA</td>
<td>One-tailed significance test</td>
<td>NA</td>
<td>30/34 participants answered questions at a level significantly better than expected by chance at the conclusion of the training</td>
</tr>
<tr>
<td>Silver &amp; Oakes (2001)</td>
<td>24 total, ages 10-18, with ASD, 2 groups of 12</td>
<td>Random pairwise group assignment based on age, gender, class</td>
<td>CAI software: Emotion Recognition Trainer includes digital photographs of faces or</td>
<td>Emotion recognition</td>
<td>NA</td>
<td>ANOVA</td>
<td>(F = 4.785, p=.041)</td>
<td>Experimental group did significantly better on strange stories and emotion recognition</td>
</tr>
</tbody>
</table>
Table A2 (continued).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Random Assignment</th>
<th>IV</th>
<th>DV</th>
<th>IOA</th>
<th>Statistical Analysis</th>
<th>Effect Size</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaka et al. (2010)</td>
<td>79 total, ages 8 to 12 years, with ASD. 2 groups</td>
<td>Random assignment to active treatment group and a waitlist control group. Randomization was stratified by mental age and IQ/Verbal skills</td>
<td>CAI software: Let's Face It! Designed to teach face recognition skills</td>
<td>Facial identification</td>
<td>NA</td>
<td>ANOVA</td>
<td>NA</td>
<td>Participants in treatment group showed significant improvement in mouth and eye recognition</td>
</tr>
<tr>
<td>Travers et al. (2011)</td>
<td>17 total, ages 3 to 6 years, with ASD</td>
<td>2 counterbalanced groups</td>
<td>CAI software: Letter recognition program designed for the study</td>
<td>Alphabetic skills, attention, undesirable behavior</td>
<td>82.1% for attention, 88.4% for undesirable behavior</td>
<td>ANOVA</td>
<td>Paired sample t-test</td>
<td>NA</td>
</tr>
</tbody>
</table>
No significant difference between groups on maintenance, attention to task, or level of undesirable behavior.

Table A2 (continued).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Random Assignment</th>
<th>IV</th>
<th>DV</th>
<th>IOA</th>
<th>Statistical Analysis</th>
<th>Effect Size</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whalen et al. (2010)</td>
<td>47 total, ages 3 to 6 years, with ASD.</td>
<td>2 groups: 22 in treatment group; 25 in control group</td>
<td>CAI software: TeachTown: Basic</td>
<td>Language, cognitive, auditory processing, and social skills</td>
<td>NA</td>
<td>ANOVA</td>
<td>NA</td>
<td>Pre-school students in the treatment group had significant increase in PPVT scores. No significant differences in K-1 scores for any measures.</td>
</tr>
<tr>
<td>Williams, Wrights, Callaghan,</td>
<td>8 total, ages 3 to 5 years, with ASD</td>
<td>Randomly assigned to</td>
<td>CAI software: books scanned into</td>
<td>Time spent attention to task</td>
<td>75-100% on timed activities</td>
<td>t-test</td>
<td>NA</td>
<td>Participants in the CAI group spend more</td>
</tr>
<tr>
<td>&amp; Coughlan (2002)</td>
<td>2 groups: Book or CAI</td>
<td>PowerPoint format with sounds and interaction embedded</td>
<td>Words spoken</td>
<td>time engaged and displayed more spoken words compared to book group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>--------------------------------------------------------</td>
<td>--------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Book: interactive books with sounds and pictures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A3

**Strength Rating Guidelines**

<table>
<thead>
<tr>
<th>Strength</th>
<th>Group Design Research</th>
<th>Single Subject Experimental Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>• All primary quality indicators;</td>
<td>• All primary quality indicators;</td>
</tr>
<tr>
<td></td>
<td>• 4&gt;Secondary quality indicators</td>
<td>• 3&gt; Secondary quality indicators</td>
</tr>
<tr>
<td>Adequate</td>
<td>• 4&gt; Primary quality indicators;</td>
<td>• 4&gt; Primary quality indicators;</td>
</tr>
<tr>
<td></td>
<td>• No unacceptable on primary quality indicators;</td>
<td>• No unacceptable on primary quality indicators;</td>
</tr>
<tr>
<td></td>
<td>• 2&gt; Secondary quality indicators</td>
<td>• 2&gt; Secondary quality indicators</td>
</tr>
<tr>
<td>Weak</td>
<td>• &lt;4 Primary quality indicators;</td>
<td>• &lt;4 Primary quality indicators;</td>
</tr>
<tr>
<td></td>
<td>• &lt;2 Secondary quality indicators</td>
<td>• &lt;2 Secondary quality indicators</td>
</tr>
</tbody>
</table>
## Group Design Research Evaluation Guidelines

<table>
<thead>
<tr>
<th>Primary Criteria:</th>
<th>Secondary Criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant characteristics</td>
<td>1. Random assignment</td>
</tr>
<tr>
<td>2. Independent variable</td>
<td>2. Inter-observer agreement</td>
</tr>
<tr>
<td>4. Dependent variable</td>
<td>4. Fidelity</td>
</tr>
<tr>
<td>5. Link between research question and data analysis</td>
<td>5. Attrition</td>
</tr>
<tr>
<td>6. Use of statistical tests</td>
<td>6. Generalization and/or maintenance</td>
</tr>
<tr>
<td></td>
<td>7. Effect size:</td>
</tr>
<tr>
<td></td>
<td>8. Social validity</td>
</tr>
</tbody>
</table>
Table A5

*Single Subject Research Design Evaluation Guidelines*

<table>
<thead>
<tr>
<th>Primary Criteria</th>
<th>Secondary Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant characteristics</td>
<td>1. Inter-observer agreement</td>
</tr>
<tr>
<td>2. Independent variable</td>
<td>2. Kappa</td>
</tr>
<tr>
<td>4. Dependent variable</td>
<td>4. Fidelity</td>
</tr>
<tr>
<td>5. Visual analysis</td>
<td>5. Generalization and/or maintenance</td>
</tr>
<tr>
<td>Authors</td>
<td>Primary QI</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Bernard-Opitz, Sriram, &amp; Nakhoda-Sapan (2001)</td>
<td>5</td>
</tr>
<tr>
<td>Bolte, et al. (2006)</td>
<td>4</td>
</tr>
<tr>
<td>Bosseler &amp; Massaro (2003)</td>
<td>5</td>
</tr>
<tr>
<td>Heimann, Nelson, Tjus, &amp; Gillberg (1995)</td>
<td>5</td>
</tr>
<tr>
<td>Hopkins, et al. (2011)</td>
<td>6</td>
</tr>
<tr>
<td>LaCava, Golan, Baron-Cohen, &amp; Myles (2007)</td>
<td>4</td>
</tr>
<tr>
<td>Moore, Cheng, McGrath, &amp; Powell (2005)</td>
<td>3</td>
</tr>
<tr>
<td>Silver &amp; Oakes (2001)</td>
<td>5</td>
</tr>
<tr>
<td>Tanaka, et al. (2010)</td>
<td>5</td>
</tr>
<tr>
<td>Travers, et al. (2011)</td>
<td>6</td>
</tr>
<tr>
<td>Whalen, et al. (2010)</td>
<td>5</td>
</tr>
<tr>
<td>Williams, Wrights, Callaghan, &amp; Coughlan (2002)</td>
<td>6</td>
</tr>
</tbody>
</table>
### Table A7

**Evaluation of Single Subject Research Design Studies**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Primary QI</th>
<th>Secondary QI</th>
<th>Strength Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coleman-Martin, Heller, Cihak, &amp; Irvine (2005)</td>
<td>5</td>
<td>3</td>
<td>Strong</td>
</tr>
<tr>
<td>Hetzroni &amp; Tannous (2004)</td>
<td>4</td>
<td>3</td>
<td>Adequate</td>
</tr>
<tr>
<td>Hetzroni &amp; Shalem (2005)</td>
<td>4</td>
<td>3</td>
<td>Adequate</td>
</tr>
<tr>
<td>Kodak, Fisher, Clements, &amp; Bouxsein (2011)</td>
<td>5</td>
<td>3</td>
<td>Adequate</td>
</tr>
<tr>
<td>LaCava, Ranking, Mahlios, Cook &amp; Simpson (2010)</td>
<td>5</td>
<td>3</td>
<td>Adequate</td>
</tr>
<tr>
<td>Mancil, Haydon, &amp; Whitby, (2009)</td>
<td>5</td>
<td>4</td>
<td>Adequate</td>
</tr>
<tr>
<td>Whitcomb, Bass, &amp; Luiselli (2011)</td>
<td>4</td>
<td>3</td>
<td>Adequate</td>
</tr>
<tr>
<td>Yaw, et al. (2011)</td>
<td>5</td>
<td>3</td>
<td>Adequate</td>
</tr>
</tbody>
</table>
Table A8

Formula for Determining Evidence-Based Practice Status Developed by Reichow et al. (2008; 2011) and Calculation of Formula for Present Review

<table>
<thead>
<tr>
<th>Total Score (based on Formula 2.1)</th>
<th>Established as EBP or Promising?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2<em>30) + (9</em>15) + (1<em>4) + (8</em>2) = 215</td>
<td>Established as EBP</td>
</tr>
</tbody>
</table>

Formula 2.1

\[(\text{Group}_S \times 30) + (\text{Group}_A \times 15) + (\text{SSEDS} \times 4) + (\text{SSEDA} \times 2) = Z\]
APPENDIX B

DETAILED METHODOLOGY
Method

Participants and Setting

The participants for this study included four children who had a determination of autism by their local education agency (LEA) or a licensed psychologist using the DSM-IV criteria (American Psychiatric Association, 2000) (see Table 1). Participants were recruited if they met the following inclusionary criteria: (a) were between the chronological age of 5 and 21 years; (b) displayed verbal skills including expressing wants and needs in complete sentences (5-7 word phrases); (c) passed the pretest measures of basic emotion recognition skills derived from Levels 1 and 2 from the mind-reading curriculum developed by Howlin et al. (1999) with 75% accuracy (d) scored 35% or lower on selected items on Level 3 of the Howlin et al. (1999) curriculum. No participant was excluded from the study based on race, ethnicity, gender, and/or socio-economic status as long as the inclusion criteria were met.

Recruitment flyers with a description of participation requirements and inclusion criteria were sent to a variety of locations including an autism intervention center, offices of special education directors of three school districts and 6 autism therapy organizations. When parents showed interest in participating in the study, an initial screening was conducted via telephone to establish the pre-requisite skills of their child. Next, I individually met with the candidates and their parents for approximately 30 minutes during which time she interacted with the potential participant in order to gain a better understanding of his/her ability. Additionally I conducted a pre-screening assessment using levels 1, 2 and 3 from the Howlin et al. (1999) curriculum (i.e., Teaching Children with Autism to Mind-Read) as stated in the inclusion criteria.
The procedures for recruiting participants and conducting the study were approved by the university’s Institutional Review Board (IRB). The letter for obtaining informed consent was provided to the parents following recruitment. The informed consent letter included the purpose and procedures of the study, foreseeable risks and benefits, and participant’s rights. The procedures for the study commenced after receipt of the signed informed consent letter from the parents. Four children participated in the study.

Evan was an 11-year old boy of Asian-American descent. He lived in a single-family household with his biological parents and younger sister. Initial screening and parent report indicated that he primarily communicated using complete sentences (e.g., “I need to get more points,” “When can we be done?”) with some repetitive phrases and echolalia. Evan was evaluated by a licensed psychologist in 2013. The results of the Wechsler Intelligence Scale for Children – IV (2003) indicated that his full-scale IQ score of 66 was significantly lower than his academic achievement scores revealed by the Wechsler Individual Achievement Test – III (2009): math problem solving - 96; word reading - 91; pseudoword decoding - 92; numerical operations - 98; spelling - 100. Possible explanation for the relatively low IQ score as noted by the licensed psychologist included inattention and distractibility during test-taking, obsessive thoughts, and difficulty with abstract reasoning. The Childhood Autism Rating Scale – II (2010) indicated a standard score of 38 which qualified Evan with autism in the mild-moderate range. During the initial emotion recognition screening (Howlin et al., 1999), Evan scored 87.5% and 100% respectively on Levels 1 and 2 and 32% on Level 3.

Ashley was a 5-year old girl of Asian-American and Caucasian descent. She lived in a single-family household with her biological parents and younger sister. Initial screening and
parent report indicated that she primarily communicated using complete sentences (e.g., “I want to open the box,” “Can I have the apple?”) with some repetitive and echolalic phrases. Ashley’s cognitive or intellectual functioning was evaluated by the LEA in 2012 using the Kaufman Assessment Battery for Children – II Nonverbal (2004) which revealed mental processing of 74 (i.e., below average) and learning/long-term retrieval of 97 (i.e., average range). Academic Achievement was assessed using the Developmental Assessment of Young Children (DAYC, 1998). The test revealed an overall standard developmental score of 77 (i.e., below average range) and a standard score of 89 on adaptive behavior (i.e., average range). Additionally she was evaluated for autism eligibility using the Childhood Autism Rating Scale – II (2010). She received a standardized score of 40 suggesting mild to moderate autism. During initial screening process Ashley scored 75% on both Levels 1 and 2 of the emotion recognition tasks and 27% on Level 3, making her eligible to participate in the study.

Trevor was a 9-year old boy of Caucasian-American descent. He lived in a single-family household with his biological parents and younger sister. Initial screening and parent report indicated that he primarily communicated using complete sentences (e.g., “Let’s take a picture,” “Can I have more goldfish crackers, please?”) with some echolalic and scripted speech. Trevor’s cognitive and intellectual functioning was evaluated by the LEA in 2013 using the Leiter International Performance Scales – Revised (LIPS-R, 1997). His full scale IQ was a 73 which is in the below average range with a strength in visual processing score at 81 and performance in the average range on the subtest “Classification”. The examiner cautioned that scores may be influenced by his level of activity and inattention during testing. His academic achievement was assessed using the Kaufman Test of Educational Achievement – II (2004), which indicated varied
scores, for example, a standard score of 105 on letter and word recognition (i.e., average range) and 71 on math concepts and applications (i.e., low range). His adaptive behavior was assessed using the Vineland Adaptive Behavior Scale – II (2005), which showed a composite score of 65 (i.e., low range). Additionally, Trevor was assessed using the Gilliam’s Autism Rating Scale (GARS, 2006) and had an overall Autism Index of 94 which suggests a very likely probability of autism. During the initial screening on the emotion recognition tasks, he scored 87.5% and 100% respectively on Levels 1 and 2 and 23% on Level 3, qualifying him for the study.

Pranav was a 5-year old boy of Indian-American descent. He lived in a single-family household with his biological parents and older sister. Initial screening and parent report indicated that he primarily communicated using complete sentences (e.g., “Can I have a candy?” or “Sister wants to play with you,”) indicating some echolalic speech. Pranav’s cognitive and intellectual functioning was evaluated in 2014 by the LEA using the Kaufman Assessment Battery for Children-II Nonverbal Index (2004) which indicated a standard score of 82 (i.e., slightly below average). The Kaufman Test of Educational Achievement – II (2004) was used to evaluate his academic achievement, indicating a standard score of 143 on letter and word recognition (i.e., superior range), 97 on math concepts and applications (i.e., average range), and 52 on written expression (i.e., significantly below average). Additionally, the Gilliam’s Autism Rating Scale (GARS, 2006) was completed and revealed a standardized score of 122 (i.e., very likely probability of autism). During initial screening, Pranav scored 75% on both Levels 1 and 2 on the emotion recognition tasks but only 5% on Level 3, making him eligible to participate in the study.
The study was conducted in the homes of individual participants. Baseline and intervention procedures were conducted in a specified location at a computer desk or other area designated by the parent as quiet and free from distractions and suitable for the study. A built-in video recording device on the laptop was used to record sessions for baseline, generalization probes, and CAI phases of the study and a handheld recording device was used to record sessions for the generalization phase.
<table>
<thead>
<tr>
<th>Participant</th>
<th>Cognitive/Intellectual</th>
<th>Academic Achievement</th>
<th>Adaptive Behavior</th>
<th>ASD Determination</th>
<th>Language/Communication</th>
<th>Level 1 &amp; 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trevor (m) – 9 Caucasian American Evaluation: 2013</td>
<td>LIPS-R: 73</td>
<td>KTEA-II:</td>
<td>VABS-II:</td>
<td>GARS: 94 (very likely probability)</td>
<td>Complete Sentences</td>
<td>87.5%; 23%</td>
<td></td>
</tr>
<tr>
<td>Pranav (m) – 5 Indian American Evaluation: 2014</td>
<td>KABC-II Nonverbal Index: 82</td>
<td>KTEA-II:</td>
<td>SIB-R:</td>
<td>GARS: 122 (very likely probability)</td>
<td>Complete Sentences</td>
<td>75%; 5%</td>
<td></td>
</tr>
</tbody>
</table>
Measurement of Dependent Variables

Guidelines for single subject research methodology by Horner et al. (2005) were utilized to operationally define and frequently measure the dependent variables throughout the study. Additionally, measurements were assessed for reliability by a secondary observer and behavior change was evaluated for both empirical and social significance to the participants as well as the field. The following dependent variables were measured for the study.

Mind-reading skills. For the purpose of this study, mind-reading was operationally defined as the ability to correctly identify the emotion of a person based on environmental cues in a picture when the facial expression was masked and no voice clues were provided. During baseline and intervention assessment, mind-reading skills were measured as a mutually exclusive correct or incorrect response. A correct response was recorded when a participant used the cursor to click the icon representing the emotion (shown through a line drawn face) that matched the social situation displayed in the picture presented on the computer screen (see Figure 1).
Figure B-1. Example of Teaching Mind-Reading Skills to Children with Autism software for CAI instruction. Includes prompting (e.g. red X) and feedback.

During generalization, mind-reading was measured when a participant, at the first attempt, correctly verbalized the emotion that the parent or sibling felt in a specific role-played scenario (described below). Non-examples of mind-reading during baseline and intervention assessment included a participant choosing an emotion that did not correctly match the social situation displayed, displaying or verbally stating an emotion but not clicking on the computer
icon, or responding to distractors (e.g., items in the picture like "his hands are up"). Mind-
reading was the primary dependent variable for this study and used to make decisions
regarding phase changes.

Data on mind-reading skills for baseline and intervention were recorded using a
frequency-based data collection system embedded in the computerized software that indicated
the number of correct responses for each assessment session. The software was programmed
to randomly select any 10 scenarios from a pool of 22 social problem scenarios at each
assessment (see Figure B-2). Assessment data were collected through this software program
immediately following the instructional session for baseline and intervention.

A_BLI_1-28's Results

Question Question 18:
  Answer #1: happy - incorrect : 19 second(s) elapsed

Question Question 22:
  Answer #1: happy - incorrect : 42 second(s) elapsed

Question Question 6:
  Answer #0: happy - correct : 16 second(s) elapsed

Question Question 3:
  Answer #1: sad - incorrect : 16 second(s) elapsed

Question Question 12:
  Answer #1: angry - incorrect : 22 second(s) elapsed

Question Question 5:
  Answer #1: angry - incorrect : 14 second(s) elapsed

Question Question 10:
  Answer #0: happy - correct : 12 second(s) elapsed

Question Question 11:
  Answer #1: angry - incorrect : 14 second(s) elapsed

Question Question 13:
  Answer #1: happy - incorrect : 12 second(s) elapsed

Question Question 19:
  Answer #1: sad - incorrect : 18 second(s) elapsed

Figure B-2. Sample log of the internal data collection system recorded by the CAI software.
Task engagement. This variable was also measured to determine whether or not task engagement of participants varied across experimental conditions involving different instructional methods to teach mind-reading skills. Task engagement was defined as the participant (a) sitting with his/her shoulders, head, and eyes oriented towards the source of instruction (i.e., person or the computer monitor); (b) having hands on the lap, table top, or the computer mouse as appropriate; and (c) verbally responding or clicking the icon on the computer screen within 10 seconds of presentation of the prompt or question. Non-examples include (a) making verbal sounds not related to the answer; (b) looking away or orienting body away from the source of instruction; (c) displaying stereotypic and repetitive movements with hands or mouse; or (d) failing to respond to a question or prompt within 10 seconds. Task engagement for baseline and CAI was measured using 30-s partial interval recording system (see sample data sheet in Figure B-3).
Direct Observation Data Sheet to Record Student Task Engagement

Data Collector: ___________________  Participant Code: ___________________

Date: ___________________

Instruction: For each of these time intervals, if the target student is engaged in the task, note a “+” and if not engaged, note a “-”

<table>
<thead>
<tr>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
<th>Interval</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (30s)</td>
<td>11.(30s)</td>
<td>21.(30s)</td>
<td>31.(30s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. (60s)</td>
<td>12.(60s)</td>
<td>22.(60s)</td>
<td>32.(60s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. (30s)</td>
<td>13.(30s)</td>
<td>23.(30s)</td>
<td>33.(30s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. (60s)</td>
<td>14.(60s)</td>
<td>24.(60s)</td>
<td>34.(60s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. (30s)</td>
<td>15.(30s)</td>
<td>25.(30s)</td>
<td>35.(30s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. (60s)</td>
<td>16.(60s)</td>
<td>26.(60s)</td>
<td>36.(60s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. (30s)</td>
<td>17.(30s)</td>
<td>27.(30s)</td>
<td>37.(30s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. (60s)</td>
<td>18.(60s)</td>
<td>28.(60s)</td>
<td>38.(60s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. (30s)</td>
<td>19.(30s)</td>
<td>29.(30s)</td>
<td>39.(30s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. (60s)</td>
<td>20.(60s)</td>
<td>30.(60s)</td>
<td>40.(60s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 minutes over</td>
<td>10 minutes over</td>
<td>15 minutes over</td>
<td>20 minutes over</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure B-3.* Direct observation recording form for task engagement during baseline and CAI phase.
A video camera built into the laptop was used to record each session which I reviewed to obtain percentage of task engagement. These data were subsequently reviewed and coded by the secondary observer for reliability.

Direct observation. I served as the primary observer for this study (given that a software program was used to deliver the intervention). The secondary observer was a doctoral candidate who was trained on the data collection procedures and operational definitions of the target behaviors prior to initiating the study. Both observers watched videos of individuals (prior to the onset of the study) using CAI in order to practice coding task engagement. Observers were trained to record the dependent variables using the operational definitions and procedures noted above. Data collection for the study was initiated only after interobserver agreement reached 90% accuracy over three practice sessions.

Interobserver agreement. Interobserver agreement for mind-reading was established by having the secondary observer review the data logs stored in the CAI software program and record the percentage of correct answers for each assessment session per participant (i.e., 100% of the logs were assessed for reliability). These percentages were then compared to the data recorded by the primary observer (Hutcherson, Langone, Ayers, & Clees, 2004).

IOA for task engagement was established by the primary observer reviewing the videos of baseline and intervention sessions for each participant and recording the duration of task engagement onto a 30 second partial interval recording system. The videos were then reviewed by the secondary observer using the same operational definitions and recording procedures. IOA for task engagement was collected for 34% of sessions across baseline and intervention phases.
IOA data for both dependent variables was calculated by dividing the number of rater agreements plus disagreements and multiplying the results by 100 to get a percentage of agreement for each session (for each dependent variable). An agreement for observance of mind-reading skills was defined as both observers documenting the same answer result (i.e., correct or incorrect) based on computer logs. A disagreement for mind-reading was defined as both observers recording a different answer result from the computer logs. An agreement for task engagement was defined as an occurrence when both observers agreed that on task behavior occurred (+) or did not occur (-) during a specified interval or period of time. A disagreement for task engagement was defined as one observer indicating the occurrence or non-occurrence of a target behavior but the other observer did not record the same event. The outcomes of IOA for mind-reading skills are as follows: Evan (M = 100); Ashley (M = 98.75; range = 90 - 100); Trevor (M = 100); and Pranav (M = 99.7; range = 90-100). The outcomes for IOA of task engagement remained high as well at: Evan (M = 98.96; range = 93.75 - 100), Ashley (M = 97; range = 85 - 100), Trevor (M = 95.83; range = 75 - 100), and Pranav (M = 98.5; range = 90.9 - 100).

IOA for generalization sessions was established by the primary observer reviewing the videos of generalization sessions and recording the percentage of correct responses to the social scenarios that were acted out by family members. These videos were then reviewed by the secondary observer and the percentage of correct responses was recorded. The secondary observer reviewed 40% of generalization sessions for IOA. IOA for generalization sessions was 100% for all participants and all sessions.
At the conclusion of the study, Cohen’s kappa was used to compute overall data reliability for mind-reading and task engagement. Cohen’s kappa is a necessary step in calculating IOA because it adjusts the observed proportion of agreement by taking into account the amount of agreement that can be expected by chance (Cohen, 1960; Horner & Kratochwill, 2012). This was accomplished by counting the number of agreements between the two observers and the number of expected agreements and applying the following formula:

\[ K = \frac{(P_o - P_e)}{(100 - P_e)} \]

where \( P_o \) = the proportion of agreements between observers and \( P_e \) = the proportion of agreements expected by chance (Cohen, 1968). The interobserver agreement across baseline and intervention phases for mind-reading skills was .98 indicating strong internal reliability. Additionally, the IOA for task engagement was .93 which also indicates strong internal reliability and meets the guidelines specified by Landis and Koch (1977) (greater than .80).

Social validity. Social validity was assessed using two methods. The first method was a questionnaire which the parents of participants completed. The questionnaire was a self-report measure regarding parental perspective on CAI and included five questions with a 5-point Likert-type scale indicating a range of response from strongly disagree to strongly agree. Questions assessed parental perceptions regarding effectiveness and acceptability of computer assisted instruction.

Results of social validity assessment are presented in Table B.2. The questionnaire completed by the mothers of all four participants indicated agreement to strong agreement that the CAI program was effective for teaching their child to predict the emotions of others and their child enjoyed using the computer. Additionally, three parents felt that the CAI added
a critical level of instruction that their child did not obtain from teacher-led instruction.

Perhaps most importantly, all four parents indicated agreement or strong agreement that their child was better able to predict the emotions of others in daily life as a result of participating in the CAI intervention and that they were satisfied with the outcome of the study.

Table B-2

Social Validity Results from Parent Questionnaire

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel like the computer assisted instruction (CAI) program was effective for teaching my child to predict the emotions of others.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>My child enjoyed using the computer to learn mind reading skills.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>I feel like computer assisted instruction adds a critical level of instruction that my child does not obtain from teacher led instruction.</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>50%</td>
<td>25%</td>
</tr>
<tr>
<td>My child is better able to predict the emotions of others in daily life as a result of participating in the CAI intervention.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>I am satisfied with the outcome of the study.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>75%</td>
</tr>
</tbody>
</table>

The second social validity measure was collected by interviewing the participants about their perspective on computer assisted instruction. The participant questionnaire used an open
ended response format to assess their enjoyment level, overall satisfaction with CAI and preference of instructional delivery method (see Table 3).

Table B-3

*Social Validity Questionnaire to be completed by Participants*

<table>
<thead>
<tr>
<th>Participant Social Validity Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you enjoy using the computer to learn about emotions?</td>
</tr>
<tr>
<td>What did you learn about different emotions?</td>
</tr>
<tr>
<td>Do you think using the computer to learn about emotions will help you recognize other’s emotions in your daily life?</td>
</tr>
<tr>
<td>Do you think the computer was easy to use?</td>
</tr>
<tr>
<td>Would you like to use the computer to learn about other things?</td>
</tr>
</tbody>
</table>

All four participants completed the interview and indicated that they enjoyed using the computer to learn about emotions. Additionally, they all indicated that they thought the CAI program was easy to use and three of the four participants indicated that they would like to use the computer to learn about additional concepts.

Materials

Social narratives. During baseline instruction, social narratives involving social situations and related emotions were read to the participant. Social narratives were created using PowerPoint and printed on white paper to present content in a booklet format. A booklet format was used to control for variation in the method of presentation of the content on mind-reading and to present the information in a format commonly used for ToM instruction (i.e.,
paper based social narratives). In other words, the amount and type of content did not change across the experimental phases, just the method of instructional delivery was varied. Each social narrative included pictures and text of social problem situations and was approximately 5 to 10 pages with 1-2 sentences of description per page (see example in Figure B-4).

**Figure B-4.** Sample social narrative for baseline procedures.

Original mind-reading curriculum. The materials for the intervention were adapted from the mind-reading curriculum, Teaching Children with Autism to Mind-Read, originally developed by Howlin and her colleagues (1999). This curriculum comprises five levels of instruction on theory of mind (ToM) skills beginning with the least complex Level 1 focused on teaching the foundation skill of identifying emotions in photographs. Level 2 raises the difficulty of the task requiring participants to identify emotions in schematic drawings; Levels 3, 4 and 5 significantly increase the complexity of the task requiring individuals to identify emotions based on social situations for perspective-taking, desire, and beliefs respectively. The
least complex levels (i.e., Levels 1 and 2, identifying emotions in pictures and schematic
drawings) have been demonstrated multiple times within the CAI literature (Golan & Baron-
Cohen, 2006; Hopkins et al., 2011; Lacava et al., 2010; Silver and Oakes, 2001) as not effective
for promoting generalized changes to real social situations (Golan & Baron-Cohen, 2006), these
two levels were used as prerequisite screening criteria for participation in this study. Level 3,
which includes more complicated perspective taking skills based on social situations, was the
focus for this study.

The goal of Level 3 was to teach individuals to predict how a character felt given the
emotional content of the scenario displayed in the picture. These emotions are referred to by
Howlin et al. (1999) as “situation-based” emotions. Schematic black and white drawings of a
variety of situations are included in this level. Although it is recognized that identifying
situation-based emotions (i.e., Level 3) is not the most complex level required for successfully
demonstrating ToM skills, it is the next level of complexity, which has not been separately
evaluated in the CAI research at this point in time, and which is a pre-requisite skill for
mastering Levels 4 and 5.

In each of the scenarios in Level 3, the character’s facial features (i.e., eyes, facial
expressions, lips) are masked, requiring the participant to determine the emotion based solely
on situational cues rather than facial or auditory cues. Each scenario contains a text description
of the scene and has an emotion question, for example: “The big dog is chasing Dan down the
road. How will Dan feel when the dog chases him?” Each picture is paired with schematic
emotion faces from which to select the correct response: happy, sad, angry, and afraid
(frightened).
Computer-based curriculum. The original mind-reading curriculum described above (Howlin et al., 1999 and referred to as the Mind-Reading Curriculum from this point onwards) was slightly modified and converted into an electronic format then saved on a university server to make it accessible from any computer or geographical location. The process of modification and conversion was steered by a collaborative team consisting of myself, the dissertation mentor, another doctoral student, and the university’s software programmer and his team. Levels 1 and 2 from the Mind-Reading Curriculum were modified by adding real photographs and clearer line drawings depicting the four emotions (i.e., happy, sad, angry, afraid). These were used for screening the participants and were part of the inclusion criteria.

For the intervention, 22 from the available 48 scenarios from the Level 3 were selected by the collaborative team referenced above based on technical requirements for the software development. A fewer number of scenarios were selected to ensure full attention of participants. The selected 22 scenarios represented each of the four emotions equitably [i.e., happy (5); sad (6); angry (6); afraid or scared (5)].

Content validation was conducted prior to the study to ensure that the scenarios incorporated into the software program from the original curriculum portrayed the emotion they were intended to represent. Content validators included three students in a doctoral program studying special education, one faculty member in the special education department, and a director of a local autism center. These individuals reviewed the 22 scenarios used for CAI and rated which emotion was portrayed. The content validation results indicated adequate agreement of the scenarios with a mean agreement of 92.5% (range = 77.35 – 100).
Once the scenarios received adequate content validation they were integrated into the computer-assisted instructional (CAI) software in two separate program files. The software utilized for instruction included audio cues and performance feedback for each correct or incorrect response option, whereas the one used for assessment purposes did not provide any audio feedback for the response selected by the participant. The audio clues were created and inserted into the software package by the collaborative team noted above and were not based on any interactive software developed by Baron-Cohen or his colleagues.

Materials for role plays. During the generalization phase of the study, siblings and parents were trained to act out or role play problem situations using a script which I developed based on previous work on teaching perspective-taking skills (Laugeson, Frankel, Gantman, Dillon, & Mogil, 2012) (see Figure B-5). Other materials used during this phase were items readily and normally available in a home setting (e.g., board game, craft supplies, food items, etc.).
Experimental Design and Procedures

A multiple baseline design across participants was used to evaluate the effectiveness of computer assisted instruction (CAI) on mind-reading skills and task engagement. The baseline phase was initiated concurrently for all participants but the phase change was staggered for one participant at a time based on their response to intervention. Once a stable rate of

---

**Situation 1**

Setting: Living room

Participants: 1 sibling and 1 parent

Activity: playing on the iPad/computer

Emotion: Angry

Action: The adult will be playing on the iPad or computer. The sibling will come and take the iPad or computer away. Parent will say “Hey!” Sibling will look at adult but not respond or give the iPad/computer back.

Assessment: Interventionist and participant will be sitting in the living room within 10 feet of the action. Interventionist will say “_________ just took the iPad from your mom/dad.” “How do you think that made your mom/dad feel?”

**Situation 2**

Setting: Kitchen

Participant: 1 sibling and 1 parent

Activity: baking

Emotion: Happy

Action: The parent will be in the kitchen and pull some cookies out of a sack and ask the sibling “Would you like a cookie?” The peer will answer “Yes, please!”

---

*Figure B-5.* Social situations script for role play during skill generalization.
responding was documented during baseline across a minimum of five consecutive sessions with no increasing trend, participant one moved from baseline to the intervention (CAI) phase. Decision making for phase change was based on the number of correct responses on mind-reading skills. This section will provide a detailed description of the experimental procedures, my background training, and procedural fidelity of the intervention. A graphic organizer of experimental procedures is provided in Figure B-6.
Figure B-6. Graphic organizer of research procedures
Baseline. Social narratives and paper-based products are common for teaching perspective-taking and mind-reading skills. Thus, these were used to collect data for baseline, the control condition that was compared with CAI (Gast, 2010). Baseline involved instructional sessions where I read a social narrative to each participant. The narrative involved a social situation requiring the participant to identify an emotion felt by one of the characters. Care was taken to ensure that the duration of each BL session was the same as session length as the intervention session (approximately 7 minutes) to control for threats to internal validity. At the end of each baseline session, participants were asked to take the computerized assessment which included 10 randomized scenarios on mind-reading skills with no feedback as discussed previously.

Computer Assisted Instruction (CAI). The intervention condition involved the participant sitting at a desk or table with a laptop computer. I started the computer program using the participant’s specified identification number. For each session participants first viewed 22 instructional trials which included embedded audio prompts (e.g., “look at the picture”), clues (e.g., “could the girl be upset?”), feedback for correct and incorrect responses (i.e., green check mark or red “x” respectively), and embedded reinforcement for correct responses (i.e., verbal praise, “Good Job!”). The 22 scenarios were displayed in systematic succession with scenarios for each emotion being displayed in groups (e.g., all scared scenarios were displayed in a row, then all happy scenarios were displayed, etc.). Once participants completed the instructional session, they were asked to take the computerized assessment on mind-reading skills which included 10 randomized questions with no feedback or prompts (identical to the baseline measurement).
Generalization probes. Generalization probes were conducted during baseline and intervention phases for all participants for two reasons: first, to assess the effectiveness of social narratives and CAI on mind-reading skills of participants based on the respective instructional methods; and second, to determine if participants would need to receive additional generalization training prior to beginning the generalization phase.

In order to complete generalization probes, I developed 16 social scenarios based on common social situations identified by parents’ of participants. Each scenario described a different social situation (e.g., “Your mom dropped her phone and it broke. How does that make your mom feel?”). For each generalization probe, the participant was required to listen to four scenarios and then determine how the character would feel in each situation (i.e., happy, sad, angry, scared). Participants were given 100% if they answered all four scenarios correctly, 75% for answering three of the four questions correctly, 50% for answering two correctly, etc. No feedback or reinforcement was provided for correct or incorrect answers. If participants had a mean performance of 50% or higher of answers correct on generalization probes during CAI, they did not receive any additional training prior to beginning the generalization phase. If a participant did not demonstrate a high level of mastery on generalization probes (M < 50%), additional training was provided prior to the onset of the generalization assessment phase.

Generalization training. Generalization training was needed only for one of the four participants (Pranav). Training involved me asking Pranav 16 social scenario questions (the same used during generalization probes; four scenarios for each emotion) and then providing feedback and an explanation for the correct answer (e.g., “your mom feels sad when she rips
her paper,” “your sister is happy when she gets to play outside with her friends”).  This was continued until the participant reached 100% accuracy on social scenarios for three consecutive scenarios.  Pranav reached 100% accuracy for three consecutive scenarios after two sessions of training and then began the generalization phase.

Generalization assessment. Generalization of mind-reading skills was conducted to evaluate the extent to which CAI was effective in teaching social problem-solving skills to participants as applied to daily life situations. The family members of each participant identified various social situations that occurred in a variety of settings (e.g., living room, kitchen, and dining room). I developed a short script (i.e., 3-4 sentences) for each of these situations (different from the scenarios used during generalization probes and training) and provided them to the parents and sibling of each of the participants with an explanation of the generalization procedure. They were subsequently trained to act out or role-play their parts (see Figure 4). Additionally, they were instructed on how to provide positive feedback to participants when they correctly (e.g., verbal praise) or incorrectly (e.g., error correction) identified an emotion based on the social situation displayed. The parents and sibling practiced acting out the social situation until they successfully completed all steps of the various social situation scripts with 90% accuracy (I recorded with a notation of “yes” or “no” for each role-play action).
During generalization, when the actors were ready to role-play, the participant was invited to join the family in the setting (i.e., living room, kitchen, or dining room). I ensured that the participant was in close proximity to the scene of the action (approximately 10 feet). For example, in one social situation, the parent was engaged with an electronic device in the living room when the sibling got up and grabbed the electronic device from the parent and started to use it herself. I asked the participant to describe how the parent felt in this situation and explain the reason for the answer. I recorded data on the accuracy of the participant’s
response. Then, reliability was assessed by the secondary observer after reviewing the same video recording.

**Procedural fidelity.** Procedural fidelity for baseline and the intervention (CAI) phase was conducted to ensure that the procedures were consistently implemented with accuracy and integrity (Gast, 2010). Because the intervention was delivered through a computerized software program, there was a low probability of error given that it was pre-programmed to deliver the same type of stimuli at each trial even if the order of presentation of stimulus items was randomized. Thus, procedural fidelity was assessed to ensure accuracy in using the correct log for each participant upon login, record the session on the laptop, and save the data on the software program when the session was completed. I used a checklist during each session to assess procedural fidelity during both baseline and intervention phases (see Table B-4 & B-5) which was then calculated by dividing the number of items on the checklist I correctly completed by the total number of items on the checklist and multiplying by 100. The outcome showed 100% accuracy for each participant on all the steps of the baseline and intervention. IOA on procedural fidelity was assessed by the parents of participants using the same checklist for an average of 36% of baseline and intervention sessions (Evan = 39%; Ashley = 37.5%; Trevor = 33%; Pranav = 37%). The IOA on procedural fidelity based on the checklist record by myself and the parent showed 100% agreement on the steps completed accurately.
Table B-4

_Procedural Fidelity Checklist for Baseline_

<table>
<thead>
<tr>
<th>Date/Session #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on laptop camera</td>
</tr>
<tr>
<td>State initial, BL or CAI, date</td>
</tr>
<tr>
<td>Show student task list: First, Next, Then</td>
</tr>
<tr>
<td>Provide two choices for social story</td>
</tr>
<tr>
<td>Student sits at computer desk facing researcher</td>
</tr>
<tr>
<td>Read 2 social stories</td>
</tr>
<tr>
<td>Student turns and faces laptop</td>
</tr>
<tr>
<td>Click on shortcut link labeled Mind Reading Software</td>
</tr>
<tr>
<td>Enter EUID and Password</td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled <strong>Assessment</strong></td>
</tr>
<tr>
<td>Enter Participant ID code</td>
</tr>
<tr>
<td>Monitor participant (must complete 10 scenarios)</td>
</tr>
<tr>
<td>Log out</td>
</tr>
<tr>
<td>Allow 5 minute break-free choice</td>
</tr>
</tbody>
</table>
Table B-5

*Procedural Fidelity for CAI*

<table>
<thead>
<tr>
<th>Date/Session #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on laptop camera</td>
</tr>
<tr>
<td>State initial, BL or CAI, date</td>
</tr>
<tr>
<td>Show student task list: First, Next, Then</td>
</tr>
<tr>
<td>Student sits at computer desk facing laptop</td>
</tr>
<tr>
<td>Click on shortcut link labeled <em>Mind Reading Software</em></td>
</tr>
<tr>
<td>Enter EUID and Password</td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled <em>Instruction</em></td>
</tr>
<tr>
<td>Enter Participant ID code</td>
</tr>
<tr>
<td>Monitor participant (must complete 22 scenarios)</td>
</tr>
<tr>
<td>Log out</td>
</tr>
<tr>
<td>Click interaction link (Earth) labeled <em>Assessment</em></td>
</tr>
<tr>
<td>Enter Participant ID code</td>
</tr>
<tr>
<td>Monitor participant (must complete 10 scenarios)</td>
</tr>
<tr>
<td>Log out</td>
</tr>
<tr>
<td>Allow 5 minute break-free choice</td>
</tr>
</tbody>
</table>
Investigator

At the time of the study, I was a fourth year doctoral candidate in special education with a focus on autism research. I had completed all the required course work for the Board Certified Behavior Analyst (BCBA®) program along with 72 doctoral credit hours including various research methods courses. Additionally, I had six years of experience teaching students with autism in the public school setting.
APPENDIX C

UABRIDGED RESULTS
Results

Visual analysis, commonly used in single subject research, was the primary method for analyzing data to determine a functional relation between the independent and dependent variables. Specifically, data were concurrently coded and graphed to analyze the stability, level and trend changes, variability, overlapping data across adjacent phases, and similarity of pattern across similar phases (Horner & Kratochwill, 2012). Additionally, effect size was computed using Cohen’s $d$ (1988) in order to determine the magnitude of effect of the intervention on the dependent variables. Results are described in relation to the specific research questions. Additionally, line graphs of results are provided in Figures C-7 and C-8.

Effectiveness of Computer Assisted Instruction on Mind-Reading Skills

The first research question assessed the existence of a functional relation between CAI and mind-reading skills. Baseline data for Evan indicated a mean of 31.66% correct answers on mind-reading skills, however, a decreasing trend was observed as the phase progressed. Given the stability in the data pattern, intervention was initiated after six consecutive sessions. Data showed an immediate effect, from an average of 13% in the last three sessions to an average of 73% for the first three sessions of intervention. Evan’s mean level of performance was noted at 82.86% during the intervention marked by ceiling performance (i.e., 100%) on the seventh session. The stability envelope indicated that 86% of the intervention data were within 20% of the median indicating stability in the pattern of learning during intervention. The percentage of non-overlapping data was acceptable at 71.4%. Overall, Evan’s data indicate an increase in mind-reading skills as a function of the intervention.
Following an increasing trend in response to intervention for Evan and a concurrent stable baseline for Ashley, intervention was initiated with her. Ashley’s baseline data show a mean level of 26% correct responses. The stability envelope indicated that 90% of data points were within 20% of the median. She also demonstrated an immediate effect in response to the intervention, moving from an average of 27% during the last 3 sessions of baseline to 83% average for the first 3 data points during intervention. The mean level change during CAI was at 90% with 83% of data within the stability envelope. During the intervention phase she demonstrated an accelerating trend and reached the ceiling of 100% correct answers after 4 sessions. There was no overlap in data across adjacent phases. Overall, data for Ashley indicate a clear functional relationship between CAI and mind-reading skills.

Once Ashley reached a stable level of performance during intervention and Trevor demonstrated a stable-flat baseline with 93.75% of data falling within the stability envelope, intervention was initiated with Trevor. His rate of responding increased from a mean baseline level of 26.25% to 84.55% during intervention. He too demonstrated an immediate response to intervention with an absolute change from 23% to 76.66%. He reached the ceiling level of 100% correct answers on the fourth session, however, his data did show some variability across the 11 intervention sessions with only 63.64% of data falling within the stability envelope. This was likely caused by a decrease in performance on the sixth session due to illness. The percentage of non-overlapping data was significant at 100%. Additionally, given the increasing trend in the data pattern starting the seventh session, overall results indicate that the CAI had a positive, functional impact on Trevor’s mind-reading skills.
When Trevor’s performance returned to the ceiling level at 100% (on intervention session 10) while Pranav’s data demonstrated a stable pattern of responding during baseline (i.e., stability envelope = 83.3%), intervention was started with him. His data appear to show a variable pattern, however, other than two sessions of data overlap with baseline (possibly caused by the initial slow response to intervention in the first session and difficulty maintaining attention in the seventh session), Pranav’s mean score increased significantly from 27.91% during baseline to 80% during intervention. Additionally, the absolute change during intervention was 40% to 100% indicating an increasing trend. The percentage of non-overlapping data was also within the moderate range at 85.71%. These results appear to indicate a functional relation between CAI and mind-reading skills for Pranav as well.
Figure C-7. Percent correct on mind-reading skills for all participants.
Effect of CAI on Rate of Task Engagement

Although task engagement was a secondary dependent variable, the second research question assessed the extent to which it was higher for participants during baseline (i.e., paper-based social narratives) vs. CAI. The results for task engagement showed higher variability across participants than what was observed for mind-reading skills (i.e., the primary dependent variable). Three of four participants showed higher levels of task engagement when CAI was used in comparison to baseline even though data showed overlap across adjacent phases for one of these participants (i.e., Trevor). Evan displayed a variable rate of task engagement during both baseline ($M = 13$; range $= 0 – 29$) and intervention ($M = 50$; range $= 30 – 68$), however, the mean level was higher during intervention with 86% of the data falling within the stability envelope. He displayed a relative change of $68 – 54$ indicating a decelerating trend yet 100% non-overlapping data indicating that CAI may have had a positive impact on task engagement compared to baseline.

Ashley did not display a significant change in task engagement between the two experimental conditions based on visual analysis of the data. She showed a mean of 56.9% during baseline which increased slightly to 73.8% during intervention. Her data were variable during both conditions and 87.5% of the data overlapped with CAI demonstrating that task engagement was not impacted by the type of instructional method used with Ashley.

Trevor displayed a very stable (93.75% within the stability envelope) mean level of 2.4% task engagement during baseline. He increased to a mean of 30% during CAI however, data demonstrated a large amount of variability with only 27% of the data falling within the stability envelope. He also demonstrated a decelerating trend for task engagement with a relative
change of 50 - 18 during CAI. In spite of the variability during intervention, visual analysis indicates Trevor’s task engagement increased during CAI as compared to baseline.

Pranav displayed a stable level of task engagement during baseline ($M = 6$; range 0 – 33) with 83% of the data within the stability envelope. The level of task engagement increased during CAI ($M = 57.5$; range = 25 - 84) even though only 36% of the data were within the stability envelope. Data show multiple trend lines with a relative change from 65 – 25. Overall Pranav’s data indicate a mean increase of 51.5% during CAI and a percentage of non-overlapping data of 91.4%.
Figure C-8. Rate of task engagement for all participants.
Effect of CAI on Response Generalization

As noted previously, the purpose of generalization assessment was to determine the extent to which mind-readings skills learned through computer-assisted instruction would transfer to daily life social situations that required problem-solving (i.e., Research Question 3). The generalization phase was initiated in a staggered manner for participants after each one showed a steadily increasing trend and an average of 80% or higher for CAI with at least one session at criterion (100%). Each participant was exposed to at least one scenario per emotion for a total of four role-play situations per session. Each social situation required the participant to label the emotion experienced by a family member after the role-play ended.

During generalization, all four participants demonstrated mastery (mean 93% or higher) on mind-reading skills when presented with social scenarios acted out by family members in typical home settings. Evan immediately showed 100% correct responses and maintained this level of mastery for all nine generalization sessions ($M = 100$). Ashley started generalization at 75% correct answers, missing only the scenario indicating a scared emotion. She reached 100% mastery in the second session and maintained this for six of the eight generalization sessions ($M = 93.75; \text{range}=75-100$). Trevor immediately reached 100% mastery during generalization and maintained this level for eight of the 10 generalization sessions ($M = 93.75; \text{range}=75-100$). Generalization assessment for Pranav was initiated after implementing 2 generalization training sessions as noted previously. He immediately responded to the generalization assessment with 100% mastery and maintained this level for nine of the 10 generalization sessions ($M = 97.5; \text{range}=75-100$).
Effect Size for Mind-Reading and Task Engagement

Effect size is now considered a critical component of single-subject research for determining the magnitude of effect of intervention (Grissom & Kim, 2005). Effect size was calculated for both dependent variables (i.e., mind-reading skills and task engagement) for all participants across the baseline and intervention phases using Cohen’s $d$ (1988) statistic. The following formula was used to calculate Cohen’s $d$ for baseline and intervention conditions:

$$D = \frac{(M_I - M_B)}{(SD_p/ \sqrt{2(1-r)})},$$

where $M_I$ represents the mean score for Intervention, $M_B$ represents the mean score for baseline, and $SD_p$ is the pooled standard deviation for both experimental phases, and $r$ is the correlation between the baseline and intervention data (Dunst, Hamby, & Trivetter, 2004). This formula for effect size is recommended when comparing correlations between phases for single subject research where the number of data points across adjacent phases are unequal (Dunst et al., 2004). The effect size for all four participants was large for mind-reading skills [Evan ($d = 3.06$); Ashley ($d = 7.49$); Trevor ($d = 4.7$); Pranav ($d = 3.33$)] and the overall outcome of the intervention demonstrated a large effect for mind-reading skills ($d = 4.15$). Additionally, the effect size was large for all participants for task engagement as well [Evan ($d = 3.49$); Ashley ($d = 1.17$); Trevor ($d = 1.84$); Pranav ($d = 3.43$)]. The overall outcome of the intervention demonstrated a large effect for task engagement ($d = 1.59$).
COMPREHENSIVE REFERENCES


189


doi:10.177/1362361310362085*


doi:10.1177/108835769601100207


*Journal of Autism and Other Developmental Disorders, 40*, 570-579.

doi:10.1007/210803-009-0901-6


*TechTrends, 52*(2), 52-58.


doi:10.1007/s10864-010-9112-7


doi:10.1007/s10803-009-0842-0


