The Effect of the Movie Time Social Learning Method on the Emotional Inferencing Skills in an Adult with Autism Spectrum Disorder

Jenna Reade
THE EFFECT OF THE *MOVIE TIME SOCIAL LEARNING* METHOD
ON THE EMOTIONAL INFERRING SKILLS IN AN
ADULT WITH AUTISM SPECTRUM DISORDER

A Thesis
Submitted to the John G. Rangos Sr. School of Health Sciences

Duquesne University

In partial fulfillment of the requirements for
the degree of Master of Science

By
Jenna Reade

August 2018
THE EFFECT OF THE MOVIE TIME SOCIAL LEARNING METHOD
ON THE EMOTIONAL INFERENCE SKILLS IN AN ADULT WITH AUTISM SPECTRUM DISORDER

By

Jenna Reade

Approved June 27, 2018

Katherine Belardi, Ph. D.
Visiting Assistant Professor of Speech-Language Pathology
(Committee Chair)

Heather Rusiewicz, Ph. D.
Associate Professor of Speech-Language Pathology
(Committee Member)

Lori Marra, M.A.
Clinical Instructor of Speech-Language Pathology
(Committee Member)

Fevzi Akinci, Ph.D., MHA
Dean, John G. Rangos School of Health Sciences

Mikael D.Z. Kimelman, Ph. D.
Chair, Speech-Language Pathology
Associate Professor of Speech-Language Pathology
ABSTRACT

THE EFFECT OF THE MOVIE TIME SOCIAL LEARNING METHOD
ON THE EMOTIONAL INFERENCING SKILLS IN AN
ADULT WITH AUTISM SPECTRUM DISORDER

By
Jenna N. Reade
August 2018

Thesis supervised by Katherine Belardi, Ph. D.

Individuals with autism spectrum disorder (ASD) have difficulty finding employment and maintaining social relationships. Professional and personal interactions require individuals to make inferences, or read between the lines. Inferencing is a challenge for individuals with ASD because it requires complex information processing. More specifically, it involves the integration of cognitive and linguistic information. There is a lack of research using complex stimuli to treat inferencing. An A-B-A-B design with one adult with ASD was used to examine the effectiveness of a movie-based method, Movie Time Social Learning (Vagin, 2012), on improving the emotional inferencing skills of one young adult with ASD. The participant answered factual and emotional inference questions about short Disney and Pixar movie clips. Visual supports, explicit teaching, and scaffolding as described in Movie Time Social Learning
were used during the treatment phases to teach the participant how to make inferences. Social validity results suggest the participant enjoyed the use of *Movie Time Social Learning*; however, post-test results and visual analysis of the data suggest there may be other factors that influenced the emotional inferencing skills. The current study may inform clinicians about treating emotional inferencing using movies and the external supports from the *Movie Time Social Learning* method (2012).
ACKNOWLEDGEMENT

This project was an unforeseen blessing along my academic career at Duquesne University. It was also fitting that there were changes during the project, but without these unexpected instances, my journey would not have felt complete. I have thoroughly enjoyed the opportunity to learn how to adapt and complete a project with the help of my peers and knowledgeable professionals surrounding me.

My thesis would not have come to its full potential without the unwavering support that I have received from my thesis committee, Dr. Belardi, Dr. Rusiewicz, and Mrs. Marra. Each member has guided my learning and assisted in various but equally significant ways.

I am also grateful to Dr. Kimelman and the entirety of the Speech-Language Pathology Department for their encouragement throughout my first experience with research.

To my colleagues and fellow graduate students that assisted me, thank you. The Class of 2018 has been more influential in this project than they even realize. I am inspired by all of you as clinicians and students.

A very important profession of gratitude to my family and friends. I recognize my fortune daily because of the acceptance and motivation you give me from over 400 miles away.

Lastly, my highest appreciation the participant and his family for their flexibility and for making this project possible.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>iv</td>
</tr>
<tr>
<td><strong>ACKNOWLEDGEMENT</strong></td>
<td>vi</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>ix</td>
</tr>
<tr>
<td><strong>LIST OF FIGURES</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>LIST OF ABBREVIATIONS</strong></td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER I</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Inferencing</td>
<td>1</td>
</tr>
<tr>
<td>Cognitive Theories</td>
<td>3</td>
</tr>
<tr>
<td>Weak central coherence</td>
<td>3</td>
</tr>
<tr>
<td>Theory of mind</td>
<td>4</td>
</tr>
<tr>
<td>Complex information processing model</td>
<td>6</td>
</tr>
<tr>
<td>Treatment</td>
<td>8</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>9</td>
</tr>
<tr>
<td>Research Question</td>
<td>10</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td>10</td>
</tr>
<tr>
<td>Recruitment</td>
<td>11</td>
</tr>
<tr>
<td>Participant</td>
<td>11</td>
</tr>
<tr>
<td>Setting</td>
<td>12</td>
</tr>
<tr>
<td>Materials</td>
<td>12</td>
</tr>
<tr>
<td>Formal assessments</td>
<td>13</td>
</tr>
<tr>
<td>Screening materials</td>
<td>15</td>
</tr>
<tr>
<td>Study stimuli</td>
<td>21</td>
</tr>
<tr>
<td>Procedures</td>
<td>26</td>
</tr>
<tr>
<td>Study design</td>
<td>26</td>
</tr>
<tr>
<td>Pre-test and clarity screening session</td>
<td>27</td>
</tr>
<tr>
<td>Baseline sessions (A1 and A2)</td>
<td>28</td>
</tr>
<tr>
<td>Treatment sessions (B1 and B2)</td>
<td>30</td>
</tr>
<tr>
<td>Post-test</td>
<td>32</td>
</tr>
<tr>
<td>Maintenance phase</td>
<td>32</td>
</tr>
<tr>
<td>Data Collection</td>
<td>33</td>
</tr>
<tr>
<td>Video recordings</td>
<td>33</td>
</tr>
<tr>
<td>Reliability</td>
<td>33</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>33</td>
</tr>
<tr>
<td>Measure of Fidelity</td>
<td>34</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: Baseline formal assessment scores.....................................................12
Table 2: Definitions of clarity.................................................................16
Table 3: Factual questions answered correctly across phases.........................37
Table 4: Emotional inference questions correct across phases.........................39
Table 5: Average factual and emotional inference scores, clarity and reading levels across all phases.............................................................48
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Movie clip with 100% clarity portraying emotion with obvious visual and verbal cues</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>Picture of Hades during the movie clip representing 75% clarity</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Movie clip with 50% clarity portraying emotion with facial expressions</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>Spotlight on Reading Comprehension: Making Inferences &amp; Drawing</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>Flesch-Kincaid reading levels of worksheets and movie transcripts (multiplied by 10 for improved visualization)</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>Graphic organizers adapted from Vagin (2012) for treatment sessions</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Library of Vocabulary Photographs word finding sample stimulus</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>Treatment design showing the distribution of the number of sessions</td>
<td>27</td>
</tr>
<tr>
<td>9</td>
<td>Reading and clarity levels across all phases</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>Cue hierarchy followed within the treatment sessions</td>
<td>31</td>
</tr>
<tr>
<td>11</td>
<td>Factual question performance and reading and clarity levels</td>
<td>36</td>
</tr>
<tr>
<td>12</td>
<td>Emotional inference and factual scores, reading and clarity levels across phases</td>
<td>38</td>
</tr>
<tr>
<td>13</td>
<td>Measures of level, trend, and variability of emotional inferencing</td>
<td>41</td>
</tr>
<tr>
<td>14</td>
<td>Immediacy effect of emotional inferencing intervention</td>
<td>42</td>
</tr>
<tr>
<td>15</td>
<td>Degree of overlap for emotional inferencing performance</td>
<td>43</td>
</tr>
<tr>
<td>16</td>
<td>Consistency of emotional inference performance across the phases</td>
<td>44</td>
</tr>
<tr>
<td>17</td>
<td>Emotional inferencing and baseline condition (word finding) scores across all phases</td>
<td>45</td>
</tr>
</tbody>
</table>
Figure 18: Results of assessments administered during the pre- and post-test sessions…46
LIST OF ABBREVIATIONS

ASD: Autism Spectrum Disorder
CASL-2: Comprehensive Assessment of Spoken Language, Second Edition
CELF-5: Clinical Evaluation of Language Fundamentals, Fifth Edition Metalinguistics
EI: Emotional inference question
KBIT-2: Kaufman Brief Intelligence Test, Second Edition
PIT: Pittsburgh Inferencing Test
TAWF-2: Test of Adolescent/Adult Word Finding, Second Edition
ToM: Theory of Mind
WCC: Weak Central Coherence
CHAPTER I

INTRODUCTION

Adults with autism spectrum disorder (ASD) have difficulty finding and maintaining employment and building meaningful relationships (e.g., Howlin, Moss, Savage & Rutter, 2013; Müller, Schuler, Burton & Yates, 2003; Skorich et al., 2016; Williams et al., 2015). Poor functional outcomes are attributed to the social-communication impairment associated with ASD (American Psychological Association, 2013). Social-communication is a complex set of skills that requires “development and coordination of the following domains: social interaction, social cognition, pragmatics (verbal and nonverbal), and receptive and expressive language processing” (Adams, 2005, p. 182). Previous research suggests individuals with ASD have impairments in all of the social-communication domains, yet there is limited research on treatments targeting inferencing, a skill that taps into all of the domains mentioned above (Lord & Paul, 1997; Mason, Williams, Kana, Minshew, & Just, 2008; Tager-Flusberg, 1981, 1996; Turner-Brown, Perry, Dichter, Bodfish, & Penn, 2008; Wilkinson, 1998; Williams, Minshew, & Goldstein, 2015). The purpose of the study is to treat the emotional inferencing challenges from an information processing perspective. More specifically, external supports and explicit instruction from the Movie Time Social Learning method (Vagin, 2012) are used to decrease the processing demands.

Inferencing

Previous research indicates individuals with ASD have difficulty with inferencing. (Jolliffe & Baron-Cohen, 1999a, 1999b; Norbury & Bishop, 2002; Preissler & Carey, 2005). Inferencing, or reading between the lines, requires integration of world
knowledge (i.e., previous experience), perspective-taking, and with comprehension of verbal and nonverbal information in order to understand a message (Bodner, Engelhardt, Minshew, & Williams, 2015). There are three types of inferences: physical, mental, and emotional. A physical inference refers to concrete problems. For example, if you saw a picture of a child walking out of an ice cream store and then saw a picture of a scoop of ice cream on the ground, one would infer the child’s ice cream fell. When making a mental inference, an individual infers intentionality, or thoughts, unrelated to emotions. A mental inference is made in the following scenario. A man wants to buy a ring but does not have enough money. After the man leaves the store, the ring goes missing. One can make a mental inference that the man stole the ring (Williams, 2009). Lastly, emotional inferences include understanding others’ emotions or reactions. An individual makes an emotional inference when he or she sees a man crying with his head down and hears that he is having a good day. Making an emotional inference in this situation would allow the individual to understand that the man crying feels upset. Social situations, in particular, require inferencing, and typically developing individuals make inferences unconsciously.

Previous research suggests that individuals with ASD have little to no difficulty making physical and mental inferences, but have challenges making emotional inferences (Bodner, et al., 2015; Happé, 1994; Kaland et al., 2005; Mason, Williams, Kana, Minshew, & Just, 2008). Individuals with ASD tend to interpret a communication partner’s message literally and make other types of pragmatic errors such as misinterpretation of idioms, humor, metaphors, and jokes (Dennis, Lazenby, & Lockyer, 2001; Happé, 1994; Jolliffe & Baron-Cohen, 1999a, 1999b; Kaland et al. 2005; Loukusa
et al., 2007; Loukusa & Moilanen, 2009). Individuals with ASD also take longer to process the information, and consequently have difficulty responding during social interactions within the time necessary to meet the conversation expectations (Mason et al., 2008; Minshew, Goldstein, & Siegel, 1997; Minshew & Williams, 2007; Williams, Minshew, & Goldstein, 2015). Together these deficits are theorized to be behavioral signs of cognitive processing challenges.

Cognitive Theories

There are multiple theories hypothesizing the underlying problem(s) resulting in the impaired social-communication behaviors in ASD. Three cognitive theories that attempt to explain the inferencing challenges in ASD are weak central coherence, theory of mind, and complex information processing.

Weak central coherence. Weak central coherence (WCC) was first described by Uta Frith (1989) and expanded by Happé and Frith (1994). The WCC theory postulates that individuals with ASD have difficulty integrating details in order to see the big picture. For example, a woman with WCC will have difficulty inferring a man feels sad, even when she sees him crying with his head down and hears that he is having a good day. This particular inference requires the woman to comprehend the man’s words according to an atypical context for those words (i.e., head down crying). In this example, a person with WCC would interpret the language literally, and not understand that the person is upset, causing the person to seem insensitive. This deficit has significant implications for social relationships. According to WCC, instead of using higher-order thinking skills to understand the main idea or gist, the brain with ASD
processes the details in their respective neural regions (Myles et al., 2002; Sansoti, Was, Rawson, & Remaklus, 2013; Skorich et al., 2016).

The WCC theory has been tested in individuals with ASD with inferencing tasks. Individuals with ASD perform lower compared to typically developing controls in treatment studies targeting inferencing. These studies include answering mental and emotional inferencing questions following a story or situation (Le Sourn-Bissaoui, Caillies, Gierski, & Motte, 2009; Norbury & Bishop, 2002; Saldana & Frith, 2007; Sansoti, Was, Rawson, & Remaklus, 2013; Tirado & Saldana, 2016), comprehending irony at the sentence and discourse levels (Loukusa & Moilanen, 2009; Williams et al., 2013), and recognizing emotion in facial expression with pictures or videos (Akmanoglu, 2015; Cassidy, Ropar, Mitchell, & Chapman, 2014). According to Norbury and Bishop (2002), individuals with ASD were unable to integrate their own knowledge with contextual information and provided a mental or emotional inference irrelevant to the context of the presented story. This lack of integration was also observed when the individuals with ASD exhibited difficulty explaining their answers to inferencing questions because they had difficulty integrating contextual and linguistic information (Loukusa et al., 2007). These difficulties were also demonstrated when individuals with ASD provided the literal interpretation for ironic statements (Loukusa & Moilanen, 2009). While the WCC acknowledges cognitive processing deficits in ASD, it fails to explain why individuals with ASD can make physical and mental inferences, but not emotional.

**Theory of mind.** A more researched theory regarding the social-communication deficits in ASD is poor theory of mind (ToM; Baron-Cohen et al., 1985; Mason,
Williams, Kana, Minshew, & Just, 2008; Tager-Flusberg, 1981, 1996). ToM deficits are thought to be the result of a processing impairment localized in an area between the left medial frontal gyrus and right temporo-parietal junction, which hinders the ability to take another person’s perspective and think about what he or she is thinking and feeling (e.g., Understanding that when someone is crying, he or she is sad; Baren-Cohen, 2000; Mason, et al., 2008; Happé, 1994; Skorich et al., 2016). Intentionality is an important skill because it increases understanding of others’ intentions and perspectives, which assists in making mental and emotional inferences – a skill needed for empathy and social relationships. Several ToM assessments and treatments have been developed and tested to improve intentionality, including the following: the comic strip test (Philpott, Rinehart, Gray, Howlin, & Cornish, 2013), thought-bubble training (Wellman et al., 2002), social skills training programs, mental state teaching, and false belief tasks (Chin & Bernard-Opitz, 2000; Ozonoff & Miller, 1995; Swettenham, 1996). These treatments target the ability to understand the intentions, thoughts and feelings of others. However, findings from these studies indicate improvements in ToM skills within direct treatment, but little to no generalization of skills post-treatment (e.g., Begeer et al., 2011; Chin & Bernard-Opitz, 2000; Feng, Lo, Tsai, & Cartledge, 2008; Ozonoff & Miller, 1995). Another ToM treatment study found that targeting ToM skills alone did not improve mental and emotional inferencing skills or the ability to respond during social interactions (McConnell, 2002). This lack of generalization to other skills and environments suggests that solely targeting ToM as a localized deficit does not improve inferencing skills (Bodner, Engelhardt, Minshew, & Williams, 2015).
**Complex information processing model.** A newer theoretical perspective acknowledges inferencing as a skill that requires cognitive and linguistic processing. The complex information processing model acknowledges the challenges individuals’ with ASD have with integrating linguistic information given the context and perspective taking. Unlike the previous cognitive theories discussed, the complex information processing model also recognizes the important role of language comprehension in inferencing. A complex information processing impairment exists when the level of cognitive and/or linguistic demands (e.g., the complexity or amount of language and/or the time constraints) requires multiple different brain areas to activate and interact with one another across the brain’s hemispheres (Williams, et al., 2006). For individuals with ASD, increasingly complex and abstract information is harder to process because it requires processing across multiple brain regions (Bodner, Engelhardt, Minshew, & Williams, 2015; Le Sourn-Bissaoui, Cailles, Gierski, & Motte, 2009). The difficulty in handling the increased complexity may be partially due to inefficient neural networks (Mason et al., 2008; Williams et al., 2013).

Neuroimaging research has confirmed that the brain with ASD processes inferences differently compared to typically developing peers. Mason, Williams, Kana, Minshew, and Just (2008) included 18 adults with ASD and 18 typically developing controls that read 30 three-sentence stories and answered simple yes/no questions for comprehension, while lying in a functional magnetic resonance imaging (fMRI) scanner. These stories required the participants to make physical, mental, and emotional inferences to answer the questions correctly. The adults with ASD demonstrated increased right hemisphere activation for all three inferencing conditions, whereas the
controls activated the right hemisphere only during mental and emotional inference processing. This abnormal right hemisphere activation suggests that the adults with ASD may be recruiting the right hemisphere homologous language areas when processing more complex inferences.

Another study examined the processing demands of textual information that contained a literal or ironic statement (Williams et al., 2013). These researchers measured the recruitment of neural regions for text content, distribution of workload, and functional connectivity during this task, related to the comprehension of the passages. The study included 15 children and 13 adults with ASD who read brief stories that included literal or ironic language while lying in an fMRI scanner. Following the stories, the participants answered yes/no questions. Results were consistent with previous research suggesting adults with ASD activate more of the right hemisphere during complex cognitive and linguistic tasks (i.e., inferencing), compared to typically developing peers. This study also indicated the adults and children with ASD incorporated left and right hemisphere areas that are necessary for language processing, but the distribution of workload was inappropriate for both conditions, compared to their typically developing peers. Results from these studies suggest the brain with ASD is working harder to process linguistic and cognitive information due to decreased communication between brain areas. The WCC and ToM based treatments only take into account one deficit – integration or perspective taking – whereas the complex information processing theory acknowledges those deficits along with language challenges. Thus, a more suitable inferencing treatment may accommodate for cognitive and linguistic
challenges by externally decreasing the level language and abstractness, while providing organization and repetition.

**Treatment**

There are no evidence-based treatments available for improving inferencing skills from a complex information processing perspective. Anna Vagin’s *Movie Time Social Learning: Using movies to teach social thinking and social understanding* method is designed to promote the use of movies for teaching social understanding (2012). More specifically, the program targets perspective-taking skills and the ability to understand emotions of characters. The movie stimuli allow clinicians, teachers, and parents to reduce the complexity level of information, directly teach how to make inferences, and “discuss the reasons behind overt social behavior without having to look directly at their own behavior,” which all aid in language comprehension (p. 3). In order to choose effective movies, Vagin suggests identifying which ‘mindreader level’ the individual’s cognitive and linguistic abilities align with. Once the movies are selected, Vagin provides a list of nonverbal strategies and visual supports corresponding to the different ‘mindreader levels’ to increase an individual’s comprehension of the movie content. Nonverbal mediating strategies include focusing on intonation, facial expressions, gestures, and direct teaching at an appropriate pace. Using visual supports to improve perspective-taking and understanding emotions allow individuals to learn from a static visual reference. Vagin (2012) also notes that having the ability to pause and replay clips can develop perspective-taking skills without the time pressures of social interactions. The framework and dynamic stimuli align with the supports needed within inferencing
interventions by reducing the cognitive and linguistic demands for with individuals with ASD.

Although Vagin’s *Movie Time Social Learning* book is commercially available, there is limited evidence about its effectiveness (2012). Scott (2015) utilized the Vagin (2012) literature for an unpublished study on social understanding in three children with ASD. Scott followed Vagin’s protocol and targeted emotional recognition, perspective-taking, and empathy with movies. The results demonstrated generalization to untreated emotion recognition. The children within the study also demonstrated understanding of character’s feelings while relating them to themselves and generalizing the concept within other environments. Advanced perspective-taking skills did not improve, but it may have been due to the short length of treatment.

**Purpose of the Study**

Emotional inferencing is an important skill for social-communication and is complex because it requires perspective-taking, verbal and nonverbal comprehension, and social cognition. Individuals with ASD have difficulty with emotional inferencing, likely due to underlying challenges with complex information processing. This study attempts to account for the complex information processing challenges by using visual supports and explicit instruction from the *Movie Time Social Learning* method (Vagin, 2012). The purpose of this single-subject study is to evaluate the effect of an adapted *Movie Time Social Learning* method on emotional inferencing skills in an adult with ASD (2012).
Research Question

This study seeks to answer the following question: does implementation of a movie-based inferencing method, an adapted version of *Movie Time Social Learning* (Vagin, 2012), improve emotional inferencing skills, as measured by emotional inference questions, in a verbal adult with ASD? It is hypothesized that an individualized intervention using movie clips will improve emotional inferencing skills. Movies are an ideal therapy tool for individuals with ASD for the following reasons: (1) movies are engaging, and engagement is the precursor to learning, (2) movies can be rewound and allow for repetition, which is critical for helping individuals with ASD comprehend complex concepts (Cassidy, Ropar, Mitchell, and Chapman, 2014; Cassidy, Mitchell, Chapman, & Ropar, 2015; Scott, 2015) and, (3) unlike worksheets, movies provide verbal and nonverbal information (e.g., facial expressions) which allows clinicians to instruct individuals with ASD explicitly regarding elements to pay attention to and how to interpret the information (Cassidy, et al., 2014). The independent variable for this research question is the adapted *Movie Time Social Learning* method (2012). The dependent variable includes the emotional inferencing skills, as measured by the number of emotional inferencing questions answered correctly throughout the baseline (A1 and A2) and treatment (B1 and B2) sessions.

CHAPTER II

METHODS

This study was a single-subject experimental design with one verbal adult with ASD. Participant enrollment began following approval from the Duquesne University Institutional Review Board.
Recruitment

Recruitment of potential participants was conducted through flyers (see Appendix A) at the Duquesne University Speech-Language-Hearing Clinic. Males and females from any racial and ethnic background were invited to participate in the study. Potential participants contacted the investigators via phone or email as indicated on the flyers.

Participant

The participant was a 24-year-old male with a diagnosis of mild expressive and receptive language disorder secondary to his primary diagnosis of ASD. The participant graduated from high-school and was unemployed at the time of the study. His native and primary language was American English. The participant was a current client at the Duquesne University Speech-Language-Hearing Clinic and case history information was collected via his clinic chart. Participant inclusion criteria were based on a previous inferencing study with individuals with ASD (ages 10-49; Bodner, Engelhardt, Minshew, and Williams, 2015). The participant must have received an ASD diagnosis by a psychologist. It was also required that the participant have a full-scale intelligence quotient (IQ) ≥85, as determined by the Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 2004) score at the pre-test session. Exclusion criteria included: concomitant disorders, such as genetic, emotional, or psychiatric disorders, visual deficits, or hearing acuity deficits (as determined by visual and hearing acuity screenings during the pre-test session). The participant’s baseline assessment scores are listed in Table 1.
### Table 1

**Baseline formal assessment scores**

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Baseline Raw Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Kaufman Brief Intelligence Test, Second Edition</em></td>
<td>91&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>Test of Adolescent/Adult Word Finding, Second Edition</em></td>
<td>69&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>Clinical Evaluation of Language Fundamentals, Fifth Edition</em></td>
<td>47/87</td>
</tr>
<tr>
<td><em>Metalinguistics</em></td>
<td></td>
</tr>
<tr>
<td><em>Comprehensive Assessment of Spoken Language, Second Edition</em></td>
<td>61/116</td>
</tr>
<tr>
<td><em>Pittsburgh Inferencing Test</em></td>
<td>36/47</td>
</tr>
</tbody>
</table>

<sup>a</sup>Standard scores (Mean = 100; SD = 15)

**Setting**

Pre- and post-testing took place in one of the treatment rooms at the Duquesne University Speech-Language-Hearing Clinic. The treatment room had a mirrored, one-way observation window and a wall-mounted video camera with a wireless integrated microphone system for video and audio recording. Baseline sessions and all but one treatment sessions took place in a quiet room in the participant’s home.

**Materials**

Materials for this study included assessment measures, screening materials, and experimental stimuli. Formal assessments were used to identify the participant’s IQ, measure baseline condition and the type of inference the participant had the most difficulty with. Screening materials were developed and used to determine the type of experimental stimuli to implement.

KBIT-2 (Kaufman & Kaufman, 2004). The primary investigator, a graduate student in speech-language pathology, hereafter referred to as “the clinician,” administered an IQ test, which determined that the participant met the eligibility criteria of a full-scale IQ score ≥85 (the cutoff used in Bodner, Engelhardt, Minshew, & Williams, 2015). The KBIT-2 included a verbal scale of receptive and expressive vocabulary items, which involved pointing to pictures and saying the missing part of a riddle, and a nonverbal scale, which included identifying the relationship in a set of pictures or patterns.

TAWF-2 (German, 2016). The TAWF-2 was administered to establish a baseline for the control behavior. The results provided information regarding word finding skills, response delay, and naming accuracy across four subtests (e.g., Picture Naming: Nouns, Sentence Completion, Picture Naming: Verbs, and Picture Naming: Categories).

Inferencing language assessments.

for inferencing skills, the *CELF-5 Metalinguistics* (2014) is a newer assessment developed by the authors of the *TLC-E* (1989). This assessment included updated and revised stimuli and provided tables for analysis of error patterns (Wiig & Secord, 2014). The subtests of interest, Making Inferences and Figurative Language, were low to moderately correlated to the corresponding subtests in the *TLC-E* (1989). The correlations were $r = .30$ and $r = .60$, respectively. Weak correlations may have been due to the norm population differences (i.e., *CELF-5 Metalinguistics* minimum testing age increased). The Making Inferences subtest from the *CELF-5 Metalinguistics* included 12 items that evaluated the participant’s ability to make inferences based on causal relationships or event chains within short stories (2014). This subtest included multiple-choice questions related to the short stories and required the participant to give another plausible reason why the situation happened in the story. The Figurative Language subtest included 17 items that evaluated the participant’s ability to interpret idioms and match each expression by their shared meaning. This subtest also included multiple-choice questions related to replacing an idiom and asked the participant to provide a definition for an expression. For both subtests, the clinician read the test item while the text was presented on a trifold flipbook to the participant. The participant had the most difficulty with emotional inferencing on the Making Inferencing Subtest and responded literally to questions on the Figurative Language subtest. The participant earned the raw score reported in Table 1, which consists of up to one point for multiple choice and up to two points for open ended questions answered correctly.

*CASL-2* (Carrow-Woodfolk, 2017). The Inference and Nonliteral Language subtests were administered. The Inference subtest included 52 items in which the
participant answered a question requiring world knowledge to make an inference. The Nonliteral Language subtest included 50 items in which the participant explained the nonliteral meaning of the presented item. The participant provided nonspecific, irrelevant, literal and incorrect emotional inferencing responses. The raw score reported in Table 1 includes up to one point for each question answered correctly.

PIT (Bodner, Engelhardt, Minshew, Williams, 2015). Following the administration of the standardized assessments to measure inferencing ability, the PIT was administered (2015). The PIT included 28 items in which the clinician asked the participant one of two types of questions, eliciting physical responses (i.e., related to physical actions) or ToM responses (i.e., emotion or non-emotion related intentions of a character). The participant earned a raw score of 36, reported in Table 1. This raw score consists of up to two points for ToM responses and up to one point for physical responses. The majority of the participant’s errors were made during emotional inferencing questions, as he often responded with the literal action or reaction of the character (i.e., physical responses).

Screening materials. The participant was screened for familiarity of movies, clarity of movies, hearing, and vision.

Movie familiarity screening. The clinician created a screening tool to obtain the participant’s familiarity of the listed movies. The screening included 33 movies with a picture of the movie cover, the title, and a Likert rating scale to determine the number of times he saw the movie (0 times - 5 or more times). See Appendix B for the list of movies. The screening included movies that used animated or human characters, and
visual and linguistic input that corresponded to Vagin’s suggestions (e.g., varying clarity of facial expressions, tone of voice, and language complexity).

**Clarity screening.** Previous research indicated that the type of character, animated or human, did not affect the responsiveness of children with ASD (Carter, Williams, Hodgins, & Lehman, 2014). Instead, performance relied on how clear the character’s thoughts and feelings were depicted. Thus, clarity ratings were developed to ensure that the movie clips shown were within the appropriate difficulty level for the participant. The clinician and Derek Tannous, another speech-language pathology graduate student in the Speech-Language Pathology Department at Duquesne University, compiled a list of elements involved in emotion portrayal (e.g., facial expression, tone of voice, and language complexity). Then, they watched movie clips, separate from the project movies, and identified the elements requiring an emotional inference. After the emotional inferencing elements were identified, the students co-constructed clarity ratings on a 0 – 100% scale to quantify the clarity difficulty of the emotional inference in the movie clip. See Table 2 for the definitions of clarity.

Table 2

*Definitions of clarity*

<table>
<thead>
<tr>
<th>Clarity Level</th>
<th>Description</th>
</tr>
</thead>
</table>
| 100%          | ● Obvious language toward the emotion  
● Tone of voice, facial expression, and body language obvious and contribute to emotion conveyed with language  
● Often all 3 elements or 2-3 elements present and exaggerated throughout  
● Volume, prosody, intonation exaggerated all agree with emotion  
● Gestures present also coincide with emotion being presented by language, facial expressions, and tone of voice.  
● Common, more simplistic and obvious emotions are portrayed (sad, happy, mad, scared, etc.)  
● Emotions do not change quickly |
- Not many character emotions to consider over time. No simultaneous emotions

75%
- Obvious facial expressions, body language, or tone of voice
- Often 1-2 visual elements present with language
- Visual elements mostly coincide with language conveying emotion
- If visual input (facial expression, body language) does not match with language, there is obvious tone of voice, gestures, or volume to aid in emotion portrayal (sarcasm, anger, etc).
- Emotions conveyed include mostly simple emotions with some contradiction or common expanded emotions (hatred, jealousy, concern, etc.)
- More natural elements (volume, prosody, intonation) and gestures present with occasional exaggeration
- Begin to change emotion (typically of one character)
- Changing of emotion does not happen quickly or without exaggeration of emotion (either language, facial expression, tone of voice, body language)
- If simultaneous emotions are present, they are common and supported by facial expression, tone of voice, or body language

50%
- Mostly natural emotion conveyed with language, facial expression, body language and tone with little exaggeration
- Occurrences of opposition occur often with somewhat obvious cues to aid in understanding
- Less elements to aid in emotional comprehension, often relying more on tone of voice and subtle facial expressions
- Emotions are more complex and change quickly
- Have to consider how the characters’ emotions are interacting and affecting each other

25%
- Little to no exaggeration of emotion within language, facial expression, body language or tone of voice. Language often contradicts the emotion conveyed by slightly exaggerated facial expressions or tone (One element present)
- When language is not contradicted, elements are present naturally
- Complex emotions (empathy, longing, regret, etc.) are often portrayed with rarely exaggerated facial expressions and tone of voice
- Quickly changing emotions within the same character or between characters’ feelings
- Emotions are often displayed with reduced time for processing and change quite quickly
0%  ● Language and/or facial expressions, tone of voice, and gestures contradict actual feeling conveyed
● Most elements usually not present or lacking natural expression
● Other aids such as volume, intonation and prosody also underused or absent.
● Complex emotions conveyed, sometimes with natural facial expressions or tone of voice, but mostly lacking expression
● Examples of complex emotions include consideration, misunderstanding
● Usually simultaneous and evolving emotions throughout, especially considering the interaction of multiple characters’ emotions

An example of a movie clip rated as 100% clear was a scene from *Finding Dory* (See Figure 1; Collins, Stanton, & MacLane, 2016). The whale and shark pictured in this scene exhibit extreme happiness and excitement. The characters’ emotions are very easy to understand because of their huge smiles, high pitch tone of voice, and body language including the whale’s arms in the air. This clip was rated as extremely clear because the character’s nonverbals matched what he was saying (i.e., “This is amazing! I can see everything!”) and the emotion portrayed was simple. An example of a 75% clear movie clip is in Figure 2. In this scene of *Hercules*, the character named Hades became angry when hearing about a possible powerless future, which was obvious with his fiery red complexion (i.e., facial expression) and tone of voice (Dewey, Musker, Clements, Musker, & Clements, 1997). While looking angry he said, “I’m cool,” creating a contradiction between his words and his nonverbals. Figure 3 depicts a scene from *Zootopia* when the rabbit character clearly felt defeated and the fox character was triumphant and then mocked the rabbit (Spencer, Howard, & Moore, 2016). This movie clip portrayed clear facial expressions but was rated a 50% because sarcasm is a more complex emotion due to its abstractness.
Figure 1. Movie clip with 100% clarity portraying emotion with obvious visual and verbal cues

Figure 2. Picture of Hades during the movie clip representing 75% clarity

Figure 3. Movie clip with 50% clarity portraying emotion with facial expressions
With these definitions in place, the clinician rated all of the study movie clips and worksheets. Interrater reliability was measured with five graduate students in the Speech-Language Pathology Department at Duquesne University, who rated the movie clips and worksheets according to the clarity rating scale. Interrater reliability for the nine movies was 87%. Any disagreements were rated by the supervisor and resolved through consensus.

The clinician administered the clarity screening during the first and second baseline sessions. The clinician prepared one movie clip within each clarity level to determine which clarity level to target during the intervention phase. The movie clip screenings included 10 emotional inference and 10 factual questions for a total of 20 questions. The participant watched the clip in 30 - 50-second segments and answered five questions corresponding to each segment. No assistance was given during the screening. Prior to the administration of the screening, the clinician and the study supervisor decided that 80% correct on emotional inferencing questions would be considered mastery. The clinician administered the first clarity screening, which was a scene requiring emotional inferencing from movie clip with 100% clarity. The participant answered the emotional inferencing and factual questions with 90% accuracy, a score within mastery. The clinician then administered a movie clip during the second baseline session, which corresponded to 75% clarity. The participant’s emotional inference accuracy was 70%, which was below the mastery level during the 75% clarity movie clip; therefore, this would be the clarity level targeted during the intervention.
Hearing screening. The clinician administered a hearing screening during the pre-test session, following standard protocol, using a pure-tone audiometer at 20 dB at 250, 500, 1000, 2000, and 4000 Hz. The participant’s hearing was within normal limits.

Visual screening. The clinician screened the participant’s vision with a picture of a movie clip during the pre-test session. The picture was presented on the same laptop used throughout the treatment sessions. The clinician asked factual questions (see Appendix C) about the image to ensure his vision did not affect his ability to participate in the movie treatment. The participant answered all questions correctly.

Study stimuli. Stimuli were used during baseline, intervention, withdrawal, replication, and maintenance phases. Stimuli included worksheets, movies, graphic organizers, and word finding picture probes.

Spotlight on reading and listening comprehension: Making inferences and drawing conclusions workbook (Johnson & LoGiudice, 2007). Worksheets from the Spotlight workbook were used during the baseline phases of this study (2007). This workbook included stories of varying readability levels (grades 2.5 - 6.9), with corresponding factual and inferencing questions. All inferencing questions were converted to emotional inferencing questions by the clinician for the purposes of this study. Each worksheet included a black and white picture, along with the text 16 - 22 sentences in length, as seen in Figure 4. Although multiple-choice answers were available within the Spotlight workbook, they were not provided to the participant (2007). Instead, the questions were presented as open-ended questions, as seen in Appendix D. The reading level of each worksheet was calculated using the Flesch-Kincaid reading scale to estimate the language level of the stories. The higher the grade/reading level the
more complex language. See Figure 5 for the average reading levels of the worksheets (multiplied by 10 for improved visualization) presented in the baseline and withdrawal phases.

**Figure 4. Spotlight on Reading Comprehension: Making Inferences & Drawing**

*Conclusions* example story (Johnson & LoGiudice, 2007)
Movies. Movies were chosen based on the participant’s movie screening responses, such that only movies that the participant had never seen were selected for the study. This was to prevent prior exposure from impacting his performance. Movies were also chosen based on their availability on Netflix or DVD from the Duquesne University Gumberg Library, to ensure the sources owned the rights to the movie. Nine different movies were used, one for each treatment and maintenance session. The clinician collected transcripts of the target three to five-minute clips to estimate the language level, as measured by the Flesch-Kincaid grade-level reading scale. See Figure 5 for the movies’ reading levels.

Figure 5. Flesch-Kincaid reading levels of worksheets and movie transcripts (multiplied by 10 for improved visualization)
Movie Time Social Learning (2012) materials. Anna Vagin outlines how to categorize the inferencing ability of participants (i.e., junior, moving up or varsity mindreader) and suggests visual and verbal supports that match the ability level. According to Vagin’s suggestions, three to five-minute clips were chosen from movies that required emotional inferencing to fully comprehend the scene. Many of the more clearly conveyed emotions occur within animations; therefore, all nine movies used in the screening and intervention were animated (p. 20). When the participant answered factual or emotional inferencing questions incorrectly, there were several visual and verbal supports (e.g., graphic organizers and feelings word lists) adapted from Vagin’s book used to directly teach emotional inferencing. See Figure 6 for graphic organizers and Appendix E for the feelings word list.
Figure 6. Graphic organizers adapted from Vagin (2012) for treatment sessions
**Word finding probe pictures.** Photographed objects were randomly chosen from 10 different categories within the Library of Vocabulary Photographs (“Library of Vocabulary Photographs,” n.d.). Ten photographs were shown, one at a time, and the participant was asked to name each object. Unlike the TAWF-2 (German, 2016), word finding probes only required the participant to label the photograph. Word finding probes were chosen based on materials available and were intended to measure the participant’s word finding skills during a simple naming task throughout all sessions. See Appendix F for example probe list and Figure 7 for an example stimulus.

![Figure 7. Library of Vocabulary Photographs word finding sample stimulus](image)

**Procedures**

**Study design.** A single subject A-B-A-B design was implemented. Pre- and post-test assessments provided a measure of change in inferencing skills following treatment. Following the procedures outlined by Kratochwill et al. (2010), first, there was an initial baseline period to collect data on the participant’s current emotional inferencing skills. Then, the movie-based intervention for teaching emotional inferencing skills was applied, withdrawn, and replicated again. Within this design, the participant provided his own control for comparison, and the outcome variable was repeatedly measured across different conditions (phases; Kratochwill et al., 2010). The second A-B
(baseline-treatment) phase served as a replication of the first A-B phase. This phase was implemented to eliminate ethical issues of designs that end in withdrawal of treatment.

During the baseline and withdrawal phases (A1 and A2), the participant attended two sessions per week for four sessions (a total of eight sessions across the two baseline sessions). During the treatment phases (B1 and B2), the participant attended two sessions per week for five sessions (a total of seven sessions were completed across the treatment phases). This study planned for 10 treatment sessions total, with five sessions in each phase; however, the second treatment phase only included two sessions due to scheduling conflicts. Each baseline and treatment session was about 30 and 60 minutes each, respectively. The participant was involved in the study for a total of 18 sessions. See Figure 8 for a representation of the study design.

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Baseline A1</th>
<th>Treatment B1</th>
<th>Baseline A2</th>
<th>Treatment B2</th>
<th>Post-test</th>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2, 3, 4, 5</td>
<td>6, 7, 8, 9, 10</td>
<td>11, 12, 13, 14</td>
<td>15, 16</td>
<td>17</td>
<td>18</td>
</tr>
</tbody>
</table>

**Figure 8.** Treatment design showing the distribution of the number of sessions

**Pre-test and clarity screening session.** Assessment administration and the clarity screening took place during a three-hour session that included breaks to reduce participant fatigue. The purpose of the pre-test session was to measure the participant’s inferencing skills before the intervention was applied, and to determine which clarity level to begin with during treatment. Prior to testing, the clinician explained study procedures to the participant and he signed the informed consent form.

The clinician anticipated completing the clarity screening during the pre-test session, but the participant became fatigued. The clarity screenings were administered during the first two sessions of the baseline phase. The average clarity level for the
screenings was 87.5% (range = 75-100%). The average reading level for the screenings was .75 (grades 0.1 and 1.4).

**Baseline sessions (A1 and A2).** Baseline data were collected in two phases (A1 and A2). The purpose of collecting baseline data was to “document a pattern of behavior in need of change with sufficiently consistent level and variability and little or no trend, to allow comparison with a new pattern following intervention” (Kratochwill et al., 2010). Therefore, baseline data were collected for four sessions in each phase. The clinician met individually with the participant for two sessions per week, under the supervision of a licensed speech-language pathologist. Each baseline session was 30 minutes long; the first two sessions were 45 minutes due to the clarity screenings. The average clarity level for the baseline phase (A1) was 28.13% (range = 25-37.5%), with three of the four sessions rated at 25% clarity levels. The average clarity level for the withdrawal phase (A2) was 21.88% (range = 12.5-25%). The average reading levels for the baseline and withdrawal phases were 3.53 (range = grades 2.6-4.75) and 3.99 (range = grades 3.35-4.75), respectively. Refer to Figure 9 for the clarity and reading levels across phases.
Figure 9. Reading and clarity levels across all phases

*Spotlight on Reading worksheet (Johnson & LoGiudice, 2007).* Throughout all baseline and withdrawal sessions, the clinician presented two stories from the *Spotlight on Reading* workbook (2007). The participant listened and read along with each 16 - 22 sentence story. He then answered open-ended factual and emotional inferencing questions following each story. No additional cues or prompts were given.

*Word finding probes.* The participant named 10 pictures of objects (See Figure 7 for Library of Vocabulary Photographs) during A1 and A2. Each probe list consisted of different nouns, which were randomly selected. Responses longer than four seconds were counted as an error, similar to the *TAWF-2* administration protocol (German, 2016).
**Treatment sessions (B1 and B2).** During B1 and B2 sessions (7 total treatment sessions completed), the participant was given an adapted version of the *Movie Time Social Learning* method (2012). The intervention was delivered in a quiet room with minimal distractions twice a week for seven sessions (B2 only completed 2/5 sessions). Each session lasted approximately 60 minutes and consisted of watching a three-to-five-minute movie clip and answering factual and emotional inference questions. The study supervisor and clinician decided that mastery within treatment would be 80% accuracy on emotional inferencing questions across two consecutive sessions with movies of the same clarity level (e.g., two consecutive sessions watching movie clips at 50% clarity). The treatment phases had average clarity levels of 65% (range = 50-75%) and 50%, respectively. The average reading level for the intervention phase was 1.82 (range = grades 0.7-4.0). The average reading level for the replication phase was 1.6 (range = grades 1.4-1.8). Refer to Figure 9 for the reading and clarity levels across the phases.

**Movie clips.** During the intervention and replication sessions, the participant watched the first 30 - 50-seconds of the clip, depending on the density of factual content and emotion depiction within the movie clip. Movies were displayed on an Apple MacBook Air 13-inch laptop, which was placed on a flat surface (e.g., therapy table or table approximately 24 inches from participant). Factual questions were asked prior to the emotional inferencing questions to ensure the participant understood the contextual basis for the emotional inferencing information (Appendix G). The clinician recorded the participant’s first response without clinician support as the emotional inferencing accuracy for that session. The participant then watched the same 30 - 50-second segment again and answered the same questions with prompts, and verbal and visual supports.
The purpose of reviewing the movies in short segments was to provide the time needed to process the complex emotional inferencing information within the movie clip, and the opportunity to pause the clip to explicitly teach the nonverbal and verbal information contributing to the emotional inference. The hierarchy of least to most restrictive verbal cues was utilized to help elicit correct emotional inference responses (see Figure 10). The clinician also utilized visual supports from Vagin (2012) including several graphic organizers, the intermediate feelings list, and specific verbal feedback to help the participant answer the questions correctly following the second viewing of the clip (Figure 6). This procedure was repeated until the remainder of the four interval segments for each movie was finished. There were 20 questions total, including 10 factual and 10 emotional inferencing questions for each movie clip. Following the completion of the movie clip, the clinician and the participant reviewed the completed graphic organizers involving factual and emotional inferencing information. Then, the participant retold the scene and explained emotional inferencing information to the clinician to ensure comprehension of the emotional inferences.

**STRATEGIES AND TECHNIQUES**

- **Cueing Hierarchy**—Select cues that are giving the client the least amount of help to elicit the desired behavior response
  - Least amount of cueing
    - Rephrasing
    - Description
    - Sentence completion; Fill-in the blank
    - Phonemic cues
    - Either/or; multiple choice
  - Most amount of cueing
    - Direct model

**Figure 10.** Cue hierarchy followed within the treatment sessions

**Word finding probes.** At the end of the intervention and replication sessions, the participant named 10 pictures of objects (See Figure 7 for Library of Vocabulary Photographs) to collect data on the control behavior. Each probe list consisted of
different nouns, which were randomly selected. Similar to the *TAWF-2* administration protocol, responses after four seconds were counted as errors (German, 2016).

**Post-test.** The purpose of the post-test session was to determine if the treatment had a positive impact on inferencing skills, as measured by improved scores on the standardized assessments. A two-hour diagnostic post-test session, five weeks following the final replication session included the *PIT* (Bodner, Engelhardt, Minshew, & Williams, 2015), *CASL-2* (Carrow-Woodfolk, 2017), *CELF-5 Metalinguistics* (Wiig & Secord, 2014), and *TAWF-2* (German, 2016). All of these were also administered from the pre-test session. The inferencing assessment administration order was randomly selected (*PIT, CASL-2, CELF-5*) to measure sensitivity in identifying the possible change of the participant’s inferencing ability over time. The *TAWF-2* measured the control behavior, word finding, which should not have resulted in significant changes because word retrieval was not targeted by this intervention. The *KBIT-2* (Kaufman & Kaufman, 2004) and visual and hearing screenings were not re-administered because these measurements are typically stable and not expected to change.

**Maintenance phase.** A 30-minute follow-up session was conducted to measure maintenance of emotional inferencing skills, five weeks following the final treatment session. A novel movie was used for this session. The participant viewed a 75% clear movie clip because that was the last clarity level mastered during treatment. The reading level for this movie clip was 0.8. The participant answered different factual and emotional inference questions (20 questions total) following each 30 - 50-second segment of the movie clip without additional cues. Word finding probes were also administered, similar to the baseline and treatment phases.
Data Collection

Data were collected for factual and emotional inferencing accuracy without visual or verbal supports, and for word finding probes for all phases.

Video recordings. Each session at the Duquesne University Speech-Language-Hearing Clinic was video-recorded on the Intelligent Stream Recorder by Paragon Development Systems. The participant’s speech was recorded using a Sennheiser bodypack wireless transmitter with a lavalier microphone, which has a frequency range of 740-776 MHz. Recordings were named in the format Participant ID_Date, to deidentify the data. Home visits were recorded using an Olympus digital voice recorder VN-5200PC.

Reliability. The study supervisor watched 20% of the sessions from the treatment phases, which were selected at random, to measure reliability of the data collection for the control and target behaviors. The study supervisor independently collected data from the video and/or audio recordings. Intraclass coefficient (ICC) estimates and their 95% confidence intervals were calculated using SPSS version 23 (SPSS Inc, Chicago, IL) based on mean-rating (k = 2), absolute-agreement, and two-way random effects model. Values ranging between .75-.90 were considered “good” and values above .90 were “excellent.” The values for the baseline phases were as follows: .84 for emotional inferencing and .96 for factual responses. The clinician and the study supervisor discussed and resolved any discrepancies in the data.

Data Analysis

The data measured were recorded and graphed according to factual and emotional inference question accuracy without supports, and word finding accuracy across all
phases. The measurements from the assessments administered at the pre- and post-test sessions were graphed according to raw or standard scores. Raw scores were used on the inferencing assessments because they were not standardized for the participant’s age.

**Measure of Fidelity**

The purpose of a fidelity measure is to ensure the clinician implemented the intervention the way in which it was intended. Two sessions from the treatment phases were randomly selected for review by the study supervisor. The supervisor coded aspects related to intervention delivery. Each aspect was rated as follows: 1 – did not implement, 2 – implemented variably, 3 – consistently and appropriately implemented (See Appendix H for fidelity checklist). Based on the fidelity data, the clinician adhered to the intervention procedures (Mean = 2.8, max possible score = 3; SD = .42).

**Measure of Social Validity**

At post-test, the clinician asked the participant five open-ended questions to assess his satisfaction with the intervention (See Appendix I for social validity questions).

**CHAPTER III**

**RESULTS**

This study examined the effect of an adapted *Movie Time Social Learning* method on the participant’s emotional inferencing skills (Vagin, 2012). First, the participant’s performance on factual questions is described, followed by his emotional inferencing performance. Next, a description of post-test scores on inferencing assessments is provided. Finally, social validity results are reported.
Results of Dependent Variables

**Factual question accuracy.** The clinician calculated the participant’s factual question accuracy for all phases. In the four baseline sessions, the participant scored an average of 67.36% on factual questions (range = 55-78.5%), with inconsistent performance across the sessions. During the five intervention sessions, the participant’s average accuracy was 82.8% (range = 44-100%). In four out of five intervention sessions, the participant scored between 90-100% accuracy, but accuracy for session 9, the last intervention session, was 44%. During the withdrawal phase, the participant’s average accuracy was 78% (range = 63-83%). In three out of four withdrawal sessions the participant answered factual questions with 83% accuracy, but during session 10 accuracy was 63%. During the replication phase, the participant scored an average of 90% accuracy (range = 80–100%). During the maintenance session, the participant’s average performance was 90% accuracy. The average accuracy in baseline and withdrawal phases was 72.69% compared to 86.4% during the treatment phases. Refer to Figure 11 and Table 3 for the participant’s performance on factual questions across all phases. Across all phases, the participant’s factual question accuracy was higher than emotional inferencing accuracy (except session 9).

Effect sizes are more robust than visual analysis because it measures how much change was made, instead of the rate at which change was made (Beeson & Robey, 2006). The extent of the change in performance was calculated using a variation of Cohen’s (1988) $d$ statistic published by Busk and Serlin (1992). The numerator involves subtracting the mean baseline phase accuracy from mean intervention phase accuracy. Then, that number is divided by the standard deviation of the baseline phase (82.8-
67.38/9.34). The effect size for this study was 1.65. To interpret this effect size Vannest & Ninch (2015) recommend using a mean-analysis of a body of literature to determine benchmarks but there is not one unavailable for this population. Using the benchmarks (1.4 = small, 3.6 = medium, 10.1 = large) from a study on preschool children receiving phonological treatment (Gierut, Morrisette, & Dickinson, 2015), there was a small treatment effect for factual responses in the present study.

Figure 11. Factual question performance and reading and clarity levels
Table 3

Factual questions answered correctly across phases

<table>
<thead>
<tr>
<th></th>
<th>Screening</th>
<th>Baseline</th>
<th>Intervention</th>
<th>Withdrawal</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>9 / 10</td>
<td>7 / 11</td>
<td>9 / 10</td>
<td>8 / 13</td>
<td>8 / 10</td>
</tr>
<tr>
<td></td>
<td>8 / 10</td>
<td>9 / 12</td>
<td>9 / 10</td>
<td>8 / 10</td>
<td>10 / 10</td>
</tr>
<tr>
<td></td>
<td>6 / 11</td>
<td>10 / 10</td>
<td>9 / 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9 / 12</td>
<td>9 / 10</td>
<td>9 / 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4 / 9</td>
</tr>
<tr>
<td>Average</td>
<td>85%</td>
<td>67%</td>
<td>83%</td>
<td>78%</td>
<td>90%</td>
</tr>
</tbody>
</table>

Visual analysis of emotional inferencing. First, the clinician calculated the emotional inferencing accuracy for each phase. The participant’s emotional inference performance is presented in Table 4 and accuracies, beginning with the baseline phase, are graphed in Figure 12. During the movie clip screening, the participant’s average emotional inferencing accuracy was 80% (range = 70-90%). In the four baseline sessions, the participant scored an average of 31% on emotional inferencing questions (range = 25-33%), with 33% correct in three of the sessions. During the intervention phase, the participant’s average performance was 66% accuracy (range = 50-80%). The average emotional inference score for the 75% clarity sessions in the intervention phase (sessions 5-7) was 76.7% compared to 47.5% on the 50% clarity sessions (sessions 8 and 9), suggesting a decrease in performance when the difficulty increased. During the withdrawal phase, the participant’s average accuracy was 39% (range = 16.5-49.5%). Three out of the four withdrawal sessions were recorded within the 41.5-49.5% range, but
During session 13, accuracy decreased to 16.5%. During the replication phase, the participant’s average accuracy was 70% (range = 60-80%). During the maintenance session, the participant’s accuracy increased to 90%. The participant’s average performance in baseline and withdrawal sessions was 35%, compared to 68% correct during the treatment phases.

**Figure 12.** Emotional inferencing and factual scores, reading and clarity levels across phases
Table 4

*Emotional inference questions correct across phases*

<table>
<thead>
<tr>
<th></th>
<th>Screening</th>
<th>Baseline</th>
<th>Intervention</th>
<th>Withdrawal</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Inferencing</td>
<td>9 / 10</td>
<td>2 / 6</td>
<td>7 / 10</td>
<td>3 / 6</td>
<td>6 / 10</td>
</tr>
<tr>
<td></td>
<td>7 / 10</td>
<td>2 / 6</td>
<td>8 / 10</td>
<td>3 / 6</td>
<td>8 / 10</td>
</tr>
<tr>
<td></td>
<td>2 / 8</td>
<td>8 / 10</td>
<td>3 / 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 / 6</td>
<td>5 / 10</td>
<td>1 / 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>80%</td>
<td>31%</td>
<td>66%</td>
<td>39%</td>
<td>70%</td>
</tr>
</tbody>
</table>

Visual analyses were used to determine whether there was a relationship between the independent variable and the outcome variable, and the strength of that relation (as described in Kratochwill et al., 2010). The interpretation of the six features (i.e., level, trend, variability, immediacy of the effect, overlap, and consistency of data patterns across similar phases) are described in the following sections.

**Predictable baseline pattern.** In determining the predictability of the baseline pattern, the measurement of interest was the participant’s emotional inference accuracy. In the four initial baseline sessions, the participant responded between 25-33% accuracy with three of the four sessions at 33% accuracy. Due to the stability of performance, the targeted behavior was judged to be established.

**Level.** According to Kratochwill et al. (2010), the level refers to the average score of the target behavior obtained during a phase. The average emotional inference accuracy during baseline and withdrawal phases were 31% and 40%, respectively. The
average score increased to 66% and 70% during the intervention and replication phases, respectively. The level for each phase is presented in Figure 13.

**Trend.** The trend, or the “slope of the best-fitting straight line for the data within a phase” is shown in Figure 13 as a dotted line (Kratochwill et al., 2010, p. 18). For the baseline phase, the data show a consistent level, with little to no trend. Once the *Movie Time Social Learning* method (Vagin, 2012) was introduced, the participant’s emotional inferencing accuracy immediately increased but decreased at a moderately steep rate across the phase when the clarity level increased in difficulty. In the withdrawal phase, the participant’s performance was variable, which resulted in a moderately steep decrease in the trend line. In the replication phase, the participant’s performance was measured across two sessions and resulted in a steep increase in the trend line. An effect size was calculated based on changes in accuracy from post-treatment to pre-treatment. Effect size was also calculated for emotional inference accuracy according to a variation of Cohen’s (1988) $d$ statistic published by Busk and Serlin (1992). The numerator involves subtracting the mean baseline phase accuracy from mean intervention phase accuracy. Then, that number is divided by the standard deviation of the baseline phase (66-31/3.5). The effect size for this study was 10. Using the same benchmarks from above (1.4 = small, 3.6 = medium, 10.1 = large) from a study on preschool children receiving phonological treatment (Gierut, Morrisette, & Dickinson, 2015), there was a medium treatment effect in the present study.

**Variability.** The variability refers to “the fluctuation of the data (as reflected by the data’s range or standard deviation) around the mean” (Kratchowill et al., 2010, p. 5). In Figure 13, the dotted lines represent plus and minus one standard deviation around the
mean. The mean percent correct for emotional inferencing during the baseline phase was 31% (SD = 3.5), 66% (SD = 13.6) for the intervention phase, 40% (SD = 13.6) for the withdrawal phase, and 65% (SD = 15) for the replication phase. The large standard deviation in the intervention phase suggests that the participant’s performance was variable, possibly attributed to change in clarity level at session 8. The large standard deviation in the withdrawal and replication phases suggest the participant’s emotional inferencing skills continued to be inconsistent.

Figure 13. Measures of level, trend, and variability of emotional inferencing

**Immediacy of effect.** The immediacy of the effect described the “change in level between the last three data points in one phase, and the first three data points in the following phase” (Kratochwill et al., 2010, p. 18). Visual comparison of the data points, as indicated by shapes, suggests there was an immediate effect on emotional inferencing between the baseline and intervention phases but not between intervention and withdrawal phases (Shown in Figure 14). Because the replication phase only had two
sessions, the immediacy of the effect could not be visually analyzed from the withdrawal to replication phase.

**Figure 14.** Immediacy effect of emotional inferencing intervention

**Degree of overlap.** The overlap of data points between adjacent phases was analyzed. Overlap is the percentage of data points from one phase that overlap with the data points from the adjacent phase. A small percentage of overlap is highly desirable because it suggests the intervention changed the dependent variable. The use of the *Movie Time Social Learning* method (Vagin, 2012) had 0 overlapping data points (0%) between the baseline and intervention phases and 2 overlapping data points between the intervention and withdrawal phases (50%), suggesting a weak effect when the intervention was withdrawn (See Figure 15). Due to the replication phase only having
two data points, the degree of overlap could not be analyzed from withdrawal to replication phases. The percent of nonoverlapping data points for all phases was 86.7%.

**Figure 15.** Degree of overlap for emotional inferencing performance

*Consistency across phases.* Consistency is a comparison of the data from similar phases to determine the extent to which the phases resemble each other. There is inconsistency in emotional inferencing performance from the baseline to intervention phase and from the intervention to withdrawal phase. Consistency is represented by the linked ovals in Figure 16.
Figure 16. Consistency of emotional inference performance across the phases

Word finding. Word finding skills were measured by photograph probes throughout all phases as a baseline condition. The participant’s average word finding performance was 85% (range = 70-100%) during baseline, 90% during intervention (range = 80-100%), 87.5% during withdrawal (range = 70-100%), and 85% during the replication phase (range = 80-90%). During the maintenance phase, the participant’s accuracy was 80%. The participant scored between 70-100% correct throughout the study, which may have demonstrated a ceiling effect (See Figure 17).
Figure 17. Emotional inferencing and baseline condition (word finding) scores across all phases

Pre- and Post-test Assessments

Three inferencing assessments were administered at pre- and post-test sessions. Raw scores were used because the assessments were not normed for the participant’s age range. The PIT (Bodner, Engelhardt, Minshew, and Williams, 2015) raw scores improved from 36 to 42, the CASL-2 (Carrow-Woodfolk, 2017) from 61 to 67, and the CELF-5 Metalinguistics (Wiig & Secord, 2014) from 47 to 54. However, it is important to consider the expected variability, or confidence intervals, for each assessment. The CASL-2 (2017) subtests have a variability total of six points and the CELF-5 Metalinguistics (2014) subtests have a total variability of four points according to 95% confidence intervals. The control behavior, word finding, was measured with the TAWF-2
(German, 2016) and the participant’s standard score increased from 69 at pre-test to 72 at post-test. Refer to Figure 18 for the pre- and post-test assessment scores.

![Figure 18. Results of assessments administered during the pre- and post-test sessions](image)

The clinician analyzed the participant’s responses on the inferencing assessments pre- and post-test. According to the results of the PIT, the participant improved accuracy with ToM-O (non-emotion related ToM responses) and slightly improved ToM-E questions (ToM responses including emotions). The participant’s performance on both subtests of the CASL-2 also slightly improved, with improvements in specificity of responses for open-ended responses (e.g., “He liked it” at pre-test to “He enjoyed playing the sport and was the best player”). The participant improved his accuracy on both subtests on the CELF-5 Metalinguistics subtests at post-test because he reduced paraphrasing responses from the information provided and performed well on figurative
language questions that included transparent meaning. On this assessment, he also increased his ability to provide specific responses that were closer to the meaning of the emotional inferences, even when incorrect.

**Maintenance Phase**

Five weeks after the last session in the replication phase, the participant watched a movie clip at 75% clarity, the last clarity level in which the participant demonstrated mastery. The participant answered 90% of emotional inferencing and factual questions correctly. The participant also named 80% of word finding probes correctly. The participant demonstrated maintenance of emotional inferencing skills (See Figures 12 and 17).

**Measure of Social Validity**

The participant attested that there was nothing within the study that he did not like and would attend another inferencing treatment, especially if movie clips or videos were involved. The participant expressed his enjoyment in participating in the study, especially with the use of the movie clips. According to the participant, “Watching movies is exactly what I like to do. I can look at them [characters] and really understand what they’re saying.” The participant also demonstrated an understanding of how this study may help him in social situations and understanding others’ feelings in the future. He states, “This will help me to maintain communication skills that I have now and help me to know how to talk to people.”
CHAPTER IV

DISCUSSION

The current study investigated the effect of the *Movie Time Social Learning* method (Vagin, 2012) on the emotional inferencing skills in an adult with ASD. It was predicted that the movie-based treatment would improve the participant’s emotional inferencing skills. To summarize the findings of this study, see Table 5. Average factual and emotional inference performances were lowest during the baseline phase and highest during the replication phase. The clarity levels were the least clear during the withdrawal phase and most clear during the intervention phase; however, it is important to note that clarity levels were controlled for treatment sessions only. The reading level was least complex during the replication phase and most complex during the withdrawal phase.

Table 5

*Average factual and emotional inference scores, clarity and reading levels across all phases*

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention</th>
<th>Withdrawal</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg Factual</td>
<td>67%</td>
<td>83%</td>
<td>78%</td>
<td>90%</td>
</tr>
<tr>
<td>Avg Emotional Inference</td>
<td>31%</td>
<td>66%</td>
<td>39%</td>
<td>70%</td>
</tr>
<tr>
<td>Avg Clarity Level</td>
<td>28%</td>
<td>65%</td>
<td>22%</td>
<td>50%</td>
</tr>
<tr>
<td>Avg Reading Level</td>
<td>3.5</td>
<td>1.84</td>
<td>4.0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Results during the withdrawal phase represent a relationship between reading and clarity levels and emotional inference performance. During session 13, the worksheets...
were rated the most complex reading level, and least clear clarity score, which resulted in a low emotional inference performance. The emotional inference accuracy in the first three sessions of the withdrawal phase suggests a possible treatment effect. The participant performed better than the baseline phase, despite the more complex reading level and less clear worksheets. However, his improved performance during the withdrawal phase may have resulted from the intervention phase stimuli supporting his emotional inference skills during more difficult worksheet stimuli.

The participant made a slight improvement during post-test and the withdrawal phase, and maintained skills at follow-up; however, there are several confounding variables that make the data difficult to interpret. The confounding variables that may have influenced the dependent variable include the following: participant factors, different stimuli conditions across baseline and treatment phases, differences in clarity level of emotional inferences, and differences in language complexity across conditions.

**Participant Factors**

It is important to explore whether there was an explanation for performance changes outside of the treatment. There was a decrease in performance during session 13. This session occurred at 6:00pm, whereas all of the other sessions took place in the morning between 10:30am-12:00pm. During the session, the participant reported feeling fatigued, which may have negatively affected his attention. It was clear during session 8 that the participant viewed the main character in *Moana* as a ‘damsel in distress,’ when in fact she depicted as a strong and independent female. The participant’s restricted perspective was apparent when he elaborated his responses and often referred to *Moana* as needing someone like Link who saves the female, Zelda, in the videogame *The Legend
Thus, his decline in performance may be related to his comprehension of the context and reduced flexibility in interpreting others’ thoughts and feelings opposing his own. Finally, the participant’s factual comprehension likely impacted his emotional inferencing performance.

**Stimuli Conditions**

Another explanation for the findings is the different stimuli conditions. The worksheets used in the baseline and withdrawal phases required the participant to listen and read simultaneously, whereas the treatment phases required the participant to listen and watch the movie. The worksheets had one black and white picture, as compared to the bright and attractive visual input from the movies. The worksheets may have been more difficult because they had fewer visual supports, and possibly did not engage the participant as much as the movies. Language complexity, as measured by the Flesch-Kincaid reading level, was more difficult during the worksheets compared to the movies. The worksheets only allowed for information regarding emotional inferences to be apparent via the clinician’s prosody and intonation, and the black and white photograph that was not always relevant to the emotional information within the story.

**Clarity Levels**

Based on the data, it is evident that the clarity of emotional inferences impacted performance. The clarity rating scale was initially developed to measure how clear the emotions were portrayed in the movie clips; therefore, it may not have been a true measure of clarity for the *Spotlight* worksheets (Johnson & LoGiudice, 2007). Moreover, the clarity level was only controlled for B1 and B2; the worksheets were randomly selected prior to the study and rated for clarity following the completion of the study.
The average clarity level for the worksheets was lower, or more difficult, compared to the movie clips (mean A1 and A2 were 28.1% and 21.9% and B1 and B2 was 60.7%). Furthermore, the participant reached the mastery criteria in the middle of the treatment phase (session 7), so the decreased performance in sessions 8 and 9 may be attributed to the more difficult clarity. The participant’s factual performance was better than emotional inferencing across all sessions, except session 9, which had the highest level of language complexity for the intervention sessions. This finding supports the results in Loukusa and Molianen’s (2009) systematic review, which stated that as language complexity increases comprehension decreases in adults with ASD.

**Flesch-Kincaid Reading Levels**

The reading level was utilized to measure the language complexity of the worksheets and movie clip transcripts. Similar to clarity, performance matched language complexity such that more complex language was associated with decreased performance. The worksheets for phases A1 and A2 had much higher average Flesch-Kincaid reading levels (grades 3.5 and 4.0) compared to the movie clip transcripts in B1 and B2 (grades 1.84 and 1.6). Of note, the reading level of the maintenance session movie transcript was the lowest of all sessions (0.8). The increased language complexity of the worksheets may have affected the participant’s comprehension during the baseline and withdrawal phases.

**Clarity and reading level interactions.** According to Table 5, there was a relationship between the clarity and reading levels and the participant’s emotional inference performance. By analyzing A1 and A2, the data suggest that more complex clarity and language level are associated with decreased emotional inference.
performance. Phases A1 and A2 required the participant to read and listen without many visual or auditory cues (e.g., facial expressions or tone of voice) to answer factual and inferencing questions. This relationship was stable across all sessions except session 3. Session 3’s *Spotlight* worksheets (Johnson & LoGiudice, 2007) clarity level average was rated as 25%, but the participant’s performance was decreased compared to other sessions with the same clarity level. The clinician analyzed the worksheets for this session and noticed one of the worksheets contained a social conflict and solution. This situation was somewhat clear to the graduate students, but because of its heavy reliance on pragmatics and social competence, it may have been more difficult/less clear for a participant with ASD. The relationship between reading and clarity levels and inference performance was demonstrated in session 13. The participant expressed his fatigue, and this session had the highest reading level (4.75) and lowest clarity score (12.5), which may have impacted his low overall performance.

**Emotional inference score, clarity and reading level interactions.** There was an improvement of emotional inferencing scores in B1 and B2, which may secondary to the ability to pause and replay the movie clips (Vagin, 2012). The dynamic stimuli were also beneficial because of the opportunities for visual input and direct teaching, the requirement of listening only (versus listening and reading), and the participant’s overall engagement in the movie clips. Intervention and replication phases had lower reading levels, higher clarity levels, and higher inferencing performance. While the data suggest the participant maintained emotional inferencing skills five-weeks post intervention, this session included the lowest reading level and the clarity level in which the participant achieved mastery.
As observed in Figures 12-17, emotional inferencing accuracy in the first three sessions of the withdrawal phase suggests the participant was able to continue using the skills that were explicitly taught during the intervention, despite the higher reading level and lower clarity level, suggesting a possible treatment effect. Learning occurred according to the participant’s ability to maintain his level of emotional inferencing accuracy during the first three sessions within the withdrawal phase (A2). Another explanation for the improvements in A2 is that his success from B1 may have helped him answer the more difficult worksheet stimuli (lower in clarity and higher in reading levels). This idea suggests that easier emotional inferences (B1) may have helped the participant understand the most difficult stimuli (A2).

According to this study, there was a relationship between clarity level and emotional inferencing performance. This means that a higher clarity level, for this individual, resulted in higher emotional inference scores on most occasions. There was also a relationship between clarity level and reading level. Worksheets and movie clips with higher clarity and lower reading level represented less difficult stimuli. Therefore, clarity level, and reading level, and inferencing score interacted such that: 1) A1—low emotional inferencing performance, low clarity levels, and high reading levels, 2) B1—high emotional inferencing performance, high clarity levels, and low reading levels, 3) A2—low emotional inferencing performance, lowest clarity levels, and highest reading levels, and 4) B2—highest emotional inferencing performance, middle clarity levels, and lowest reading levels.
Pre- and Post-test Assessments

The PIT (Bodner, Engelhardt, Minshew, Williams, 2015), CELF-5 Metalinguistics (Wiig & Secord, 2014), and CASL-2 (Carrow-Woodfolk, 2017) were administered as baseline inferencing measurements, and again during the post-test session to measure improvement. The participant demonstrated a slight improvement in skill on the inferencing assessments, noted by an approximate seven-point gain on each assessment. Although the raw scores increased, it may be due to the pretest bias (i.e., only four months between administrations) or the typical expected variability in scores (i.e., confidence interval of assessments). Of note, although the scores at the post-test did not greatly improve, the participant’s responses included approximately 15-18 increased attempts at emotional inferencing. The assessment scores coincided with the withdrawal phase results that demonstrated consistency in emotional inferencing ability.

Limitations

The results of this study provided initial information regarding the impact of an adapted Movie Time Social Learning method (Vagin, 2012) for an adult with ASD who demonstrated difficulty making emotional inferences. Although the results demonstrated an improvement in skills, it is important to consider potential internal and external threats to validity that might have occurred during the study. These threats included factors that affected the cause and effect nature of the study and the ability to generalize the findings.

Potential threats to internal validity included the history of the participant, as he received continuous speech-language therapy services throughout the study, although emotional inferencing was not targeted in treatment. It is also important to consider the maturation of the participant, although spontaneous and rapid learning at his age is not
typical. Due to the post-test session occurring 5-weeks post-treatment, it may have introduced a pretest effect. Also, word finding was not a strong control behavior since the participant’s performance was high. While the clarity rating scale was created to provide an objective measure of clarity, there was limited reliability. Finally, the confounding variables described above makes the results difficult to interpret.

While this type of treatment is typically delivered with Spotlight worksheets and movies in a quiet setting in clinical practice similar to this study, the study outcomes are based on one participant, so the results cannot be generalized to the ASD population. However, as initial research, the results obtained from this single-subject design were informative for future research.

**Clinical Implications**

The results of this study have direct implications for practicing clinicians. According to these findings, language complexity and clarity influence performance. Thus, it is important for clinicians to consider these factors when choosing treatment materials. The clarity definitions created for this study may be a starting point for clinicians to determine their client’s current level of functioning, so that the stimuli are within the zone of proximal development (i.e., achievable with support). The external supports used in this study (i.e., rewatching and pausing of movies, explicit teaching, and graphic organizers) may reduce the information processing demands so that the individual can learn how to make emotional inferences.

**Improvements**

In order to measure the effectiveness of the Movie Time Social Learning method (Vagin, 2012), several improvements need to be made. One improvement for the study
design is to set a criterion for factual question accuracy to ensure comprehension of the movie clip is not affecting the participant’s emotional inferencing performance. Additionally, future studies should control for clarity and/or language complexity and use one stimulus condition (worksheets or movies) to determine whether changes in performance are due to the Movie Time Social Learning method (2012). Continued development and reliability of the clarity rating scale is also needed.

**Future Research**

With advancements in technology, there is a potential to use virtual reality for teaching emotional inferencing skills during real social-communication situations and social interactions. Using this context may improve generalization of the skill toward real social situations (Ip et al., 2018).

**Summary**

The aim of this study was to improve emotional inferencing skills in an adult with ASD. Despite the inconclusive results, the participant demonstrated increased understanding of emotional inferences, and reported improvement and enjoyment during the movie-based treatment. Further investigation of movie-based inferencing treatments is needed to determine the effectiveness at improving emotional inferencing skills with the ultimate goal of helping individuals with ASD to maintain employment and build social relationships (Howlin, Moss, Savage, & Rutter, 2013; Müller, Schuler, Burton & Yates, 2003; Williams et al., 2015).
References


discourse by high-functioning older children, adolescents, and adults with autism.

*Journal of Autism and Developmental Disorders, 45*(9), 2721-2733.


Appendix A

Recruitment Flyer

Dynamic Video Inferencing Treatment Study

**Purpose:** This study will examine the impact of an individualized inferencing treatment on the inferencing skills of adolescents and adults with Autism Spectrum Disorder.

**To participate, you must:**

- Be between the ages 12 to 35 years old
- Speak fluent English
- Have an Autism Spectrum Disorder diagnosis from a licensed psychologist or psychiatrist
- Pass a hearing screening
- Pass a vision screening

**Time required:**

- About 20 treatment sessions lasting about 30-45 minutes each
- 2 evaluation sessions (one pre-training and one post-training) lasting about 2-2.5 hours each
- 1 maintenance session (2-4 weeks after treatment) lasting about 45 minutes

**Study Activities:**

- Complete pre-treatment tests
- Learn about inferencing or “reading in between the lines” by watching videos on iPad
- Practice inferencing skills by acting them out
- Complete post-treatment test and questionnaire

**Location:**

- Sessions will take place at Duquesne University or at your home.

You will not be paid for your participation.

**If you are interested, please contact**

Katie Belardi, PhD, CCC-SLP
Duquesne University Department of Speech-Language Pathology
(412) 396 4217
belardik@duq.edu
Appendix B

Movie questionnaire example

Movies You Might Know

1. Zootopia
How many times have you seen the movie?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never seen</td>
<td>1 time</td>
<td>2-3 times</td>
<td>4-5 times</td>
<td>More than 5 times</td>
</tr>
</tbody>
</table>

69
Appendix C

Visual screening questions

**Finding Dory**

1. Point to all of the fish.

2. What colors are they?

3. Point to the middle fish's eyes.

4. Point to a rock.

5. What is the middle fish doing?
Appendix D

*Spotlight* worksheet questions for one session

**Baseline A Session 1**

**Level 1: Story 10**

1. What did Dalton carry his supplies in?
   a. Beach bag

2. How long did it take Dalton to build his sand castle?
   a. 2 hours

3. Dalton is a castle expert. What does that mean?
   a. Dalton knows a lot about castles

4. Dalton’s eyes almost popped out. What does that mean?
   a. Dalton was amazed

5. (EI) How do you think Dalton felt about sandcastles?
   a. He loved making sandcastles
   b. He was interested/creative

6. How did Dalton know the windows of his castle were perfectly square?
   a. The sides, tops and bottoms all looked straight

7. Why was only part of Dalton’s castle destroyed?
   a. The waves didn’t reach all of the castle?

8. (EI) How do you think Dalton felt when he saw some of his castle knocked down?
   a. Sad/disappointed

9. (EI) How might Dalton’s feelings change?
   a. Determined to fix castle

**Level 2: Story 8**

1. What is the main idea of this story?
   a. Mr. Klingman gives strange assignments
2. What subject does Mr. Klingman teach?
   a. Social studies

3. How many rolls of tape could each group use?
   a. 3

4. How did this assignment fit with what the students were studying?
   a. They made inventions because they were studying about inventors

5. Although the students laughed, they probably weren’t surprised by the assignment. Why?
   a. Mr. Klingman gives out strange assignments

6. (EI) How do you think the students felt about the assignment?
   a. They enjoyed the assignment although it was strange
   b. They liked it because they made fun and silly inventions

7. (EI) Why did Mr. Klingman want to give that assignment?
   a. To encourage the students to be inventors in a fun way

8. (EI) How might the students’ feelings change the next time Mr. Klingman has an assignment?
   a. They might be excited to do something fun while they learn
Appendix E

Vagin (2012) feelings word list used during the treatment phases

Intermediate Feelings

<table>
<thead>
<tr>
<th>Affraid</th>
<th>Angry</th>
<th>Happy</th>
<th>Sad</th>
<th>Surprised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afraid</td>
<td>Afraid</td>
<td>Afraid</td>
<td>Afraid</td>
<td>Surprised</td>
</tr>
<tr>
<td>Anxious</td>
<td>Anxious</td>
<td>Anxious</td>
<td>Anxious</td>
<td>Surprised</td>
</tr>
<tr>
<td>Cautious</td>
<td>Cautious</td>
<td>Cautious</td>
<td>Cautious</td>
<td>Surprised</td>
</tr>
<tr>
<td>Frantic</td>
<td>Frantic</td>
<td>Frantic</td>
<td>Frantic</td>
<td>Surprised</td>
</tr>
<tr>
<td>Freaked out</td>
<td>Freaked out</td>
<td>Freaked out</td>
<td>Freaked</td>
<td>Surprised</td>
</tr>
<tr>
<td>Frightened</td>
<td>Frightened</td>
<td>Frightened</td>
<td>Frightened</td>
<td>Surprised</td>
</tr>
<tr>
<td>Jumpy</td>
<td>Jumpy</td>
<td>Jumpy</td>
<td>Jumpy</td>
<td>Surprised</td>
</tr>
<tr>
<td>Nervous</td>
<td>Nervous</td>
<td>Nervous</td>
<td>Nervous</td>
<td>Surprised</td>
</tr>
<tr>
<td>Panicked</td>
<td>Panicked</td>
<td>Panicked</td>
<td>Panicked</td>
<td>Surprised</td>
</tr>
<tr>
<td>Scared</td>
<td>Scared</td>
<td>Scared</td>
<td>Scared</td>
<td>Surprised</td>
</tr>
<tr>
<td>Stressed</td>
<td>Stressed</td>
<td>Stressed</td>
<td>Stressed</td>
<td>Surprised</td>
</tr>
<tr>
<td>Terrified</td>
<td>Terrified</td>
<td>Terrified</td>
<td>Terrified</td>
<td>Surprised</td>
</tr>
<tr>
<td>Threatened</td>
<td>Threatened</td>
<td>Threatened</td>
<td>Threatened</td>
<td>Surprised</td>
</tr>
<tr>
<td>Worried</td>
<td>Worried</td>
<td>Worried</td>
<td>Worried</td>
<td>Surprised</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interested</td>
<td>Hurt</td>
<td>Sorry</td>
<td>Unsure</td>
<td>Sure</td>
</tr>
<tr>
<td>Interested</td>
<td>Hurt</td>
<td>Sorry</td>
<td>Unsure</td>
<td>Sure</td>
</tr>
<tr>
<td></td>
<td>Hurt</td>
<td>Sorry</td>
<td>Unsure</td>
<td>Sure</td>
</tr>
<tr>
<td></td>
<td>Ignored</td>
<td>Stupid</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sneaky</td>
<td>Liked</td>
<td>Okay</td>
<td>Unfriendly</td>
<td>Wanting</td>
</tr>
<tr>
<td>Mischievous</td>
<td>Liked</td>
<td>Okay</td>
<td>Unfriendly</td>
<td>Wanting</td>
</tr>
<tr>
<td>Sneaky</td>
<td>Liked</td>
<td>Okay</td>
<td>Unfriendly</td>
<td>Wanting</td>
</tr>
<tr>
<td></td>
<td>Liked</td>
<td>Okay</td>
<td>Unfriendly</td>
<td>Wanting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                |             |             |          |               |

73
Appendix F

Sample lists for word retrieval probes presented with colored photographs

1. Lemon
2. Fork
3. Couch
4. Ear
5. Crayons
6. Stapler
7. Screw
8. Sponge
9. French fries
10. Window
Appendix G

Questions for one B1 session

Brave
1:24:25-1:20:49

1:24:25-1:23:29

• 1. What does Merida’s horse do when she asks if it’s hungry?
  o Hits her with his tail
  o Supports: verbal attentional cues, verbal choice if needed

• 2. What does Merida take from the kitchen?
  o Apple for herself and a tray of cookies behind her back
  o Supports: verbal attentional cues, verbal choice if needed

• (I) 3. When the father starts telling the story, how do the kids feel about it?
  o Bored, they know the story, he tells it all the time, tired of hearing it
  o Supports: verbal cues (facial expressions/body language), emotional word bank, graphic organizer

• (I) 4. How do the children’s feelings change once Merida starts telling the story?
  o They’re excited to hear it; Like when she tells it
  o Supports: verbal cues (tone/facial expression), emotional word bank, pictures of emotions, graphic organizer

• (I) 5. How does the father feel when Merida finishes his story?
  o Upset; she told his favorite part
  o Supports: verbal cues (tone/facial expression), emotional word bank, graphic organizer


• 6. What was the mother telling the boys to stop doing?
  o Play with their food
  o Supports: verbal attentional cues, verbal choice if needed

• 7. What does one of the little boys pretend to do when his brother holds some food by his face?
  o Throw up/grossed out
  o Supports: verbal attentional cues, verbal choice if needed
8. At the end of the clip, there is a woman walking in holding a platter. What’s on the platter?
   - Letters
   - Supports: verbal attentional cues, verbal choice if needed

9. How does Merida’s mother feel about princesses having weapons?
   - They shouldn’t have them; not appropriate for a princess
   - Supports: verbal cues (tone), graphic organizer

10. Merida was excited to tell her family that she drank from fire falls. How did everyone react towards her news?
    - Dad: only ancient kings—joking with her
    - Brothers: so impressed
    - Mother: not paying attention
    - Supports: verbal cues (tone, facial expression, body language), emotional word bank, pictures of emotions, graphic organizer

11. What did Merida do with the cookies?
    - Slide them under the table for her brothers to eat
    - Supports: verbal attentional cues, verbal choice if needed

12. What did her mother say about the letters?
    - They all accepted
    - Supports: verbal attentional cues, verbal choice if needed

13. How did the little boys feel about getting excused from dinner?
    - Excited, happy, ran away as quickly as they could
    - Supports: verbal cues (body language, facial expression), pictures of emotions, graphic organizer

14. How does Merida feel about the letters?
    - Confused/unsure at first; then she thought she did something wrong
    - Supports: verbal cues (tone of voice, facial expression), emotional word bank, pictures of emotions, graphic organizer

15. Why does her father spit out his drink when Merida’s mother said that he’ll discuss something with her?
    - He’s surprised, wasn’t expecting it, nervous
1:21:34-1:20:49

- 16. Who ends up explaining what the letters say?
  - Mom
  - Supports: verbal attentional cues, verbal choice if needed

- 17. What does mom explain?
  - Each clan will send men to compete for her for marriage
  - Supports: verbal attentional cues, verbal choice if needed

- 18. At the end of the clip, what does Merida do?
  - Leave the table
  - Supports: verbal attentional cues, verbal choice if needed

- (I) 19. How does Merida feel about having suitors/men from each clan compete for her?
  - Upset, angry, mad, doesn’t want to do it
  - Supports: verbal cues (attentional, body language, tone of voice), emotional word bank, pictures of emotions, graphic organizer

- (I) 20. How does Merida’s father feel about this news for her?
  - He feels bad for her and calls after her
  - Supports: verbal cues (tone, facial expression), emotional word bank, pictures of emotions, graphic organizer

**Review or Summary of What you covered**

- Client and clinician will **review completed materials** (graphic organizers) to ensure understanding of factual and inferential events that occurred during the clip.
- By completing the Me Too! Activity (using sticky notes/ideas written down throughout), the clinician will be able to help the client relate to the feelings just discussed and **draw greater connections** in his/her own life.
Appendix H

Fidelity of Implementation Form

**Fidelity of Implementation Form**

All items are scored from 1-3.

1 – Did not implement the intervention procedure.

2 – Implemented the intervention procedure variably

3 – Consistently and appropriately carried out the intervention procedure

Date of Video: ____________ Participant ID: ___________

1. The movie clip had the type of inference appropriate for client.

2. The clinician introduced the movie.

3. The clinician asked the client to independently make an inference when the movie clip was finished.

4. The clinician scaffolds with questions appropriate for the participant’s language level.

5. The clinician maintained client engagement during the session.

6. The clinician and client completed a retell activity based on the movie clip.
Appendix I

Social validity Questions

1) Did you like the watching the movies? What did you like best? Was there anything you didn’t like? What was it?

“Yes! I liked watching movies. Watching movies is exactly what I like to do. I can look at them [characters] and really understand what they’re saying now. No, I liked everything. I was excited to see what movie I would watch and what we would talk about.”

2) Did you like coming to therapy knowing we would work on inferencing?

“Yes. I really liked watching movies.”

3) Would you attend another inferencing treatment?

“Yes. Learning about peoples’ faces and tone of voice helped me understand.”

4) What did you learn how to do?

“I reviewed what I have already known and now I know how to look at peoples’ faces and know their emotions.”

5) Are there any other skills you want to work on?

“I want to continue working on this. This will help me to maintain communication skills that I have now and help me to know how to talk to people.”