A Comparison of Intelligent Agent Technology and Traditional Preference Assessment Implementation Methods for Preservice Speech-Language Pathologists

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Curriculum and Instruction

by

Brenna Griffen
University of Arkansas
Bachelor of Science in Child and Family Sciences, 2009
Harding University
Master of Education in Special Education, 2019

August 2023
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

____
Elizabeth R. Lorah, Ph.D.
Dissertation Director

____
Christine Holyfield, Ph.D.
Committee Member

____
Christian Goering, Ph.D.
Committee Member

Abstract

Identifying preferred stimuli is an initial step in many evidence-based educational programs for young children. Preference assessments, such as the Multiple Stimulus Without Replacement (MSWO), provide an empirically validated way of identifying and ranking these stimuli. Traditional methods of training professionals to implement MSWO often require the presence of an expert trainer, involve lengthy instructional time, require additional training to transfer skills into the clinical or classroom setting, and necessitate follow-up training to maintain skills over time. Intelligent agent technology may overcome these challenges by providing professionals with easily accessible, consistent instruction. The purpose of the current study was to compare the use of intelligent agent technology with pen and paper self-instructional methods in training preservice speech-language pathologists to implement MSWO with young children. The results demonstrate significant increases in implementation fidelity for two out of five participants and slight increases for the remaining three during the intelligent agent condition. Additionally, the participants collectively scored the results of the MSWO incorrectly nearly half of the time while using traditional methods. In contrast, all participants were able to score and interpret the results accurately during every session using intelligent agent technology. There was a significant reduction in duration of implementation for two participants, a moderate reduction for two participants, and a slight reduction for the remaining participant while using intelligent agent technology. Results of the follow-up survey suggest that all participants found intelligent agent technology had a higher treatment acceptability and was more effective at producing socially significant outcomes than traditional methods. Recommendations for clinicians and future research are discussed.

Keywords: preference assessment, technology, preservice training, speech-language pathologist

Acknowledgements

I would like to express my deepest appreciation to the many people who have supported me throughout this journey. I could not have come this far without you. First, words cannot express my gratitude to my dissertation director, Dr. Elizabeth R. Lorah. Your invaluable guidance and mentorship have meant the world to me. Your expertise, insightful feedback, and constant encouragement have been an unwavering source of inspiration. I truly appreciate all the time you have dedicated to furthering my education and career.

Next, I would like to acknowledge my committee members: Dr. Christine Holyfield and Dr. Christian Goering. You have always challenged me to explore diverse perspectives and think outside of my own discipline and experience. Your feedback and support have played a significant role in refining my ideas and shaping me into the educator I am today. Additionally, this study would not have been possible without the help and support of the all the people at Guiding Technologies.

I am extremely grateful to my family. Your constant love and support have instilled in me a sense of confidence and determination that I never knew I possessed. Thank you for always encouraging me to do my best, believing in me, and giving me the necessary time and space to focus on studying, writing, and revising.

Lastly, this endeavor would not have been possible without my friends and fellow classmates. I am truly grateful for all your encouragement, moral support, and understanding throughout this process. The countless hours we spent studying and writing together have encouraged me to persevere even during the hard days. Thank you all for your unwavering support and belief in my abilities.

Table of Contents

Chapter One: Introduction

The purpose of this study was to compare the use of intelligent agent technology with traditional self-instructional methods during preservice speech-language pathologists' (SLPs) training and implementation of multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessments for young children. Stimulus preference assessments (SPA), including the MSWO, are used to identify stimuli (items) that may serve to increase engagement and appropriate behaviors when incorporated into a variety of instructional programs for young children (Logan & Gast, 2001). The success of these programs relies on the correct identification of preferred stimuli, which can be measured by analyzing a pattern of choosing when the same options are offered consistently over time (Bishop & Kenzer, 2012; Graff & Karsten, 2012a; Logan & Gast, 2001; Virues-Ortega et al., 2014). High levels of fidelity and accuracy in implementing SPA reduce the likelihood of misidentifying preferred stimuli leading to the use of ineffective teaching strategies.

For example, if a practitioner identifies a balloon as a preferred stimulus and gives it to a learner as a consequence of correct completion of math problems, however it is not a preferred activity, correct completion of math problems may not increase but instead result in problem behavior. In contrast, if the learner was given an accurate preferred stimulus, such as a toy car, the correct completion of math problems might increase (Cooper et al., 2020).

Due to the wide use and application of SPA, professionals need effective and efficient methods of learning how to conduct them (Graff & Karsten, 2012a). Traditional training methods often require access to an expert trainer and large amounts of training time (Graff & Karsten, 2012a; Kirkpatrick et al., 2019). Additionally, skills are often rehearsed with other adult trainees and can necessitate further training time when implemented with children with intellectual and developmental disabilities (IDD; Ausenhus & Higgins, 2019; Higgins et al., 2017; Pence et al., 2012; Shapiro et al., 2016). In many studies, there is little to no record of the professionals' skill maintenance over time (Graff & Karsten, 2012b; Hansard & Kazemi, 2018; Shapiro et al., 2016; Pence et al., 2012; Weldy et al., 2014). The current study sought to address these limitations and add validity to the use of technology in training programs for professionals working with children with IDD.

Intelligent agent technology is a goal-oriented computer program that actively responds to user input and coaches the user to decide which, if any, action is appropriate. A common example of intelligent agent technology is the global positioning systems (GPS) that provide directions to a user while driving in a car. The intelligent agent is the technology that actively receives feedback about the position of the car and changes the directions to the user when the position or direction of the car changes. This differs from other data-based technologies that only provide a static read-out of directions. When applied to instruction with preservice professionals, intelligent agent technology can provide real-time assistance to them through complex, evidence-based, instructional programs (Nosek, 2017).

One instructional package that uses intelligent agent technology is the Guidance, Assessment, and Information System (GAINS®). This system includes audio and visual guidance that coaches professionals through complex, evidence-based programs for individuals with IDD (Nosek, 2017). The intelligent agent is accessed through a tablet-based application, which provides visual guidance on the screen of the tablet and audio guidance through a Bluetooth headset. Because GAINS provides in-the-moment direction that adjusts with the learner, there is no need for an expert instructor and professional training can occur while the implementor is with the targeted learner. In addition, this content can be accessed at any time, even after initial skills are learned. Although this technology is designed by experts, it does not necessitate experts to be present during its use. Because this technology removes the need for role-play and feedback from an expert, training times may be drastically reduced. Preliminary research has demonstrated that GAINS can be used to increase implementer fidelity during evidence-based instruction programs, including least-to-most prompting, time delay, and total task chaining procedures (Griffen et al., accepted, pending revisions). Further research is needed to evaluate the use of intelligent agent technology during preservice professional instruction and establish training methods for SPA that will address these specific limitations.

Professionals use many evidence-based teaching strategies for children with IDD that require knowledge of which stimuli a learner prefers (Graff & Karsten, 2012a). Preference can be defined as the relative strength of behaviors between two or more choices (Virues-Ortega et al., 2014). If preferences are not identified accurately, teaching programs will not be as successful (Bishop & Kenzer, 2012; Logan & Gast, 2001). Preferred stimuli can be used across a variety of instructional strategies aimed at skill acquisition, such as discrete-trial training, differential reinforcement, and functional communication training; and behavior reduction, such as reinforcement and response interruption and redirection (Hume et al., 2020; Steinbrenner et al., 2020).

Reinforcement is any change in stimulus that increases the future occurrence of the behavior that immediately precedes it. A reinforcer refers to the specific stimulus (i.e., food or item) that is added or removed in the process (Cooper et al., 2020). As an instructional strategy, reinforcement allows the instructor to manipulate the environment by adding preferred stimuli or removing aversive stimuli after a behavior occurs. This will increase the occurrence of the behavior in the future. Preferred stimuli can also be incorporated into learning activities to increase engagement (Logan & Gast, 2001). The success of these programs relies on the correct identification of preferred stimuli, which can be measured by analyzing a pattern of choosing when the same options are offered consistently over time (Virues-Ortega et al., 2014); (Bishop & Kenzer, 2012; Graff & Karsten, 2012a; Logan & Gast, 2001).

The most empirically validated practices for identifying potential reinforcers are different types of formal SPA (Kodak et al., 2009; Lanner et al., 2010; Logan & Gast, 2001; Mangum et al., 2012). Therefore, many behavior modification programs begin with SPA as an initial foundational step (Bishop & Kenzer, 2012; Graff & Karsten, 2012a). In addition, some research demonstrates that preferences of young children or children with IDD changes frequently over time (Bishop & Kenzer, 2012; Logan & Gast, 2001; Mason et al., 1989). Therefore, SPA need to occur regularly and often to ensure that reinforcers are appropriately identified (Graff & Karsten, 2012a).

Children with IDD often require specialized instruction across multiple domains with a variety of professionals to learn social communication skills (American Psychiatric Association [APA], 2013; Hume et al., 2020). Due to social communication impairments, practitioners often face additional barriers when attempting to identify preferred stimuli (items) for these individuals. They may have difficulty clearly articulating their preferences or answering questions, such as “What do you want?” or “Which do you like?” (Alcalay et al., 2019; APA, 2013; Brodhead et al., 2016; Chappell et al., 2009).

Research has demonstrated a low correspondence rate between stimuli identified through vocal nomination and formal SPA, with items identified through SPA being perceived as more effective by teachers (King, 2016). As a result of restrictive and repetitive behavior patterns, the number of leisure stimuli these individuals interact with can be limited and lead to rapid satiation effects (Alcalay et al., 2019).

Preferred stimuli can be identified through a variety of informal and formal methods. Informal methods generally consist of asking the individual or their caregivers what they like or prefer. Formal methods consist of surveys, checklists, or direct observation of interaction between the individual and different stimuli (Verschuur et al., 2011). Many methods of direct observation involve systematic manipulation of stimuli in order to generate a preference hierarchy or ranking of stimuli as most-to-least preferred. Reinforcers identified through these methods tend to produce a higher degree of efficacy and remain more stable over time when compared to informal methods (Didden & de Moor, 2004; Verschuur et al., 2011; Windsor et al., 1994). Despite 100% of professionals working with children with autism spectrum disorder (ASD) reporting using some form of informal or formal SPA, only 50.5% have received formal training (Graff & Karsten, 2012a). Because the success of much evidence-based instruction relies on the use of preferred items, it is essential to train staff to conduct SPA with fidelity (Bishop & Kenzer, 2012; Graff & Karsten, 2012a).

Traditional training methods rely on the use of Behavioral Skills Training (BST). BST is a four-step process that includes written instructions, modeling, role-play or practice, and feedback (Kirkpatrick et al., 2019). Research has demonstrated that this practice is effective in teaching professionals to implement a wide range of skills, including SPA (Lerman et al., 2004; Luck et al., 2018; Kirkpatrick et al., 2019; Roscoe & Fisher, 2008). However, this practice requires the presence of an expert trainer, can be time consuming for all involved, often requires additional training to transfer skills from adults to children with IDD, and may not reliably maintain over time without additional training (Graff & Karsten, 2012b; Pence et al., 2012; Shapiro et al., 2016).

Other methods of training include self-instructional manuals (Graff & Karsten, 2012b; Hansard & Kazemi, 2018; Ramon et al., 2015; Shapiro et al., 2016), telehealth (Ausenhus & Higgins, 2019; Higgins et al., 2017; Lim et al., 2020; Sump et al., 2018), video modeling (Lipschultz et al., 2015; Miljkovic et al., 2015; Nottingham et al., 2018; Rosales et al., 2015), pyramidal training (Pence et al., 2012), group training (Weldy et al., 2014), and computer-aided instruction (Weston et al., 2020; Wishnowski et al., 2017). All these training methods have attempted to remediate the need for an expert trainer, with varying degrees of success. However, most of these methods still involve lengthy training time for professionals. Because many studies used other trainees for the practice component, it necessitated extra practice to transfer skills to children with IDD and in many cases, these skills did not reliably transfer (Ausenhus & Higgins, 2019; Higgins et al., 2017; Pence et al., 2012; Shapiro et al., 2016). Therefore, a method of training that overcomes these limitations would be of high practical value to professionals working with children.

The aim of the current study was to compare the use of an intelligent agent technology application to pen and paper self-instructional methods during training and implementation of MSWO for preservice SLPs. Fidelity of implementation, duration of assessment and training, and social validity were be measured. The specific research questions examined were:

1. Will using intelligent agent technology to conduct a MSWO preference assessment increase preservice SLPs' fidelity of implementation when compared to pen and paper?

2. Will using intelligent agent technology to conduct a MSWO preference assessment decrease the amount of time spent in generating a preference hierarchy when compared to pen and paper?

3. Do preservice SLPs view the intelligent agent technology as having a higher treatment acceptability than pen and paper?

4. Do preservice SLPs view the intelligent agent technology as more effective at producing socially significant outcomes than pen and paper?

5. Do child participants report a higher degree of treatment acceptability for use of intelligent agent technology over pen and paper?

Chapter Two provides a literature review and contains a more in-depth analysis of the research conducted on preference assessments. Chapter Three contains more information about the methodologies that were used in this study and information about the participants. Chapter Four presents a detailed review of the results for each participant. Chapter Five presents a discussion of the results, including relation to previous research, implications for clinicians, and future research directions.

Chapter Two: Literature Review

This chapter presents research on the importance of and uses for stimulus preference assessments (SPA), types of SPA, strategies used to train professionals to conduct SPA, and information on technological tools that can be used for SPA. The limitations of current literature on these topics are reviewed to highlight areas in research and professional training that current practices have yet to address. In addition, a brief overview of intelligent agent technologies, specifically the GAINS application, is provided.

Preferences

Preference can be defined as the relative strength of behaviors between two or more alternatives and can be measured by analyzing a pattern of choosing when the same options are offered consistently over time (Virues-Ortega et al., 2014). This is consistent with principles of behaviorism, which state that the interactions of any organism with the environment can be explained, predicted, and controlled. All behavior, including choice-making, can be predicted based on patterns of behavior over time (Baum, 2004). An individual's choices can be explained and predicted by examining that individual's history of interaction with those alternatives and the consequences of those interactions over time. Therefore, preferences can change as those interactions with the alternatives vary based on contexts, environments, and other factors (Johnston, 2014).

Instructional Practices Using Preferences

Many evidence-based practices used with children with intellectual and developmental disabilities (IDD) require professionals to determine which stimuli the learner will demonstrate a preference for (Graff & Karsten, 2012a; Hume et al., 2020; Steinbrenner et al., 2020). By incorporating preferred stimuli into an individual's instructional programs, professionals can increase social engagement, skill acquisition, and other desirable behaviors (Hu et al., 2021; Logan & Gast, 2001; Phillips et al., 2017). This may be especially important for speech-language pathologists (SLPs), who play a vital role in the development and education of many young children.

SLPs focus on teaching communication and literacy skills, such as understanding, speaking, writing, and reading, by developing personalized treatment plans tailored to each individual's needs. SLPs also strive to improve social communication and interaction by targeting skills related to back-and-forth conversations, joint attention, linguistic and prelinguistic communication, peer play, and responses to invitations and greetings (Wilkinson, 2017). By teaching foundational communication skills, they enable individuals to fully participate in social, academic, and professional activities improving quality of life and enhancing overall well-being.

To achieve these goals, SLPs utilize a vast array of evidence-based practices (Wilkinson, 2017). One of the most used instructional strategies to increase desirable behaviors is reinforcement, which relies on the use of consequent strategies and often involves presenting an individual with preferred items contingent on specific behaviors (Hume et al., 2020; Steinbrenner et al., 2020). Research has consistently demonstrated that reinforcement can improve outcomes across the wide range of developmental domains, including social interaction and communication (Gillon et al., 2017; Steinbrenner et al., 2020).

Reinforcement

Reinforcement is based on the principles of operant conditioning (Baum, 2004) and can be defined as the addition or removal of a stimulus that increases the future probability of the behavior that immediately precedes it, in the presence of a stimuli that precedes the behavior. The reinforcer, a noun, refers to the specific stimulus (i.e., food or item) that is added or removed (Cooper et al., 2020). As an instructional strategy, reinforcement allows the instructor to arrange the environment by adding positive stimuli or removing aversive stimuli after a behavior occurs, in the presence of specific instructional stimuli. This will increase the occurrence of the behavior in the future. Common examples include presenting a student with a sticker (positive reinforcement) after answering a question correctly or removing a work task after a student appropriately asks for a break (negative reinforcement). In these examples, the student's behavior of answering correctly or asking appropriately will increase in the future.

Reinforcement occurs regularly within the natural environment. For example, when a child is hungry, they might ask a caregiver for a snack. When the caregiver gives the child a snack, the child's behavior of asking for a snack when hungry will increase in the future. Another example of natural reinforcement is opening a window when it is hot. The cool breeze from the open window reduces the heat in the room. In the future, one will open the window more often when it is hot. Reinforcement commonly occurs in social situations, as well. When a child sees a peer and greets the peer, the peer smiles, responds to the greeting, and gives the child attention. When the child sees the peer in the future, his behavior of saying hello will increase.

Reinforcement has been incorporated into instructional programs to teach social interaction skills (Hu et al., 2021), play skills (Nelson et al., 2017), personal hygiene skills (Carter et al., 2019) and vocational skills (Lattimore et al., 2006; McDuff et al., 2019; Graff &

Karsten, 2012a) for individuals with IDD. For instance, in a program aimed at increasing personal hygiene skills, a practitioner might present a child with a Mario action figure upon completion of brushing his teeth. Reinforcement-based strategies can also be used in programs to decrease challenging behaviors, including aggression (Phillips et al., 2017; Wallace et al., 2013), stereotypy (Butler et al., 2021), self-injurious behavior (Phillips et al., 2017; Wallace et al., 2013), noncompliance (Nergaard & Couto, 2021), food refusal (Tereshko et al., 2021) and property destruction (Fuhrman et al., 2016; Graff & Karsten, 2012a). In this case, reinforcement might be used in a program to reduce aggression by allowing a child to engage with bubbles after every half hour of not engaging in aggression.

In a survey of 406 professionals working with individuals with IDD across a variety of settings, all respondents indicated using at least one category of reinforcer, including social praise, tokens, breaks from work, edibles, toys, free play, access to physical activities, sensory items, and community-based activities. The majority of respondents indicated that they use several different categories of reinforcers to teach a variety of skills, including language or communication skills, social interaction, play skills, community safety skills, personal hygiene skills, and vocational skills (Graff & Karsten, 2012a). In addition, reinforcement has been used successfully in over 106 studies with children with ASD and is considered an evidence-based teaching practice by the National Clearinghouse on Autism Evidence and Practice (Steinbrenner et al., 2020). In an international survey of SLPs, over 30% reported regularly using evidence-based interventions, including reinforcement, which indicates its widespread acceptability (Gillon et al., 2017). Since reinforcement occurs naturally within the environment and has a wide base of empirical support, it is generally considered a common and essential practice for professionals working with individuals with IDD.

Given the wide variety of potential reinforcers, professionals must make decisions about which types of stimuli are more likely to operate as reinforcers on an individual basis. Research has demonstrated that there is a strong correlation between preference for a stimulus and its ability to function as a reinforcer (Call et al., 2012; DeLeon et al., 2009; DeLeon & Iwata, 1996; Fisher et al., 1992; Glover et al., 2008; Mangum et al., 2012; Paclawskyj & Vollmer, 1995; Roane et al., 1998). In many cases, identification of preferences is a first and essential step to implementing a reinforcement program (Bishop & Kenzer, 2012).

Other Instructional Programs Using Preferences

Beyond the use of reinforcement, there are other reasons that preference identification is important for professionals working with children with IDD. Preferred stimuli can be integrated into social or vocational activities to increase engagement, such as using preferred toys during peer play or teaching cooking skills with preferred foods (Logan & Gast, 2001). They can be used to identify prelinguistic behaviors (i.e., eye gaze, smiling, gestures, etc.) that can be increased or maintained during instruction (Logan & Gast, 2001). In addition, when both neutral and preferred stimuli are identified, instructors can use them in teaching contingency responding, one-to-one correspondence, and choice making (Logan & Gast, 2001). Professionals can identify novel stimuli with similar properties to established preferences to increase responding, establish new preferences, and decrease habitation that may suppress responding. Often, while attempting to identify preferences, professionals may discover stimuli perceived as aversive and avoid using these in programs (Logan & Gast, 2001).

Challenges in Preference Identification

Because preferences are a foundational component of a wide range of evidence-based interventions, many early childhood experts recommend beginning instructional programs by identifying preferred stimuli (Bishop & Kenzer, 2012). However, this can present professionals with many challenges. One major barrier is the rapidity of shifting preferences over time in often unpredictable ways (Bishop & Kenzer, 2012). Preferences for young children have been observed to shift frequently over weeks and sometimes within the same day or session (Bishop & Kenzer, 2012; Logan & Gast, 2001; Mason et al., 1989). For adults with IDD, over half displayed instability in preferences over a sixteen-month period (Bishop & Kenzer, 2012; Zhou et al., 2001). When comparing stimuli identified during a one-time SPA with stimuli identified during a daily SPA, the stimuli identified as highly preferred matched in 30% or less of observations (DeLeon et al., 2001). This indicates that the majority of preferences changed on a daily basis when compared with the initial SPA. In addition, neutral stimuli can become preferred over time (Logan & Gast, 2001). There may be many factors that contribute to frequently shifting preferences, including context and exposure (Chappell et al. 2009; Gottschalk et al., 2000; McAdam et al., 2005; Vollmer & Iwata, 1991).

Motivation

A motivating operation (MO) is any event that momentarily alters the status of a consequence as a reinforcer or punisher and alters the probability of all behavior that has been reinforced or punished by that consequence. MOs can have satiation effects that decrease the likelihood that a response will occur or deprivation effects that increase the likelihood that a response will occur (Cooper et al., 2020). In other words, when a person is exposed to a preferred stimulus for an extended period of time, they may experience satiation effects that inhibit responses that will result in access to that stimulus. In contrast, if a person has not been exposed to a preferred stimulus for an extended period, they may experience deprivation effects that increase the likelihood of responses that will result in access to that stimulus (Chappell et al., 2009).

For example, if a student has just spent half an hour playing with bubbles, they may be less likely to engage in behavior that would result in access to more bubbles, such as asking for bubbles. In contrast, if a student has not played with bubbles in a week, they may be more likely to engage in behaviors that would result in access to bubbles, such as asking for them. Deprivation and satiation occur naturally in everyday contexts. If someone eats donuts for breakfast on the way to work and there is a box of donuts in the break room, they will be less likely to eat the donuts in the break room. However, if the same person has not had donuts in several days, they will be more likely to eat the donuts in the break room. Another example is when a friend extends an invitation to watch a movie. If the person has recently seen the movie, they are less likely to accept the invitation. Whereas, if they have not seen the movie in several weeks, they may be more likely to accept the invitation.

Research has demonstrated that results of SPA can be altered by exposing participants to or depriving participants of stimuli for an extended period beforehand (Chappell et al. 2009; Gottschalk et al., 2000; McAdam et al., 2005; Vollmer & Iwata, 1991). These results have been demonstrated consistently across both individuals with IDD and typically developing children (McAdam et al., 2005). Shifting MOs due to satiation and deprivation effects may account for the variation in preferred stimuli often observed with young children (Hanley et al., 2006).

Additional factors influencing preferences include specific features of stimuli, instructional settings, and magnitude or duration of access to the stimuli. Because preferences may vary across features of a stimulus, the results of SPA may be less accurate when stimuli are assessed in an array of items with limited features (Fritz et al., 2020). For example, an individual may prefer cherry Starburst® over orange gummy bears but may prefer orange gummy bears over lemon Starburst®. If assessed in an array with only cherry Starburst®, the instructor may assume that the individual prefers all Starburst® over all gummy bears, even though this is not the case. Preferences may also vary across instructional settings, with some stimuli only functioning as reinforcers in specific environments and some across multiple contexts (Pino & Dazzi, 2005). For instance, an individual may prefer a toy train at home, but may not interact with the same toy train in a classroom setting.

The reinforcing efficacy of preferences may be further influenced by duration or magnitude of access (Kodak et al., 2009). For example, ten seconds of a preferred video may not be a sufficient reinforcer to increase task completion, but ten minutes of the same video may function as reinforcer. Similarly, a mini-Snickers® may be an insufficient reinforcer, but a full-size Snickers® might be. In addition, the idiosyncratic nature of preferences is influenced by individual reinforcement history. These preferences can be limited to the specific setting or context in which those events occurred and can have profound effects on behavioral repertoires (Chappell et al., 2009). For example, a child may have a father that plays baseball with him every day at home, giving him social praise, hugs, and attention while playing. The child may show a strong preference for baseball at home, due to the history of social reinforcement with his father. However, the same child may show no preference for baseball at school because there is no history of social reinforcement from his father in that setting.

Satiation and deprivation effects, magnitude and duration of access, shifting instructional contexts, and specific features of stimuli may contribute to the unpredictable and constantly changing nature of MOs in young children (Bishop & Kenzer, 2012). These effects can interact with one another, making it difficult for professionals to identify which MOs are currently in place. Therefore, professionals working with young children are in greater need of an effective and efficient means of consistently identifying preferences.

Further Challenges Specific to Individuals with IDD

Professionals encounter a greater number of challenges when working with children with IDD, such as autism spectrum disorder (ASD). One diagnostic criterion for ASD is persistent difficulties in social communication and interaction across multiple contexts (American Psychiatric Association [APA], 2013). Furthermore, children with ASD have a higher prevalence of comorbid conditions, meaning two or more coexisting medical or psychiatric conditions that occur simultaneously and interact with one another. Among the most commonly occurring comorbid conditions for individuals with ASD are language disorders and impairments (Gillon et al., 2017). When considering the increased prevalence of comorbid language disorders and the communication difficulties inherent in ASD, it is unsurprising that 93.7% of school based SLPs report having at least one child with ASD on their caseload (American Speech-Language-Hearing Association [ASHA], 2022). On average each school based SLP serves 11 children with ASD, which constitutes approximately 23% of their caseload (ASHA, 2022). Because of impairments in language and communication, professionals, such as SLPs, face additional barriers in identifying preferences for individuals with ASD.

Difficulties in communication may prevent children from being able to clearly articulate their preferences or answer questions, such as "What do you like?" (Alcalay et al., 2019; APA, 2013; Brodhead et al 2016; Chappell et al., 2009). King (2016) examined the correspondence between preferences identified through SPA and vocal nomination in children with emotional disorders. Results indicated that the children answered more math questions when the stimulus identified through SPA was offered as a consequence than when the stimulus identified through vocal nomination was offered. In a follow-up survey, the classroom teachers indicated that SPA was perceived as more effective and easier to use in the classrooms (King, 2016).

Individuals with IDD may also have a limited repertoire of leisure stimuli with which they frequently interact (Alcalay et al., 2019). One of the core diagnostic criteria for ASD is patterns of restricted and repetitive behavior, including stereotyped or repetitive motor actions or speech, ritualized behavior patterns, rigid eating habits, and/or perseverative interests (APA, 2013). These repetitive patterns of behavior often lead to highly restricted, fixated interests that can limit the number of preferred stimuli available to assess and lead to rapid satiation effects when the same stimuli are consistently used over time (Alcalay et al., 2019; APA, 2013). Restrictive patterns of eating behavior and limited diets may inhibit the use of a variety of edible reinforcers, as well (Tereshko et al., 2021). These limitations on preferences can further increase the effects of satiation and prevent engagement in novel stimuli.

Stimulus Preference Assessment Measures

In order to overcome the challenges of identifying preferences, research has established several means of assessment, including informal and formal methods. Informal methods do not use a systematic strategy or measurement tool and do not have empirical evidence demonstrating their validity. However, methods of formal assessment include a variety of systematic procedures that have demonstrated varying degrees of efficacy across many studies (Graff & Karsten, 2012a).

Informal Assessment Method

Informal methods include observing the child, asking parents/caregivers what the individual likes, and asking the individual what he/she likes (Graff & Karsten, 2012a). Research has indicated that instructor predicted preferences may not be accurate in predicting the reinforcing efficacy of stimuli (Brodhead et al., 2017; Cannella-Malone et al., 2005; Cannella-Malone et al., 2013; Logan and Gast 2001; Resetar & Noell, 2008). In addition, there has been a statistically insignificant correlation between informal and formal assessment methods across multiple studies (Didden & de Moor, 2004; Verschuur et al. 2011; Windsor et al., 1994). Even when examining the preferences of typically developing children, eight out of nine participants showed a small to no correlation between items reported as preferred by caregivers and items selected in formal SPA. Overall, the items identified in the direct SPA were more likely to function as reinforcers than items identified by the teacher (Cote et al., 2007). A limited amount of evidence suggests that parents may be better than clinicians or teachers at identifying effective reinforcers, but this finding has been variable across studies (Didden & de Moor, 2004; Verschuur et al. 2011; Windsor et al., 1994). In spite of these limitations, informal assessment methods remain popular among professionals as 88.3% of professionals report using at least one informal method of assessing preference (Graff & Karsten, 2012a).

Formal Methods

In contrast to informal methods, formal assessments use a systematic method to identify stimuli that are more likely to have the reinforcing efficacy to increase or maintain responding at levels that can produce socially significant long-term outcomes (Call et al., 2012). These methods vary by the level of interaction with the individual being assessed and can be divided into indirect and direct methods (Verschuur et al., 2011). Formal, indirect methods consist of using published surveys, interviews, questionnaires, and checklists to obtain information from parents or caregivers regarding the individual's preferences (Graff & Karsten, 2012a; Verschuur et al., 2011). The most common and empirically validated is the Reinforcer Assessment for Individuals with Severe Disabilities (RAISD; Fisher et al., 1996). It is estimated that at least 36.5% of professionals use these types of formal, indirect assessment methods (Graff & Karsten, 2012a).

Assessments that consist of some type of direct observation of interaction between the individual and stimuli, include data collection, and make a comparison between relative stimuli, are classified as formal, direct assessments (Mangum et al., 2012). Direct SPA can be further divided into two categories: selection or approach-based procedures and engagement-based procedures (Kang et al., 2013; Kodak et al., 2009). Selection-based procedures involve the individual touching, pointing to, or in another way indicating which stimuli is preferred (Kang et al., 2013). Prelinguistic responses, such as eye gaze, are sometimes used with individuals who cannot physically reach out to items (Cannella-Malone et al., 2013). Selection-based responses may be more useful with children with IDD (Virues-Ortega et al., 2014). Engagement-based procedures measure choice behavior based on the duration of interaction with each stimulus (Kang et al., 2013). Both types of direct SPA are associated with higher levels of reinforcer predictive validity compared to indirect SPA (Cote et al., 2007; Logan & Gast, 2001). Because direct methods of SPA are only useful when stimuli that are likely to be preferred for the individual are included, indirect or informal methods are often used beforehand to identify potential stimuli for inclusion (Cannella-Malone et al., 2013; Fisher et al., 1996; Graff & Karsten 2012a; Kang et al., 2013; Weeden & Poling, 2011). Over half of professionals working with children with disabilities report using some method of direct SPA (Graff & Karsten, 2012a).

An advantage of direct assessments is the production of a preference hierarchy, or a ranking of stimuli from most to least preferred. Stimuli that are interacted with more often than the others presented are called “highly preferred” whereas stimuli that are interacted with less often are called “less preferred” (Mangum et al., 2012). Preference hierarchies are unique to each individual and are an essential component of arranging effective learning when relying on reinforcement (Chappell et al., 2009). If one stimulus loses its reinforcing efficacy, an instructor can begin using the next stimulus in the hierarchy, providing instructors with more flexibility to accommodate shifts in MO during session (Chappell et al., 2009; Kang et al., 2013).

Free Operant Assessments. The most common form of engagement-based SPA is the free operant assessment (FO; Roane et al., 1998). In a FO, stimuli are placed in a designated space or on a tabletop and the individual can interact with any of the stimuli for a set duration, typically 5 minutes (Roane et al., 1998). An observer measures the duration of engagement with each object and a preference hierarchy is created by ranking stimuli according to the duration of interaction for each stimulus (Graff & Karsten, 2012a). Engagement-based SPA may be most useful for individuals with IDD or problem behaviors maintained by denied access to stimuli (Virues-Ortega et al., 2011; Kang et al., 2010). However, other researchers have hypothesized that the length of access to stimuli during the assessment may contribute to satiation effects that are likely to cause shifts in preference and the reinforcing efficacy of selected stimuli (Kodak et al., 2009). Ceiling effects can be observed when the individual engages with only one stimulus, making it difficult to generate a preference hierarchy (Kodak et al., 2009). Only 19.8% of professionals reported using FO (Graff & Karsten, 2012a). In a literature review of clinical decision-making related to SPA for individuals with IDD, ten of forty-nine studies (20.4%) used FO methods to identify reinforcers (Virues-Ortega et al., 2014).

Single-Stimulus Assessments. The first developed and most basic of selection-based SPA is the single-stimulus preference assessment (SS; Pace et al., 1985). In the SS, stimuli are presented one at a time across a set series of trials and an observer measures approach behavior. A preference hierarchy is generated by calculating the percentage of approach responses per stimulus (Graff & Karsten, 2012a). This assessment tends to be used more frequently with children who have difficulty attending, scanning an array of stimuli, or making choices (Cannella-Malone et al., 2013; Logan and Gast, 2011; Virues-Ortega et al., 2011). SS is perhaps the most time-consuming form of SPA and the evaluation time increases as more stimuli are added (Cannella-Malone et al., 2013). Perhaps due to their time intensive nature, only 14.8% of professionals report using SS methods to assess preferences (Graff & Karsten, 2012a) and only five of forty-nine studies (10.2%) used SS methods to identify reinforcers (Virues-Ortega et al., 2014).

Paired Stimulus Assessments. Another form of direct, selection-based procedure is the paired stimulus, or paired-choice stimulus assessment (PS; Fisher et al., 1992). During the PS, stimuli are presented two at a time across a series of trials. Individuals can select only one stimulus at a time and responses are recorded. After completion of the trials, a percentage of approach responses per stimulus is calculated, and converted into a hierarchy of preferences (Fisher et al., 1992). Responses on the PS tend to remain consistent over time and have a high rate of correspondence with reinforcer efficacy (Call et al., 2012). In some cases, the PS can be used with pictures instead of three-dimensional stimuli, which adds to its utility and ability to assess a wider range of interactions (Davis et al., 2010). This SPA is ideal for individuals who do not reliably choose between multiple stimuli or have difficulties scanning an array of stimuli (Cannella-Malone et al., 2013: Virues-Ortega et al., 2014). Although less time intensive than the

SS, the PS tends to require a longer duration of time to conduct than other forms of direct, selection-based assessments and the duration increases as more stimuli are added (Virues-Ortega et al, 2014). Therefore, it is surprising that more professionals (36.4% in total) reported using the PS than any other form of direct SPA (Graff & Karsten, 2012a). Also, PS was used in 29 of 49 (59.2%) research studies, making it the most utilized form of SPA in the literature (Virues-Ortega et al., 2014).

Multiple Stimulus With Replacement Assessments. Another form of direct, selection-based SPA is the multiple stimulus with replacement (MSW; Windsor et al., 1994). In the MSW, stimuli are presented in an array of three to eight, depending on the number of stimuli being assessed, across a series of trials. After each trial, the individual is allowed to interact with the stimulus for about fifteen to thirty seconds, and the stimulus is placed back in the array and all stimuli are rotated before the next trial. Preference hierarchies are established by calculating the percent of approaches per stimulus across sessions (Virues-Ortega et al., 2014; Windsor et al., 1994). MSW may be more accurate than FO and less time consuming than PS (Windsor et al., 1994). It may be most useful for students who are able to scan an array of stimuli and make selections (Cannella-Malone et al., 2013). However, MSW may not reliably identify reinforcers when compared to the PS and in some instances only one or two highly preferred stimuli are selected for the duration of the assessment, which inhibits creation of a preference hierarchy (Windsor et al., 1994). Only 22.8% of professionals report using MSW (Graff & Karsten, 2012a) and only one study out of forty-nine (2.0%) reported using MSW to identify reinforcers (Virues-Ortega et al., 2014).

Multiple Stimulus Without Replacement. In order to overcome the disadvantages of MSW, DeLeon and Iwata (1996) modified the procedure, creating the multiple stimulus without replacement assessment (MSWO). During the MSWO, an array of three to eight stimuli is placed in front of the individual and they are instructed to select one. An observer records the selected stimuli, and the individual is allowed to interact with it for fifteen to thirty seconds. The stimulus is then removed completely and not placed back in the array. The remaining stimuli are rearranged and represented. This process continues until all the stimuli are selected or the individual stops responding (DeLeon & Iwata, 1996). Preference hierarchies are generated by calculating the percentage of approach responses per stimulus (Graff & Karsten, 2012a). Due to the nature of selecting stimuli from an array, many researchers hypothesize that the participant's ability to scan an array and make choices may be prerequisite skills to successful implementation (Cannella-Malone et al., 2013). Occasionally some individuals will display positional bias during the MSWO, by consistently selecting the stimuli in the same position in the array (i.e., the middle item, the far-left item, etc.). This can create an inaccurate preference hierarchy and lead professionals to make erroneous decisions about the reinforcing value of stimulus (Karsten et al., 2011).

Even given these constraints, the MSWO has many advantages over other forms of SPA. Research has demonstrated that MSWO can be less time consuming than other forms of selection-based SPA and averages around 15 minutes depending on the number of stimuli assessed (Cannella-Malone et al., 2013; DeLeon & Iwata, 1996; Kang et al., 2013; Virues-Ortega et al., 2014). The original MSWO procedure (DeLeon & Iwata, 1996) consisted of five trials, or presentations of the full array of stimuli. However, studies have demonstrated that the assessment can be shortened from five trials to only three while still maintaining the ability to consistently identify effective reinforcers (Brodhead et al., 2016; Carr et al., 2000; Conine et al., 2021; Graff & Ciccone, 2002). The three-trial procedure is commonly referred to as a brief

MSWO. Across multiple studies, Kang et al. (2013) found that MSWO more accurately predicted reinforcing effects of identified stimuli and provided consistency of choice results across sessions at a higher rate when compared to other SPA. Because stimuli are removed after selection, MSWO can produce a more distinct preference hierarchy with a wider range of preference values (Karsten et al., 2011; Virues-Ortega et al., 2014). Perhaps due to these advantages, MSWO was used in thirteen of forty-nine (26.5%) studies, making it the second most frequently utilized SPA in the literature (Virues-Ortega et al., 2014).

Despite these advantages, MSWO may be the least utilized SPA in applied practice, with only 9.9% of professionals reporting its use (Graff & Karsten, 2012a). This response is especially concerning given that lack of time was identified as one of the major barriers to regular implementation of direct SPA (Graff & Karsten, 2012a). Considering that the MSWO appears to be the least time consuming and most effective form of direct SPA, it may be that professionals lack the procedural knowledge to perform it effectively.

Training Professionals to Conduct SPA

When asked to identify the purpose of SPA, Graff and Karsten (2012a) found that 6% of respondents reported not knowing what SPA are, another 5% gave a clearly inaccurate answer, and 27% did not answer the question. This indicates that potentially 38% of professionals do not understand the purpose of SPA. When asked if they were familiar with the term SPA, 41% said they were not familiar with the term. About half of respondents reported never having received training on SPA (Graff & Karsten, 2012a).

In terms of SPA implementation, 68.3% of respondents reported that they never conduct full-scale SPA, although 40% report offering two or more choices prior to teaching at least once per day. Another 30% reported never offering choices, which is troubling considering the high degree of correlation between preference and reinforcers (Call et al., 2012; DeLeon et al., 2009; Logan & Gast, 2001). If professionals remain unknowledgeable about and untrained in conducting SPA, they may not be able to deliver the most effective reinforcers or behavior change programs. The most cited barriers to implementation were lack of time and lack knowledge of SPA procedures (Graff & Karsten, 2012a).

Therefore, it is incumbent on researchers to identify strategies to successfully teach professionals to implement SPA in order to deliver effective interventions for behavior change, skill acquisition, and communication development to children with IDD. A means of addressing these barriers is to identify methods of training that do not require direct services of an individual with expertise in conducting SPA (Graff & Karsten, 2012a). In addition, training programs that are cost-effective and portable could reach a greater number of individuals (Graff & Karsten, 2012a). To this end, many researchers have examined training programs to address some or all of these barriers, with varying degrees of success.

Traditional Training Methods

Perhaps the most common method of training involves some or all components of Behavioral Skills Training (BST). This training package is commonly used to train implementation of a wide variety of skills, including discrete-trial teaching, function-based assessments, and SPA (Alexander et al., 2015; Kirkpatrick et al., 2019; Kretlow & Bartholomew, 2010; Rispoli et al., 2011). BST consists of four main components: instruction, modeling, rehearsal, and feedback. The first component, instruction, generally involves a written or oral set of directions given to the trainee. Modeling can occur in-vivo or with the use of a video model recorded by the instructor. The next step is rehearsal and generally occurs between the instructor and trainee or two trainees with the instructor observing. Lastly, the instructor provides feedback, consisting of both positive and constructive evaluations of the trainee's performance. This process is sometimes repeated in full or part if a trainee's skills are not adequately displayed (Kirkpatrick et al., 2019; Luck et al., 2018). Although several studies have demonstrated that BST can be used to successfully train implementers to conduct SPA (Lerman et al., 2004; Luck et al., 2018; Kirkpatrick et al., 2019; Roscoe & Fisher, 2008), BST has several limitations in natural settings. First, it requires an instructor who is experienced and knowledgeable about the procedures being taught. Second, it can require hours of time to train individuals using all four steps. Third, often due to time and setting constraints, the role-play and rehearsal portion is completed using fellow trainees instead of the persons with IDD for whom the SPA is needed. This necessitates further time to evaluate if learned skills can transfer to the intended population and in some cases, additional training if the transfer was not successful (Graff & Karsten, 2012b; Pence et al., 2012; Shapiro et al., 2016). Fourth, many studies have not collected long-term maintenance data, so it remains unknown if fidelity of implementation will maintain over time without the need to revisit training procedures. In order to overcome some or all of the limitations of BST, studies have examined using pyramidal training, group training, self-instruction manuals, computer-aided self-instruction, video modeling, and telehealth to train implementation of SPA procedures.

Use of Pyramidal Training and Group Training

Pence et al. (2012) sought to overcome lack of access to instructors by using a pyramidal model of training for staff in schools. Prior to beginning the study, Pence et al. (2012) trained three students of Applied Behavior Analysis in conducting PS, MSWO, and FO. These students then trained six special education classroom staff using instruction, modeling, practice and feedback. Afterwards, the school staff used this method of training to teach preschool staff to conduct the SPA. Data were collected on the feedback the instructors provided during each step of the SPA. In total, fifteen of the eighteen participants mastered conducting the SPA in two or fewer sessions. Training for each individual lasted an average of 90 minutes with a range of 60-120 minutes. Although this model was generally effective, fidelity of implementation fell below mastery for several participants during the generalization phase (Pence et al., 2012). Weldy et al. (2014) evaluated a method of group training with a thirty-minute instructional video that included models to further reduce the need for long training sessions and access to expert trainers. Seven of nine participants met mastery criteria for the MSWO after the initial viewing but two required a second viewing of the video to reach mastery criteria.

Both of these instructional methods required substantial training time for each participant that many applied settings may not be able to accommodate. Neither study collected long-term follow-up to examine if fidelity would remain consistent over time or across more repetitions of teaching in the pyramidal model.

Use of Self-Instruction and Job Aids

To determine if antecedent only interventions would be effective in training implementation and scoring of SPA without the need for expert trainers, research has evaluated the use of a self-instruction package, or job aid (Graff & Karsten, 2012b; Hansard & Kazemi, 2018; Ramon et al., 2015; Shapiro et al., 2016). A job aid is any tool that provides instruction and guidance to individuals while performing specific tasks, including checklists, flowcharts, decision trees, templates, troubleshooting guides or manuals. Their purpose is to increase efficiency and productivity by providing reminders and information during job performance (Rossett & Schafer, 2006). The original SPA training procedure consisted of a job aid with written instructions from the methods section of previous research for baseline, the written instructions and a data sheet for phase one, and the addition of an enhanced self-instruction manual for phase two. The self-instruction manual contained written instructions with a step-by-step task analysis, pictures, examples, diagrams, and jargon-free language. Eleven special education teachers were trained to conduct PS and MSWO with students of Applied Behavior Analysis reading scripts and acting as simulated child participants. Following mastery, generalization probes were conducted with children from the school. During baseline, professionals conducted the SPA with an average of 38% accuracy. With the addition of the data sheet, participants showed an increase to 59% accuracy, but none were able to reach mastery criteria until the introduction of the enhanced self-instruction manual. Using this manual, professionals reached an average of 99% accuracy of implementation. When conducting a follow-up survey, all participants preferred the use of the enhanced self-instruction manual (Graff & Karsten, 2012b).

In a follow-up study, Shapiro et al. (2016) evaluated the effectiveness of the same self-instructional manual with undergraduates and in-home behavior technicians. Five out of seven undergraduates and four out of five behavior technicians met mastery criteria with the manual alone. The remaining three participants met mastery criteria only after exposure to modeling and feedback sessions. All PS were conducted with adults using simulated child scripts. Participants were allowed to review the manual for thirty minutes before each session and could refer to it during the PS. Generalization probes with children with disabilities were conducted only for the behavior technicians one week after meeting mastery criteria. Three of five participants needed further training either with the manual or the manual with feedback in order to maintain mastery criteria during generalization (Shapiro et al., 2016). These results provide further support that self-instructional manuals alone may not be effective in teaching some professionals to conduct SPA.

To address the limitations of the previous two studies, Hansard and Kazemi (2018) conducted another follow-up study using the same self-instructional manual with the addition of technology, in the form of video and voice-over feedback. The thirty-minute training video was divided into three sections: set-up, implementation, and scoring/interpretation. All four undergraduate students met mastery criteria for conducting the SPA with an adult as a simulated child after watching the video one time and were able to score and interpret the results accurately. These results indicate that technological interventions may be successful at training professionals to accurately conduct SPA without the need for exposure to in-person expert trainers (Hansard & Kazemi, 2018). However, the study did not contain a generalization phase to evaluate if skills could be transferred from simulated to actual children with IDD without further training. Across this series of studies (Graff & Karsten, 2012b; Hansard & Kazemi, 2018; Shapiro et al., 2016), there was no collection of long-term maintenance data and so it is unknown if fidelity of implementation would remain consistent across time without reviewing training materials. Additionally, the first two studies still included the use of expert modeling and feedback.

Self-Instruction Using Technology

Another study evaluated the use of an online computer-aided self-instructional training package on implementation of MSWO for university students and school staff (Wishnowski et al., 2017). Although all participants showed improvement after the program, they did not all reach mastery criteria or demonstrate maintenance of mastered skills during the follow-up. The training process lasted about eighty-five minutes for each participant. Another potential solution to eliminate the need for feedback from expert instructors involves the use of video self-monitoring (Weston et al., 2020). Four graduate students in Applied Behavior Analysis were trained to conduct PS and MSW using a written task analysis, model from an instructor, and rehearsal with an adult as a simulated child. Then, the students performed the SPA with a child with ASD and either received feedback from the instructor or reviewed the video of themselves and completed a fidelity checklist on their performance. Overall, results showed a comparably high increase in performance for both the feedback and the self-monitoring conditions, which suggests that video recordings with self-evaluation may be a way to increase fidelity of implementation without relying on an expert instructor. However, this study used graduate students who may be more motivated to perform skills with fidelity, identify errors, and make corrections than other individuals working with children with ASD. In addition, although training times were not documented, the initial training to conduct the SPA consisted of instruction, modeling, and rehearsal with an expert instructor.

Use of Video Modeling

Other studies have incorporated video models as a part of training packages to eliminate the need for expert models to be present (Lipschultz et al., 2015; Miljkovic et al., 2015; Nottingham et al., 2018; Rosales et al., 2015). Miljkovic et al. (2015) found that a video modeling intervention alone was insufficient for any of the participants to meet mastery criteria on accuracy of implementation of MSWO procedures. However, all participants were able to meet mastery criteria when they were given a written manual after watching the video models. According to the follow-up survey, participants felt that the manual and the video model were needed in order to have confidence in conducting SPA with clients. These results suggest that technology and written instructions together may be sufficient in training professionals to conduct SPA without access to feedback from an expert instructor.

In a similar study, Lipschultz et al. (2015) evaluated a video model with voiceover instructions to train four individuals to conduct PS, SS, and MSWO. Following the video model, all four participants showed immediate and substantial increases in fidelity of implementation, with participants needing two to six repetitions of the video to meet mastery criteria and one participant requiring an extra feedback phase. Participants needed an average total of 82 minutes (range 38-116 minutes) of training before reaching mastery criteria. The initial training used adults as simulated children and all three participants who were evaluated were able to generalize rehearsed skills to children with IDD. Nottingham et al. (2018) attempted to reduce training time by adding a feedback condition to the same video modeling with voiceover instructional package. All participants showed mastery in four to seven sessions and the feedback component did not add a significant amount of training time, however it was provided by an expert instructor. Given that the participants showed similar rates of mastery in comparison with Lipschultz et al. (2015), the addition of a feedback component did not appear to enhance or hasten instruction.

Rosales et al. (2015) used a different video modeling package with embedded audio instruction to train teachers of children with ASD to conduct PS and MSWO. Two of the three teachers met mastery criteria after two to four training sessions, while the third required an additional feedback phase. Adults as simulated children were used during intervention, but all showed generalization to students with ASD following mastery during training. Only two of the three participants maintained mastery levels of implementation during the one-month follow-up probes (Rosales et al., 2015). Overall, the results of these studies demonstrate that some professionals can be trained using video models, but some may require more explicit instruction from experts. Furthermore, not all professionals are able to generalize their skills to the children with IDD and skills may or may not maintain with fidelity over time.

Use of Telehealth

Another way that technology has been utilized in training professionals to conduct SPA is through telehealth (Ausenhus & Higgins, 2019; Higgins et al., 2017; Lim et al., 2020; Sump et al., 2018). Higgins et al. (2017) used a training package that included a multimedia presentation, descriptive feedback from previously recorded baseline sessions, and scripted role-play with immediate feedback provided via telehealth. The total training time with instructor on telehealth was 90 minutes. Immediately following, the participant conducted the SPA with an adult as a simulated child. One participant met mastery after one training session, but the other two participants required one or two extra 90-minute training sessions. Altogether training and assessment lasted 4.1 hours, 4.3 hours, and 6 hours for each participant. Only two of three participants were able to generalize skills at mastery level to a child with disabilities without further training. However, all participants maintained skills at mastery level during the one to two-month follow-up probes (Higgins et al., 2017).

In order to reduce the amount of training time, Ausenhus and Higgins (2019) conducted an additional study that eliminated the multimedia presentation and provided written instructions with real-time feedback via telehealth. Using this method, training times were significantly reduced to between 31.1 minutes to 46.0 minutes per participant. All four participants met mastery criteria in two to three sessions. Adults as simulated children were used during training and only three of the four participants were able to generalize skills to children at mastery level accuracy. All participants' skills maintained at mastery level at the two-week follow-up probes.

The results of these studies indicate that telehealth can be useful in reducing training time and providing access to an expert trainer who is not physically present. However, skills during initial training did not reliably generalize to children with disabilities. Training time with the expert instructor was lengthy in many cases and resulted in only a small number of trained individuals.

Research has examined many solutions to overcome the barriers of traditional training methods, but none have been completely successful in eliminating the need for an expert instructor, significantly reducing instructional time for all participants, providing methods of skill generalization to children with IDD, and maintaining consistently high levels of fidelity over time. Therefore, more effective and efficient methods of training professionals are warranted and should be empirically evaluated.

Intelligent Agent Technology

A new, promising technology for training caregivers and professionals working with children with ASD is an intelligent agent software platform Guidance, Assessment, and Information System (GAINS®). GAINS, part of Guiding Technologies™, is designed by behavior and computer scientists through Temple University and funded by The National Science Foundation (NSF) and National Institutes of Health (NIH; Nosek, 2017). An intelligent agent is a goal-oriented computer program that actively responds to user input and coaches the user to decide which, if any, action is appropriate. This technology allows GAINS to move beyond basic data collection and include guidance that provides real-time assistance to instructors through complex, evidence-based, instructional programs for individuals with ASD (Nosek, 2017). Because the program is created and programmed by experts, it eliminates the need for an expert to be present on site for training. In addition, the program can be created once and used across any number of professionals, significantly reducing the burden of time on the expert trainer.

GAINS provides in-the-moment expert direction at each step for implementers that adjusts as the learner requires. This process occurs during instruction with the learner, allowing the implementer to focus on the learner, while redistributing work and time resources away from independent training to instruction with the learner (Nosek, 2017). Because training occurs with the learner, there is no need for a separate generalization phase. While using GAINS, instruction is provided via a tablet or smart phone-based application that allows for interaction through the device screen, as well as with Bluetooth devices, such as headsets or buttons. Directions are displayed on the screen of the device and when used in conjunction with Bluetooth headsets, can be heard through the headset audio. The user-friendly interface allows the implementer to tap or swipe and allows for maintenance of paperless records, while producing graphical summaries of reports and assessment results (Nosek, 2017). Because guidance remains consistently available on the device, implementation fidelity can be maintained at high levels over an extended amount of time. Preliminary research has demonstrated that GAINS can be used to increase implementer fidelity during least-to-most prompting, time delay, and total task chaining procedures (Griffen et al., accepted, pending revisions).

Purpose Statement

The current study sought to compare the use of an intelligent agent technology to traditional pen and paper self-instructional methods during training and implementation of MSWO for preservice SLPs. Data was collected on fidelity of implementation, assessment duration, and social validity, including treatment utility, accuracy, and efficacy. Preservice SLPs used pen and paper methods, as well as intelligent agent technology to conduct and score

MSWO. The next chapter reviews the methodologies used in this study, participant information, definitions of dependent measures, descriptions of materials and setting, and procedural details. The results of the study are presented in Chapter Four and Chapter Five provides a discussion and interpretation of the results.

Chapter Three: Methodology

Purpose

The purpose of this study was to compare the use of intelligent agent technology to typical pen and paper self-instructional methods during preservice speech-language pathologists' (SLPs) training and implementation of multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessments. Both implementation strategies will be measured in terms of fidelity of implementation, duration, and social validity. In addition, this study also evaluated the perceptions of child participants in the use of intelligent agent technology and pen and paper for preference assessments.

Research Questions

1. Will using intelligent agent technology to conduct a MSWO preference assessment increase preservice SLPs' fidelity of implementation when compared to pen and paper?

2. Will using intelligent agent technology to conduct a MSWO preference assessment decrease the amount of time spent in generating a preference hierarchy when compared to pen and paper?

3. Do preservice SLPs view the intelligent agent technology as having a higher treatment acceptability than pen and paper?

4. Do preservice SLPs view the intelligent agent technology as more effective at producing socially significant outcomes than pen and paper?

5. Do child participants report a higher degree of treatment acceptability for use of intelligent agent technology over pen and paper?

Participants

Adult Participants

Participant demographic information is reflected in Table 1. There were five female participants between the ages of 20 and 22. Three participants identified as White/Non-Hispanic or Latino. One participant identified as White and Hispanic or Latino and one participant identified as Hispanic. Four participants had no experience working in the field of special education or communication disorders and one participant had less than one year of experience. No participants had experience conducting SPA or any training in conducting SPA. All participants were Communication Sciences and Disorders majors, with three juniors, one sophomore, and one first year graduate student. This information was collected prior to beginning the study using a written survey (see Appendix A). Inclusion criterion were that each participant was a student at the university in the Communication Sciences and Disorders department and had no history of conducting or training in MSWO procedures regardless of experience working with children.

Simulated Learner

The simulated learner was an adult who performed the role of a child learner participating in the preference assessment for all sessions, except generalization probes. A doctoral student in the field of special education played the role of the simulated learner for all participants across conditions. The simulated learner did not provide any feedback on procedures and performed the actions as outlined in the prewritten scripts (adapted from Kuhn, 2017; see Appendix B).

Child Participants

The child participants were two neuro-typically developing White/Non-Hispanic or Latino males, ages four and five, recruited from the community. Demographic information were collected for the children, including participant age and ethnicity prior to beginning the study (see Appendix C). They only participated in generalization probes at the beginning and end of the study.

Informed Consent

This study was approved by the University of Arkansas Institutional Review Board (IRB) Committee under Protocol # 2208415588 (see Appendix D). Informed consent was obtained prior to beginning the study by meeting face-to-face with each adult participant and with the parent/guardian of each child participant prior to beginning the study. The primary investigator discussed the nature of the study, what type and how data would be collected, how videos would be taken, how the data and videos would be used and how all information pertaining to the study would be used. This information was also provided in a consent form that adult participants and parent/guardians of child participants signed after the face-to-face meeting. A copy of the adult and child consent forms are included as Appendix E and Appendix F, respectively.

Settings and Materials

Prior to beginning the study, an informal stimulus preference assessment (SPA) and free operant (FO) was conducted with caregivers of the child participants to identify seven to eight stimuli that the child participant interacts with (or consumes) consistently to use during the MSWO generalization probes. Each generalization session was conducted at a child-sized table. Items for the simulated learner were chosen at random and sessions were conducted at an adultsized table. The study used a brief MSWO preference assessment with each session consisting of three trials, or presentations of the full array of stimuli.

During pen and paper self-instructional condition, written instructions of procedures adapted from DeLeon and Iwata (1996) and used in Higgins et al. (2017) was given to participants (see Appendix G). In addition, the participant was provided with a paper data sheet (see Appendix H) to collect data on the learner's responses and help with scoring the MSWO. This data sheet was adapted from those used in Higgins et al., 2017.

During the intelligent agent condition, the GAINS application was accessed via tablet-based devices. Participants were able to access audio instructions through AfterShokz® OpenMove™ Wireless Bluetooth headsets. Videos of sessions were recorded via a separate tablet-based device and used to help facilitate data collection.

The adult simulated learner read from scripts adapted from Kuhn (2017; see Appendix B) during all sessions except the generalization probes. One of the three different scripts were chosen at random prior to beginning the session and the adult participants were blind to which script the simulated learner was using.

Implementer and Training

The primary investigator was a doctoral student in Curriculum & Instruction with twelve years of experience working with children with IDD, with four and half of those years focused in ABA and two and a half years as a Board Certified Behavior Analyst (BCBA). The primary investigator was trained in performing all types of SPA during coursework, practicum experiences, and clinical training using both didactic training and Behavioral Skills Training (BST) from other professionals. An additional graduate student in the Communication Disorders department was used as a secondary observer to collect interobserver agreement (IOA). The primary investigator reviewed data collection methods with the secondary observer and the secondary observer practiced scoring sample videos until proficient in data collection methods.

Dependent Measures

The primary dependent measure was fidelity of implementation of MSWO procedures, as assessed with the Fidelity Checklist provided in Appendix I (adapted from Pellegrino, 2019). In this checklist, MSWO procedures are divided into twenty-three steps. Due to the complex nature of the procedure, each step is further broken down into discrete components, for a total of one hundred and forty components across three trials. An independent observer marked a "+" for each component performed independently and correctly, a "– " for each component performed incorrectly, and a "N/A" if the component did not apply. An independent and correct response was defined as completing the component as defined on the Fidelity Checklist. For example, step five (item selection) contains three components. The first component is "the learner is allowed up to 10s for selection." If the learner selected an item before the 10s or the participant removed the items after 10s, the observer scored a "+" for that component. If the participant removed the items before 10s passes, the observer scored a "-" for the component. The second component of step five (item selection) is "the learner is blocked from accessing items if more than one item is selected." If the learner did not attempt to select multiple items, the observer marked "N/A." However, if the learner attempted to select multiple items and access was blocked, the observer marked "+". If the learner attempted to select multiple items and access was not blocked, the observer marked "-". After each trial, the number of components marked with "+" and "N/A" was counted. One session consisted of three trials. Following the third trial, the total number of components marked with a "+" or "N/A" from all three trials was added together and divided by the total number of components (one hundred and forty) and then multiplied by 100 to obtain a percentage correct implementation for the entire session.

The secondary dependent measure was the duration of the assessment. A timer was started as soon as the child participant or simulated learner was seated at the table. The timer was stopped after the participant completed the MSWO and calculated the results, generating a preference hierarchy. Videos of sessions were recorded by the primary investigator and used to assist with data collection and IOA.

Social validity was measured using a System Usability Scale survey (Bangor et al., 2008; see Appendix J) administered to each adult participant upon completion of the study. The survey was divided into three sections: Experiences with GAINS, features of GAINS, and using GAINS for preference assessment. It contained a total of 33 statements and used a Likert-type scale of responding, in which the respondent marked “strongly disagree,” “disagree,” “neutral,” “agree,” and “strongly agree” for each statement.

A second measure of social validity was a questionnaire given orally to the child participants, included as Appendix K. The questionnaire consisted of five questions related to the treatment acceptability of using pen and paper, intelligent agent technology, conducting preference assessments, and participating in the study. The child participants were also shown a visual of a green smiley face, yellow face with a straight mouth, and red frowning face and allowed to respond to the questions by pointing at the faces instead of or in addition to responding verbally.

Interobserver Agreement and Procedural Fidelity

Interobserver agreement (IOA) data were collected for 52% of sessions in the intelligent agent condition, 60% of sessions in the pen and paper condition, and 50% of sessions during generalization across participants by watching the video recordings. IOA for the primary measure of fidelity of implementation were calculated by dividing the number of agreements on the Fidelity Checklist by the number of agreements plus disagreements and multiplying by 100 to obtain a percentage of agreement. IOA for duration were counted as correct if the total duration counted by each observer was within 10s of each other.

IOA data were calculated for Jill for 60% of sessions during the intelligent agent condition and averaged 95.38% (range 91.24%- 98.54%) for fidelity and 100% for duration. In the pen and paper condition, IOA data were taken during 60% of sessions and averaged 91.24% (range 90.51%- 91.97%) for fidelity and 100% for duration. IOA data were taken during 50% of generalization sessions and were 89.05% for fidelity and 100% for duration during the intelligent agent condition. IOA data were 84.67% for fidelity and 100% for duration during generalization for the pen and paper condition.

IOA were calculated for Yazmina for 60% of sessions during the intelligent agent condition and averaged 95.38% (range 94.89%- 95.62%) for fidelity and 100% for duration. In the pen and paper condition, IOA were taken during 60% of sessions and averaged 92.94% (range 88.32%- 95.62%) for fidelity and 100% for duration. IOA were taken during 50% of generalization sessions and was 96.34% for fidelity and 100% for duration during the intelligent agent condition. IOA data were 93.43% for fidelity and 100% for duration during generalization for the pen and paper condition.

IOA were calculated for Molly for 60% of sessions during the intelligent agent condition and averaged 94.65% (range 89.78%- 97.81%) for fidelity and 100% for duration. In the pen and paper condition, IOA were taken during 80% of sessions and averaged 93.25% (range 84.67%- 96.35%) for fidelity and 100% for duration. IOA were taken during 50% of generalization sessions and were 100% for fidelity and 100% for duration during the intelligent agent condition. IOA were 84.67% for fidelity and 100% for duration during generalization for the pen and paper condition.

For Shawna, IOA were calculated in 40% of sessions during the intelligent agent condition and averaged 92.34% (88.32%, 96.35%) for fidelity and 100% for duration. In the pen and paper condition, IOA were taken during 60% of sessions and averaged 94.89% (range 92.70%- 98.54%) for fidelity and 100% for duration. IOA data were taken during 50% of generalization sessions and were 95.62% for fidelity and 100% for duration during the intelligent agent condition. IOA were 91.97% for fidelity and 100% for duration during generalization for the pen and paper condition.

For Adriana, IOA were calculated in 40% of sessions during the intelligent agent condition and averaged 86.86% (85.40%, 88.32%) for fidelity and 100% for duration. In the pen and paper condition, IOA were taken during 40% of sessions and averaged 89.05% (83.94%, 94.16%) for fidelity and 100% for duration. IOA data were taken during 50% of generalization sessions and were 83.94% for fidelity and 100% for duration during the intelligent agent condition. IOA were 89.78% for fidelity and 100% for duration during generalization for the pen and paper condition.

Procedural fidelity consisted of a five-item questionnaire that was completed by the primary investigator following every session (see Appendix L). The questionnaire concerned the primary investigator's implementation of the intervention without offering feedback or guidance before, during or after the MSWO and the simulated learner's adherence to the written script. IOA on procedural fidelity were collected by the secondary observer for 52% of sessions in the intelligent agent condition, 60% of sessions in the pen and paper condition, and 50% of sessions during generalization across participants by watching the video recordings. IOA were calculated by taking the number of agreements on the questionnaire divided by the number of agreements plus disagreements (five total) and multiplied by 100 to obtain a percentage of agreement.

Procedural fidelity data were documented for 100% of sessions across all participants and averaged 100% across all participants for all conditions. IOA for procedural fidelity were collected in all the sessions where IOA on primary and secondary measures were collected and was 100% for all participants for all conditions.

Experimental Design

This study used a single-subject alternating treatment design (Ledford & Gast, 2018), consisting of two conditions: 1) pen and paper, 2) the intelligent agent application (GAINS). Each condition consisted of a minimum of five sessions for each participant. Prior to beginning the study, each adult participant was randomly assigned to begin in either the pen and paper or GAINS condition. By counterbalancing the order of interventions across participants, it assisted in controlling for sequence effects. Afterward, sessions were alternated between the two conditions with two to four days between each condition to control for rapid alternation effects. In addition, generalization probes with children were conducted at the beginning and the end of the study. After five sessions of each condition, the final generalization probe occurred. At this point, the study was concluded.

Procedures

MSWO Procedures

MSWO procedures were held consistent across both conditions of the study and the generalization probes. At the beginning of each session, the participant prompted the learner to sit at the table and the participant sat across or beside the learner. Each session began with a pre-session exposure condition, where the participant presented stimuli one at a time to the learner and allowed the learner to access the item for thirty seconds. If the learner interacted with (or consumed) the stimulus, that stimulus was used during the MSWO. If the learner did not interact with (or consume) the stimulus, that stimulus was removed from the assessment. This continued until five items were selected for use in the MSWO.

Next, the participant placed the five selected items in a line on the table about five inches apart within reach of the learner. The participant instructed the learner to “pick one.” The learner was given ten seconds to choose a stimulus. If the learner selected an item, he/she was given thirty seconds of access to the item (or able to consume the item, if edible). The participant recorded the learner’s response on the data sheet or on the tablet-based device. After thirty seconds, the item was removed from the immediate area. The remaining stimuli were rearranged by taking the item at the left end of the line and moving it to the right end and shifting the other stimuli, so they were equally spaced in an array in front of the child. The participant again instructed the learner to “pick one.” The learner was given ten seconds to choose an item and given thirty seconds of access to the item (or until consumed). Once again, the response was recorded, the item removed, and the remaining items rotated. The assessment continued in this manner until all items were selected or the learner makes no selection within ten seconds. If no selection was made, all remaining stimuli were removed, and the participant recorded “no selection" for those trials. If the learner selected more than one item, the learner was allowed to access the first item contacted. If two or more items were selected simultaneously, the learner was not allowed to access either item and the trial was reinitiated with the items rotated.

This procedure was repeated twice more so that the learner had access to all five stimuli in the array for a total of three times. Following the completion of the assessment, the participant finished filling out and scoring the data sheet (pen and paper condition) or reviewed the graphs provided by GAINS (the intelligent agent condition).

Generalization Probes

Generalization probes occurred for both the pen and paper condition and the GAINS condition at the beginning and the end of the study. During the generalization probes, the learner was a child participant, aged four-to-five years old. Prior to each generalization session, seven to eight stimuli that the child learner interacts with (or consumes) regularly were collected for the child based on an informal preference assessment using caregiver or participant report and free operant (FO) observation. Seven to eight stimuli were available to the participant, even though only five were used in the MSWO to allow for replacement stimuli if the child learner did not interact with (or consume) a stimulus during the pre-session exposure.

Pen and Paper Condition

During all sessions (except generalization probes) in the pen and paper condition, the learner was an adult simulated learner. The simulated learner was told which script (selected at random) to follow prior to beginning each session. Adult participants were given written instructions of the procedures for a MSWO preference assessment, adapted from DeLeon and Iwata (1996; Higgins et al., 2017; see Appendix G) as a job aid. The participants were given as much time as they needed to read the instructions and told to let the adult simulated learner know when they were ready to begin. They were allowed to keep and refer to the written instructions throughout the session. The participants were told to conduct the assessment to the best of their ability but were not given any further instructions or feedback. If they asked questions, the primary investigator would respond that they were to do whatever they thought was best. Adult participants were given the paper data sheet (adapted from Higgins et al., 2017; see Appendix H) to fill out and record the learner's selections. The session was recorded and scored by the primary investigator according to the Fidelity Checklist (see Appendix I). As soon as the learner sat down at the table to begin the assessment, a stopwatch was started and kept running until the participant had completed and scored the assessment to generate a hierarchy of preferences among the stimuli selected.

GAINS Condition

During all sessions (except generalization probes) in the GAINS condition, the learner was an adult simulated learner. The simulated learner was told which script (selected at random) to follow prior to beginning each session. The participant used the GAINS application on a tablet device with a Bluetooth headset. The application provided guidance on conducting the MSWO via audio instructions on the Bluetooth headset. It also provided visual guidance using words and diagrams on the screen of the tablet device. Prior to the first session using GAINS, each participant was given about five minutes to access the application on the tablet and complete a sample program not related to the study. This allowed time to work through any technical difficulties that participants may encounter and ensured that the application was working properly. The primary investigator did not provide any guidance, feedback or answers to questions before, during, or after each trial or session. The session was recorded, and the Fidelity

Checklist was filled out by the primary investigator. Since the GAINS application scores the MSWO results, the primary investigator asked the participant which item they would select to use in their program for the day to determine the participant's interpretation of the results. In addition, as soon as the learner sat down at the table to begin the assessment, a stopwatch was started and kept running until the participant had completed and scored the assessment to generate a hierarchy of preferences among the stimuli selected.

Social Validity Survey

When all adult participants completed the final generalization probe, they were given a copy of the System Usability Scale survey (Bangor et al., 2008; see Appendix J). They were asked to answer each question honestly and to the best of their ability.

The second measure of social validity evaluating the children's perspectives was administered orally to each participant following completion of each day they were involved int the study. The child participants were also shown a visual of a green smiley face, yellow face with a straight mouth, and red frowning face and allowed to respond to the questions by pointing to the faces instead of or in addition to responding verbally. Both the questions and the visuals are included as Appendix K.

Data Analysis

To understand how the data will be used to address each of the research questions, a brief overview for each question is presented.

Question 1*: *Will using intelligent agent technology to conduct a MSWO preference assessment increase preservice SLPs' fidelity of implementation when compared to pen and paper?

Data collected on the primary dependent variable (fidelity of implementation) were used to address this research question. Data are displayed in graphical format (see Figures 1, 2, and 3) and analyzed with visual analysis (Ledford & Gast, 2018). Descriptions of level, trend, variability, consistency, effect size, and immediacy are provided. Effect size was calculated using improvement rate difference (IRD; Parker et al., 2011), a measure of effect size using nonoverlapping data. An IRD of 0.00-0.72 is considered small, a score of 0.72-0.90 is considered moderate, and a score of 0.90 or above is considered large (Parker et al., 2011). If the data displayed an increasing trend, low variability, consistency, immediacy, and a moderate to high effect size in the intelligent agent condition over the pen and paper condition, it would demonstrate that the intelligent agent technology may increase fidelity of implementation.

Question 2: Will using intelligent agent technology to conduct a MSWO preference assessment decrease the amount of time spent in generating a preference hierarchy when compared to pen and paper?

The secondary dependent measure of duration was used to address this research question. The data are displayed in a graphical format (see Figures 4, 5, an 6) and the average duration for each condition with each participant was calculated. Data collected on the secondary dependent variable were also analyzed with visual analysis (Ledford & Gast, 2018). A comparison between conditions was be made by examining the data patterns of each condition and evaluating for differentiation. Descriptions of level, trend, variability, consistency, and immediacy are provided. If the intelligent agent condition displayed a shorter administration time, then it would demonstrate that the intelligent agent technology may lead to shorter assessment durations.

Question 3: Do preservice SLPs view the intelligent agent technology as having a higher treatment acceptability than pen and paper?

Data from the follow-up survey given to adult participants were analyzed to address this research question. This survey allowed respondents to choose from the following answer choices: strongly disagree, disagree, neutral, agree, and strongly agree. After completion of the survey, each response was assigned a numerical value ranging from 1 to 5 with 1 being “strongly disagree” and 5 being “strongly agree.” Responses were compiled, and results calculated using an average and range of responses for each question. A total of 24 items on the follow-up survey relate to the treatment acceptability of the intelligent agent technology.

Question 4: Do preservice SLPs view the intelligent agent technology as more effective at producing socially significant outcomes than pen and paper?

The remaining items (nine total) on the follow-up survey were used to analyze if the intelligent agent produces socially significant outcomes. Data were analyzed in the same manner as in the previous question, by assigning each response a numerical value and calculating the average and range for each question.

Question 5: Do child participants report a higher degree of treatment acceptability for use of intelligent agent technology over pen and paper?

To address this research question, a questionnaire was administered orally to each child participant following completion of each day they were involved in the study. The child participants were allowed to respond verbally and/or by pointing to a visual of a green smiley face, yellow face with a straight mouth, and red frowning face. Results from these questions were analyzed by calculating the percentage of responses in each category (i.e., smiley face, straight face, and frowning face) for each question.

Results from the study are presented in the next chapter, followed by a discussion of their implications and future research ideas in Chapter 5.

Chapter Four: Results

This chapter discusses the results of the study using visual analysis (Ledford & Gast, 2008) and descriptive statistics when applicable. First, this chapter presents the dependent measures of implementation fidelity and duration of assessment. These measures reflect the first two research questions. Next, the results of the follow-up survey for the adult participants are described, as they relate to research questions three and four. Finally, the results of the follow-up survey given to the child participants are presented, which correspond to the final research question.

Implementation Fidelity

Figures 1, 2, and 3 display a graphical representation of fidelity results for all five participants. Overall, participants averaged 93.97% (range 82.48%- 100%) fidelity while using the intelligent agent technology. They averaged 83.79% (range 62.04%- 98.54%) during the pen and paper condition. In addition, four out of five participants initially scored higher while using the intelligent agent technology than pen and paper. During generalization sessions with child participants, adult participants averaged 93.94% (range 83.21%- 100%) using the intelligent agent condition and 83.65% (range 68.61%- 95.62%) using pen and paper. In total, participants made errors while scoring the assessment 16 out of 35 times (45.71%) during the pen and paper condition, including four out of ten (40.00%) generalization sessions, which resulted in inaccurate identification of a preference hierarchy. No participants made errors in identifying the correct preference hierarchy during any of the sessions in the intelligent agent condition. Effect size was calculated using improvement rate difference (IRD; Parker et al., 2011), a measure of effect size using nonoverlapping data. An IRD of .00- 0.72 is considered small, a score of 0.72-

.90 is considered moderate, and a score of .90 or above is considered large (Parker et al., 2011). Two participants had an IRD score of 1.00, indicating a large effect size.

Jill

As seen in Figure 1, Jill averaged 92.60% (range 89.78%- 96.35%) fidelity while using the intelligent agent technology. She averaged 73.72% (range 62.04%- 81.75%) fidelity while using pen and paper. In terms of visual analysis, Jill's data for the intelligent agent condition illustrate a high level with little variability, no visible trend and a high degree of consistency. Initially, Jill scored higher in the intelligent agent condition. Jill's data for the pen and paper condition show a moderately high level with some variability, no visible trend and a moderate degree of consistency. IRD for Jill was 1.00, indicating a large effect size with no overlapping data points.

During the generalization sessions with child participants, Jill averaged 91.24% (89.78, 92.70) while using the intelligent agent technology. She averaged 81.39% (81.75%, 81.02%) during the pen and paper condition. Jill made errors when scoring the MSWO in three out of seven (42.86%) sessions during the pen and paper condition. In contrast, she was able to correctly identify the preference hierarchy in all seven sessions during the intelligent agent condition and made no errors (0%).

Yazmina

As depicted in Figure 1, Yazmina averaged 96.98% (range 89.05%-100%) fidelity while using the intelligent agent technology. She averaged 94.36% (range 82.48%-99.2%) fidelity while using pen and paper. In terms of visual analysis, Yazmina's data in the intelligent agent condition demonstrate a high level, a slightly increasing trend, and a high degree of consistency with no variability. Initially, Yazmina scored higher in the intelligent agent condition. Yazmina's data in the pen and paper condition illustrate a high level, no visible trend and a high degree of consistency with little variability. IRD for Yazmina was 0.29, indicating a small effect size.

In terms of generalization, Yazmina averaged 96.72% (94.89%, 98.54%) in the intelligent agent condition and 92.34% (91.97%, 92.7%) in the pen and paper condition during sessions with child participants. Yazmina scored the assessment incorrectly in two out of seven (28.57%) sessions during the pen and paper condition, including during one generalization session. However, Yazmina was able to correctly identify the preference hierarchy in all seven sessions using the intelligent agent technology and had an error rate of 0%.

Molly

As represented in Figure 2, Molly averaged 96.45% (range 89.05%- 100%) fidelity while using the intelligent agent technology. She averaged 85.71% (range 75.91%- 96.35%) fidelity while using pen and paper. In terms of visual analysis, Molly's data demonstrate a high level, slightly increasing trend, no variability and a high degree of consistency under the intelligent agent condition. In contrast, Molly's data during the pen and paper condition indicate a moderately high level, no clear trend, a small degree of variability and a moderate degree of consistency. Molly initially scored higher in the intelligent agent condition than the pen and paper condition. Molly's data obtained an IRD score of 0.71, indicating a small to moderate effect size.

During the generalization sessions with child participants, Molly averaged 97.45% (94.89%, 100%) in the intelligent agent condition and 78.47% (81.02%, 75.91%) in the pen and paper condition. Molly made errors when scoring the assessment in one out of seven (14.29%)

sessions during a generalization session. In contrast, Molly was able to identify the correct preference hierarchy in all seven sessions during the intelligent agent condition and made errors 0% of the time.

Shawna

As illustrated in Figure 2, Shawna averaged 96.77% (range 87.59%- 100%) fidelity while using the intelligent agent technology. She averaged 93.41% (range 89.78%- 97.81%) fidelity while using pen and paper. In terms of visual analysis, Shawna's data illustrate a high level, slightly increasing trend, no variability and a high degree of consistency in the intelligent agent condition. Shawna's data during the pen and paper condition demonstrate a high level, slightly decreasing trend, some variability, and a moderate degree of consistency. Initially, Shawna scored higher during the pen and paper condition than during the intelligent agent condition. IRD for Shawna was 0.57, indicating a small effect size.

Shawna scored 97.08% in both generalization sessions with child participants for the intelligent agent condition. 92.70% (95.62%, 89.78%) was her average during the pen and paper generalization sessions. Shawna made errors when scoring the assessment in three out of seven (42.86%) sessions during the pen and paper condition. Whereas she identified the correct preference hierarchy in all seven of sessions during the intelligent agent condition and made errors 0% of sessions.

Adriana

Figure 3 depicts the results of implementation fidelity for Adriana. She averaged 87.06% (range 82.48%- 91.24%) fidelity while using the intelligent agent technology. She averaged 71.74% (range 64.96%- 78.1%) fidelity while using pen and paper. In terms of visual analysis, Adriana's data for the intelligent agent condition show a high level, an increasing trend, no variability and a high degree of consistency. In contrast, Adriana's data during the pen and paper condition illustrate a moderately high level, relatively stable trend, minimal variability, and a high degree of consistency. Initially, Adriana scored higher in the intelligent agent condition than the pen and paper condition. Adriana's data obtained an IRD score of 1.00, which indicates a large effect size with no overlapping data points.

In terms of generalization, Adriana averaged 87.23% (83.21%, 91.24%) in the intelligent agent condition and 73.36% (68.61%, 78.10%) in the pen and paper condition during sessions with the child participants. Adriana made errors in scoring the assessment in all seven (100%) sessions during the pen and paper condition, including both generalization sessions. In contrast, Adriana made no errors (0%) at any time during the intelligent agent condition and was able to identify the preference hierarchy correctly each time.

Duration of Assessment

Figures 4, 5, and 6 represent a graphical depiction of duration results for all five participants. Overall, the assessments utilizing intelligent agent technology averaged 12 min 52.7s (range 8 min 32s – 17 min 36s) while the assessments that used pen and paper averaged 19 min 11.3s (range 12 min 19s – 28 min 32s). Additionally, all five participants initially performed the MSWO more quickly with the intelligent agent than the pen and paper assessment. During the generalization sessions with child learners, adult participants averaged 12 min 11.5s (range 8 min 32s- 17 min 36s) during the intelligent agent condition and 19 min 13.2s (range 12 min 19s – 28 min 32s) during the pen and paper condition in generalization sessions with child learners. Effect size for duration was measured using IRD. Four participants had IRD scores falling in either the large or moderate range.

Jill

As depicted in Figure 4, Jill averaged 14 min 20.4 s (range 11 min 40s -17 min 36s) during the intelligent agent technology and 18 min 37.4 s (range 15 min 1s – 23 min 33s) during the pen and paper condition. Visual analysis for Jill's data illustrates a positive low level, decreasing trend, little variability and a high degree of consistency in the intelligent agent condition. Whereas her data in the pen and paper condition demonstrate a moderate level, a slightly decreasing trend, some variability, and some degree of consistency. IRD for duration for Jill was 0.71, indicating a small to moderate effect size. Initially, Jill performed the MSWO assessment faster during the intelligent agent condition than during the pen and paper condition. During generalization sessions with child participants, Jill averaged 14 min 47.5s (17 min 36s, 11 min 59s) using the intelligent agent and 16 min 42s (15 min 1s, 18 min 23s) using the pen and paper.

Yazmina

As illustrated in Figure 4, Yazmina averaged 13 min 9 s (range 11 min 35s- 15 min 8s) during the intelligent agent technology and 17 min 52.6 s (range 14 min 48s – 23 min 28s) during the pen and paper condition. In terms of visual analysis, Yazmina's data under the intelligent agent condition show a positive low level, stable trend, minimal variability and a high degree of consistency. Initially, Yazmina performed more quickly under the intelligent agent condition. In contrast, her data during the pen and paper condition demonstrate a moderate level, a relatively stable trend, and a moderate degree of variability and consistency. IRD for Yazmina was 0.86, indicating a moderate effect size. In terms of generalization, Yazmina averaged 12 min 27s (11 min 35s, 13 min 19s) during the intelligent agent condition. She averaged 19 min 38.5s (23 min 28s, 16 min 19s) during the pen and paper condition with child participants.

Molly

As seen in Figure 5, Molly averaged 12 min 15.4 s (range 9 min 49s – 15 min 34s) and 21 min 13.4s (range 17 min 54s – 28 min 32s) during the intelligent agent technology and the pen and paper condition, respectively. In terms of visual analysis, Molly's data under the intelligent agent condition demonstrate a positive low level, a slight decreasing trend, a slight degree of variability and a high level of consistency. Initially, Molly was able to perform the MSWO more quickly under the intelligent agent condition. Her data during the pen and paper condition illustrate a moderate level, an increasing trend, some variability and a moderate degree of consistency. IRD for Molly was 1.00, which indicates a large effect size with no overlapping data points. In generalization sessions with child participants, Molly averaged 10 min 7s (9 min 49s, 10 min 25s) during the intelligent agent condition and 24 min 5.5s (19 min 39s, 28 min 32s) during the pen and paper condition.

Shawna

As shown in Figure 5, Shawna averaged 11 min 35.9 s (range 8 min 32s – 13 min 10s) and 17 min 10.3 s (range 12 min 19s – 27 min 42s) during the intelligent agent technology and the pen and paper condition, respectively. Shawna's data in the intelligent agent condition demonstrate a positive low level, a slight decreasing trend, some variability and a high degree of consistency. Initially, Shawna was able to complete the MSWO more quickly under the intelligent agent condition. Her data during the pen and paper condition illustrate a moderate level, a decreasing trend, a high degree of variability and a low level of consistency. IRD for duration for Shawna was 0.86, indicating a moderate effect size. During generalization sessions with child participants, Shawna averaged 9 min 39s (8 min 32s, 10 min 46s) and 13 min 52.5s (12 min 19s, 15 min 26s) using the intelligent agent and pen and paper respectively.

Adriana

As presented in Figure 6, Adriana averaged 13 min 2.9 s (11 min 22s – 15 min 1s) during the intelligent agent condition and 21 min 2.7s (range 18 min 4s – 24 min 1s) during the pen and paper condition. Adriana's data in the intelligent agent condition demonstrate a positive low level, a stable trend, little to no variability and a high degree of consistency. Initially, Adriana was able to perform the MSWO more quickly under the intelligent agent condition than under the pen and paper condition. Her data under the pen and paper condition illustrate a moderate level, a stable trend, limited variability and moderate degree of consistency. IRD for Adriana was 1.00, indicating a large effect size with no overlapping data points. In terms of generalization, Adriana averaged 13 min 56.5s (14 min 44s, 13 min 9s) during the intelligent agent condition and 21 min 53.5s (21 min 27s, 22 min 20s) during the pen and paper condition in sessions with child participants.

Follow-up Survey for Adult Participants

Following the completion of the last session, each adult participant was given a follow-up survey (see Appendix J) that contained thirty-three positively worded statements about the acceptability and social significance of the intelligent agent technology. This survey allowed respondents to choose from the following answer choices: strongly disagree, disagree, neutral, agree, and strongly agree. After completion of the survey, each response was assigned a numerical value ranging from 1 to 5 with 1 being "strongly disagree" and 5 being "strongly agree."

Treatment Acceptability

Of the 33-items on the survey, 24 of the items pertained to the treatment acceptability of the intelligent agent technology when compared to pen and paper self-instruction (i.e. research question 3). The results of these specific 24 items are reflected in Table 2. Overall, the responses averaged 4.54 with a range of 1 to 5. The lowest scoring items involved the use of the audio assistance on the application and included, "Audio is enough. It is not necessary to read the display" and "Audio assistance is useful," which scored a mean of 2.6 (range 1 -5) and 3.8 (range 3-5), respectively. All other items scored an average between 4.0 and 5.0. The highest scoring items involved the ease of interaction, including the ability to learn to use it and the ease of reading the display screen. One participant added a write-in comment, "I think that the app was far easier to use, and you do not have to do math!"

Socially Significant Outcomes

The remaining nine items on the follow-up survey pertained to whether the intelligent agent technology was more effective at producing socially significant outcomes than pen and paper self-instruction. The results of these nine specific items are listed in Table 3. Overall, the responses had a mean of 4.84 and a range of 4 to 5. All means of the items were between 4.6 and 5.

Follow-up Surveys for Child Participants

The social validity questionnaire for the children contained five items (see Appendix K). Although there were only two child participants, they completed the social validity questionnaire after every day they were involved with the study, resulting in two different survey responses for each child and a total of four responses. Table 4 represents the children's responses. The first question asked if the children liked having a choice in items to play with and to eat. The children answered that they did enjoy having choices 100% of sessions. When asked if they liked it when the teacher used the paper, the children responded that they did like it 75% of the time, but did not like it 25% of the time. When asked if they liked it when the teacher used the tablet, 75% indicated that they did and the other 25% were neutral. When asked if they preferred the tablet or the paper, 75% indicated a preference for the tablet, while 25% remained neutral. Finally, all child participants (100%) indicated that they had fun during the study.

Chapter five provides a discussion on the results, implications for practitioners, limitations of the study and future research directions.

Chapter Five: Discussion

The purpose of the study was to compare the use of typical pen and paper self-instruction to intelligent agent technology during preservice speech-language pathologists' (SLPs) training and implementation of multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessments. Fidelity of implementation, duration, and social validity were evaluated for both types of assessment. In addition, this study examined the perceptions of child participants in the use of intelligent agent technology and pen and paper stimulus preference assessments (SPA).

Limitations of Traditional Training Methods

As previously outlined in Chapter Two, even successful traditional SPA training methods can present significant limitations in the natural environment. First, past training methods often require an instructor who is experienced and knowledgeable about the procedures of the assessment. Even in studies the employed some form of technology, many still required experts for the initial training phase (Ausenhus & Higgins, 2019; Higgins et al., 2017; Weston et al., 2020). In the current study, participants were able to reach high levels of implementation fidelity upon initial use of the intelligent agent and did not require any modeling, feedback or instruction from an expert.

Second, it often requires hours of time to train individuals to perform the SPA to fidelity before they begin implementation. In previous research, training time varied among participants and methods, ranging from 30 minutes to three hours and in many cases, participants needed multiple training sessions to reach a high level of fidelity (Higgins et al., 2017; Lipschultz et al., 2015; Pence et al., 2012; Weldy et al., 2014; Wishnowski et al., 2017). In applied settings, this could represent a tremendous burden for practitioners and clinicians whose time is limited and valuable. In the current study, the participants required no initial training to use the GAINS application aside from five minutes of guidance on how to log-in and begin a session. All five participants were able to complete the MSWO with high levels of fidelity during the initial session.

A third limitation of traditional training methods is that skills are taught in contrived environments and require additional time and training to generalize to the intended target population, usually children (Ausenhus & Higgins, 2019; Graff & Karsten, 2012b; Higgins et al., 2017; Pence et al., 2012; Shapiro et al., 2016). Although initial training with simulated adults was successful for some participants, many studies had mixed results when attempting to transfer skills to the intended target population (Ausenhus & Higgins, 2019; Higgins et al., 2017; Shapiro et al., 2016; Pence et al., 2012). One major finding of this study was that when the intelligent agent technology was used with children, the adult participants maintained the same high levels of implementation fidelity.

Lastly, many studies have not collected long-term data on implementation fidelity so it is unknown if practitioners will be able to maintain fidelity over a long period of time (Graff & Karsten, 2012b; Hansard & Kazemi, 2018; Shapiro et al., 2016; Pence et al., 2012; Weldy et al., 2014). In studies that included follow-up data, not all participants were successful in maintaining their skills with high levels of fidelity (Rosales et al., 2015; Wishnowski et al., 2017). With traditional training methods, it would be necessary for skills to be monitored and retrained if fidelity fell below appropriate levels. Because the intelligent agent technology used in the current study provides the same instructions for the practitioner every time and there is no need for it to be removed, clinicians will be able to maintain the same high levels of implementation fidelity over time.

The results of the current study are consistent with previous research which suggests that pen and paper self-instruction alone is insufficient to produce consistent outcomes during SPA training and implementation (Graff & Karsten, 2012b). Even with the addition of a job aid that could be referenced during implementation, some professionals have remained unable to acquire the necessary skills to use SPA with fidelity, necessitating additional time and resources spent on training (Graff & Karsten, 2012b; Shapiro et al., 2016). Other technologies, such as telehealth and computer-aided instruction, have improved outcomes, but even successful interventions have presented with similar limitations as traditional pen and paper self-instruction methods, including the length of initial training phases and the requirement of an expert trainer (Hansard & Kazemi, 2018; Lipschultz et al., 2015; Miljkovic et al., 2015; Rosales et al., 2015; Weston et al., 2020; Wishnowski et al., 2017). The current study provides a technological solution that may overcome these limitations and achieve high levels of fidelity during SPA implementation.

This chapter aims to provide an analysis of the findings of each research question presented in this study. It also details the limitations of the current study. In addition, clinical implications and recommendations for future research are discussed.

Question 1: Will using intelligent agent technology to conduct a MSWO preference assessment increase preservice SLPs' fidelity of implementation when compared to pen and paper?

The results of this study indicate that intelligent agent technology increased preservice SLPs' fidelity of implementation when compared to pen and paper for all five participants. A visual analysis and effect size measure determined that the increase in fidelity was statistically significant for two (i.e. Jill and Adriana) out of five participants. For the remaining three participants (i.e. Yazmina, Molly, and Shawna), although there was an increase in implementation fidelity, it was not considered a statistically significant difference in comparison with the pen and paper condition.

Implementation Fidelity

Both Jill and Adriana had an IRD score of 1.00, which indicates not only a large effect size, but that their data contained no overlapping data points. This means that their lowest score in the intelligent agent condition was higher than their highest score in the pen and paper condition. Ceiling effects made determining a difference in implementation fidelity for Yazmina, Molly, and Shawna difficult, due to high levels of responding across both the pen and paper condition and the intelligent agent conditions. One reason for such high levels of implementation fidelity seen across both conditions for Shawna may be due to her higher level of education. She was the only graduate student included in this study. Perhaps because she was more accustomed to working independently and has been given more autonomy to read and follow written instructions without explicit training, she was better able to implement the SPA initially under the pen and paper condition. It is significant to note that the remaining four participants (i.e. Jill, Yazmina, Molly, and Adriana) initially scored higher while using the intelligent agent technology than pen and paper self-instruction. This could indicate that the intelligent agent is better able to immediately produce higher levels of fidelity for most professionals.

An important finding of the current study is the high levels of implementation fidelity demonstrated during the generalization sessions with children in the intelligent agent condition. For all five participants, high levels of fidelity were demonstrated with child participants without the need for additional training or feedback from an expert. As previously stated, many traditional training methods have been unable to produce these results.

Scoring and Interpretation Errors

It is worth noting the difference in interpretation and scoring errors between the two conditions. During the pen and paper condition, the results of the SPA were scored incorrectly in 16 out of 35 (45.71%) times. This included one participant (i.e. Adriana) who was unable to correctly score the SPA during any of the pen and paper condition sessions. In each of these instances where an incorrect score was given, participants produced an incorrect preference hierarchy and were unable to identify most preferred and least preferred items correctly. In an applied setting, these scoring and interpretation errors could lead to unproductive treatment programs due to practitioners making erroneous decisions about which items to use as a part of instructional programs. Essentially, miscalculating the scores of the SPA produces invalid results. Even the participants with high levels of implementation fidelity (i.e. Yazmina, Molly, and Shawna) scored the SPA incorrectly in at least one pen and paper session. In contrast, all participants were able to accurately identify the correct preference hierarchy for every SPA conducted using the intelligent agent technology. One participant also noted on the follow-up survey that she preferred the intelligent agent condition because it was “far easier to use, and you do not have to do math!” This suggests that participants may have found the scoring of the pen and paper assessment particularly challenging and unenjoyable.

Question 2: Will using intelligent agent technology to conduct a MSWO preference assessment decrease the amount of time spent in generating a preference hierarchy when compared to pen and paper?

The current study's results indicate that use of the intelligent agent technology to conduct MSWO preference assessments decreased the amount of time spent in generating a preference hierarchy when compared to pen and paper for all five participants. A visual analysis and effect size measure determined that the decrease in duration had a moderate to large statistical significance for four (i.e. Molly, Yazmina, Shawna, and Adriana) out of five participants. For the remaining participant (i.e. Jill), although there was a decrease in implementation duration, it was not considered statistically significant when compared to the pen and paper condition.

On average, participants completed the MSWO using the intelligent agent 7 minutes faster than pen and paper methods. This represents a 32.84% reduction in duration of assessment. Both Molly and Adriana had an IRD score of 1.00, which indicates both a large effect size and no overlapping data points. For these two participants, their longest session using the intelligent agent was still shorter than their fastest session using the pen and paper assessment. Additionally, all five participants initially performed the MSWO more quickly with the intelligent agent than the pen and paper assessment. These results remained consistent during generalization sessions with child participants. In these sessions, adult participants were able to generate a preference hierarchy with the intelligent agent on average 7 minutes (32.84%) faster than with pen and paper. These results are especially significant given that the pen and paper condition resulted in an inaccurate preference hierarchy 45% of the time. One could argue that a preference hierarchy was never obtained during these sessions.

Question 3: Do preservice SLPs view the intelligent agent technology as having a higher treatment acceptability than pen and paper?

Based on the results of the follow-up survey, it was determined that preservice SLPs view the intelligent agent technology as having a higher treatment acceptability than pen and paper self-instruction, as represented in Table 2. The items related to treatment acceptability had an average response of 4.54 out of 5. The only items on the survey with any responses lower than 4 involved the use of the audio assistance via the Bluetooth headset. Anecdotally, many of the participants chose not to wear the Bluetooth headset after the first intelligent agent session or turned the audio down on the tablet so they could not hear the audio. On all items that discussed having a choice between the two assessment methods, all participants indicated that given a choice, they would prefer to use the intelligent agent technology over pen and paper assessments. Additionally, the highest scoring items involved the ease of interaction, including the ability to learn to use the application and the ease of reading the display screen. Based upon these findings, it is determined that the participants viewed the intelligent agent as having a high degree of treatment acceptability.

Question 4: Do preservice SLPs view the intelligent agent technology as more effective at producing socially significant outcomes than pen and paper?

Based on the results of the follow-up survey, the preservice SLPs view the intelligent agent technology as more effective at producing socially significant outcomes than pen and paper self-instruction, as documented in Table 3. All items scored between a 4 and a 5 out of 5, with an overall mean of 4.84. All participants strongly agreed that the intelligent agent technology improved SPA administration, enhanced their job performance, and that SPA were useful. These results indicate that intelligent agent technology was viewed as having a high degree of social significance.

Question 5: Do child participants report a higher degree of treatment acceptability for use of intelligent agent technology over pen and paper?

To evaluate the child participants' views about the treatment acceptability of the intelligent agent technology, each child was given a follow-up survey after each day they were involved in the study. The results indicate that the children reported a higher degree of treatment acceptability for the use of intelligent agent technology over pen and paper, as reflected in Table 4. Although most of the time, the children said they liked it when the teacher used both the pen and paper and the tablet, one child indicated that he did not like it when the teacher used the pen and paper. In the same survey, the child was neutral toward the use of the tablet. In addition, when asked if they preferred the tablet or the paper, most indicated a preference for the tablet and one remained neutral. No children indicated a preference for the pen and paper at any point during the study. These results reflect a relatively high degree of treatment acceptability for the intelligent agent technology.

Clinical Implications

Professionals working with young children must make a myriad of decisions every day for each child they encounter about which items to include in educational programs. Because using preferred items during instruction can increase social engagement and skill acquisition, professionals, such as SLPs, can leverage these items to achieve increased communication development (Hu et al., 2021; Logan & Gast, 2001; Phillips et al., 2017). Improving social communication and interaction early in life can have tremendous impact on quality of life and independence in adulthood (Drmic et al., 2018). For this reason, identifying preferred items is an essential and common practice among SLPs and other professionals (Bishop & Kenzer, 2012; Graff & Karsten, 2012a). Research has consistently demonstrated that professionals are unable to accurately identify these items without the use of SPA (Cote et al., 2007; King, 2016). In addition, after using formal SPA, professionals have found them more effective and easier to use than other informal methods of reinforcer identification (King, 2016). Therefore, providing SLPs and other professionals with a means of easily and quickly performing regular SPA could be considered an essential job function and foundational to the implementation of many evidence-based practices, such as reinforcement (Graff & Karsten, 2012a; Hume et al., 2020).

Increasing Professionals' Procedural Knowledge

According to a survey of professionals, one of the most significant barriers to implementing SPA was lack of training and procedural knowledge (Graff & Karsten, 2012a). Previous training methods have relied heavily on the presence and availability of experts to train individuals. Many professionals live in areas where these experts are not available. Additionally, the fees that experts charge for hours spent in training may exceed the financial resources of some professional organizations. The results of this study illustrate that with the use of intelligent agent technology, individuals may not need any training or feedback from experts to be able to implement SPA with consistently high levels of fidelity. All five participants were able to complete the SPA with fidelity at or above 80% with the initial use of the intelligent agent technology without any training or consultation from an expert. Technologies that eliminate the need for experts would be of particular value in areas where experts are limited, while simultaneously being cost-effective to reach a greater number of individuals.

Saving Professionals' Time

Another major barrier to regular implementation of SPA identified by professionals was a lack of time (Graff & Karsten, 2012a). In the current study, implementors saved about 7 minutes per assessment with the use of intelligent agent technology. Although perhaps unsurprising given the ease of scoring and data collection, this finding has immense practical value for practitioners. SPA are generally the first step in a much longer program and should be used regularly to identify shifting motivations throughout clinical treatment (Bishop & Kenzer, 2012). If an SPA is conducted daily, the use of the intelligent agent technology would save a practitioner over 35 minutes per week, or over two hours per month. This represents a significant amount of time that could be utilized in teaching the child new skills, while still allowing the practitioner to identify the most effective reinforcers in the moment.

Instructional time is perhaps even more valuable to SLPs, who may only see a child for one to two hours per week (Gillon et al., 2017). The median caseload size for a school-based SLP is 48 children. They spend on average 22 hours a week in direct intervention with clients, which is around 50% of their time (ASHA, 2022). This means that if they are spending 22 hours a week with 48 students, they are only seeing each student on average less than half an hour a week. Their remaining time is spent completing documentation, performing diagnostic assessments, consulting with families and stakeholders, and supervising preservice SLPs. Because they have such high caseload sizes and so many demands on their time, SLPs could really benefit from time-saving technologies, such as the intelligent agent technology used in this study.

In addition to saving time during implementation, the intelligent agent technology can save time by eliminating the need for individual training before implementation. By providing directions during each step of the SPA while the professional is working with the child learner, it allows the implementer to focus on the learner, dedicating work and time resources to the learner rather than spending them on independent training (Nosek, 2017). All participants in the current study were able to meet a high level of fidelity upon initial use of the intelligent agent technology during implementation. There was no need to spend time training the participants before beginning to implement the SPA. Because high levels of fidelity were maintained throughout the entire study, there was no need for additional training. Considering that training in previous studies required between thirty minutes to three hours before initial SPA implementation, this represents a significant reduction in training time and resources, further reducing the barrier of lack of time to regular SPA implementation.

Increasing Accuracy of SPA Results

One of the most significant findings of the current study was that participants reported inaccurate preference hierarchies for almost half of the pen and paper assessments. Preference hierarchies are an essential component of arranging effective learning environments, are unique to each individual, and change on a daily basis (Chappell et al., 2009). One of the main advantages of preference hierarchies is that if one item loses its reinforcing efficacy, a professional can easily transition to the next item in the hierarchy as a learner's MO shifts throughout the day (Chappell et al., 2009; Kang et al., 2013). If these hierarchies are not identified correctly, professionals will not be able to make effective treatment decisions. This could result in a lack of skill acquisition for a learner, if an item given as a consequence to a desired behavior lacks reinforcing efficacy. Intelligent agent technology could prevent professionals from making these erroneous decisions by assisting them in identifying preference hierarchies with significantly higher rates of accuracy than traditional methods.

Respecting Professionals' Preferences

Given the difficulty in scoring and the increased duration of assessment time, it is perhaps unsurprising that the results of the follow-up survey indicated that participants found the intelligent agent technology to be more effective, easier to use, and more accurate than pen and paper self-instructional methods. Because SPA should be conducted regularly and by such a large number of professionals (Graff & Karsten, 2012a), having a method that is preferred and easier could be a tremendous resource. The increase in effectiveness, ease, and accuracy might also result in professionals conducting SPA more regularly, producing improved instructional outcomes for their students and clients. In addition, providing professionals with the tools to perform their jobs more consistently, accurately, easily and quickly, could significantly reduce the stresses they face, leading to higher job satisfaction and retention rates.

Research demonstrates that the preferences of young children shift rapidly over times in often unpredictable ways (Bishop & Kenzer, 2012; DeLeon et al., 2001). This can further complicate the decisions that practitioners need to make, producing an even greater need for conducting SPA frequently and making a tool for quick, easy, accurate SPA even more valuable to professionals. According to the results of this study, intelligent agent technology may be able to assist professionals in staying current on the preferences of their clients or students and maintain a higher level of engagement throughout the learning process.

Limitations

Carryover Effects

One significant limitation of the current study is the possibility of carryover effects from one treatment condition to the other due to the nature of the alternating treatment design.

Carryover effects are a type of multitreatment interference where one experimental condition may influence the performance under another experimental condition due to the similar nature of treatment conditions (Ledford & Gast, 2018). It remains possible that instructions in one condition could have affected performance and the way that instructions were interpreted or remembered in other treatment conditions. However, four out of five participants had higher initial fidelity scores under the intelligent agent condition, which suggests that there may not have been significant carryover effects. Other types of multitreatment interference, such as rapid alternation effects and sequence effects were controlled for by alternating the initial treatment condition across participants and allowing at least forty-eight hours between sessions.

Participant Limitations

Another limitation is the use of a first-year graduate student (i.e. Shawna) and upperclassmen (i.e. Jill, Yazmina, and Adriana) as adult participants. Fidelity scores may have been higher throughout both conditions due to their higher levels of education. The intelligent agent technology is designed for use with individuals with lower levels of education, such as registered behavior technicians. Despite their higher levels of education, they all still committed errors in scoring during the pen and paper assessment and duration of assessment was much shorter in the intelligent agent condition for all of them, which may have practical application even for users with higher levels of education.

The generalizability of this study is further limited by the child population to which the primary investigator had access. Both child participants presented with vocal capabilities and typical development. Therefore, results may not generalize to child populations with IDD or limited communication abilities.

Observer Limitations

An additional limitation is that the primary and secondary observer could not be blind to treatment conditions because it was clear on the video recording if the participant was using a tablet or a pen and paper. In order to avoid as much bias as possible, the secondary observer completing interobserver agreement did not learn about the application or any other studies that have been completed with the application. However, the primary observer was very familiar with the application and had previously conducted a study using a different program of the application (Griffen et al., accepted, pending revisions).

Inconsistencies with Child Participants

The inconsistent use of edible stimuli may have caused a difference in the duration of some generalization sessions. Edible stimuli were chosen due to the behavioral issues of denied access for one of the child participants. Although, this difference may have been offset because it applied for that child participant during both the intelligent agent condition and the pen and paper condition for different participants. Additionally, the child participants did not always play with the selected item for a full 30 seconds and handed it back to the adult participant early during multiple sessions. However, the simulated learner scripts included no selection of an item, as well. Duration times for generalization sessions and adult simulated learner sessions are comparable so it is possible that neither of these issues had an impact on duration data.

Future Research Directions

This study was a preliminary study to evaluate the effectiveness of intelligent agent technology for training and implementation of MSWO. Future research could extend these findings by using a different population of professionals or caregivers as implementors. Given that caregiver report of reinforcers is characteristically low (Cote et al., 2007; Didden & de Moor, 2004; Verschuur et al., 2011; Windsor et al., 1994), the intelligent agent application may allow caregivers to identify more effective reinforcers and preferred stimuli. As previously stated, the current study included participants that represent a fairly high level of education and training that some professionals working with children might not have. Future research could evaluate if individuals with at or below a high school level education would be able to perform the SPA with the same high levels of fidelity using the intelligent agent technology.

This study only included five individuals as adult participants, which represents a small scale of the population of professionals that work with young children. Additionally, the participants in this study were all female between the ages of twenty and twenty-two, which represents a very restricted population. Future research should examine the use of intelligent agent technology with older individuals, who may have less experience with technology and more barriers to its use. Including males as participants would extend the findings of this study, as well. Although, this study included two participants of color, a greater diversity of ethnicities would provide additional generalizability. More repetitions are needed to add validity to these findings and be able to generalize results across the population of professionals that work with young children.

The current study did not include a measure of reinforcer assessment to evaluate if the items identified as preferred during the SPA were efficacious as reinforcers. However, previous research has demonstrated that there is a high correlation between items identified as preferred during SPA and their ability to function as reinforcers (Call et al., 2012; DeLeon et al., 2009; DeLeon & Iwata, 1996; Fisher et al., 1992; Glover et al., 2008; Mangum et al., 2012; Paclawskyj

& Vollmer, 1995; Roane et al., 1998). Future studies could incorporate this type of measure to add validity to the intelligent agent technology.

Although this study did attempt to generalize results of training to child populations, more research is needed in testing intelligent agent technology with children with IDD or limited communication abilities. Due to the primary investigator's limited access to child participants, MSWO sessions were conducted MSWO outside of typical learning programs and educational sessions. Future research could assess if adult participants are able to maintain high levels of fidelity during daily SPA conducted as a part of a typical school or therapy session.

In addition, research could examine if participants achieve the same high levels of fidelity during other types of SPA, such as paired stimulus, multiple stimulus with replacement, or free operant preference assessments, while using intelligent agent technology. In applied settings, professionals need to be able to conduct a variety of SPA with high levels of fidelity so evaluating the use of intelligent agent technology for additional forms of SPA holds high practical value for practitioners.

Conclusion

This study demonstrated that intelligent agent technology produced higher levels of implementation fidelity with shorter assessment durations than pen and paper preference assessments for five preservice SLPs. Results demonstrate that participants were able to generalize skills to child participants with comparatively high levels of fidelity using the intelligent agent technology. Pen and paper assessments resulted in a significant number of errors in scoring, producing inaccurate preference hierarchies almost half the time. In contrast, the intelligent agent produced an accurate preference hierarchy every time that was easily interpreted by participants.

The intelligent agent technology was viewed as having a higher treatment acceptability and producing more socially significant outcomes than the pen and paper assessment by all five adult participants. All participants responded that if given a choice, they would continue to use the intelligent agent technology to complete SPA. In addition, they all strongly agreed that the intelligent agent improved SPA administration, enhanced their job performance, and was useful. Child participants reported a higher degree of treatment acceptability for the intelligent agent technology than the pen and paper assessment. These results hold significant value and meaning for professionals, such as SLPs, working with young children.

Professionals have cited a lack of procedural knowledge and a lack of time as the greatest barriers to regular SPA implementation (Graff & Karsten, 2012a). Intelligent agent technologies, such as the one used in this study, could overcome these barriers by providing time effective training methods that do not require modeling, feedback, or interaction from an expert. The training methods used in the current study occurred during the implementation of the SPA and would not require professionals to be removed from the children and settings they work in for training. Participants reported that the instructions provided by the intelligent agent were easy to read and follow. SLPs and other professionals would save additional time by being able to conduct SPA more quickly and effectively while using intelligent agent technology.

All participants displayed difficulty in scoring the pen and paper assessments and reported a preference for the intelligent agent technology over the pen and paper assessment. If professionals view intelligent agent technology as easier, more accurate, and more effective, they may be more likely to conduct SPA more regularly. This could lead to improved instructional methods and outcomes for children and significantly reduce the stress of the SLPs and other professionals that work with them.

Intelligent agent technology addresses the limitations presented by traditional SPA training methods by not requiring an expert to conduct or oversee training. In addition, it has the potential to eliminate, or at least significantly reduce, initial training time for SPA implementation. Participants in this study were able to conduct the SPA with high levels of fidelity upon initial use of the intelligent agent technology without any independent training. Participants did not require any additional training to generalize high levels of fidelity with an adult simulated learner to young children. Finally, intelligent agent technology provides the same instructions for the user every time it is used. Because there is no need to remove it, high levels of implementation fidelity can be maintained over an extended period. Implications of using intelligent agent technology include many advantages for SLPs, other professionals, and stakeholders working with young children. Further research should be conducted to determine if intelligent agent technology is effective in increasing SPA implementation fidelity in caregivers and teachers, in more natural settings, and with children with IDD or communication difficulties.

References

Alcalay, A., Ferguson, J.L., Cihon, J.H., Torres, N., Leaf, J.B., Leaf, R., McEachin, J., Schulze, K.A., & Rudrud, E.H. (2019). Comparing multiple stimulus preference assessments without replacement to in-the-moment reinforcer analysis on rate of responding. *Education and Training in Autism and Developmental Disabilities, 54*(1), 69-82.

Alexander, J. L., Ayres, K. M., & Smith, K. A. (2015). Training teachers in evidence-based practice for individuals with autism spectrum disorder: A review of the literature. *Teacher Education and Special Education*, *38*(1), 13–27. doi:10.1177/0888406414544551

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Speech-Language-Hearing Association. (2022). *2022 schools survey report: SLP caseload and workload characteristics.* www.asha.org/Research/memberdata/Schools-Survey/

Ausenhus, J.A., & Higgins, W.J. (2019). An evaluation of real-time feedback delivered via telehealth: Training staff to conduct preference assessments. *Behavior Analysis in Practice, 12*, 643-648. https://doi.org/10.1007/s40617-018-00326-1

Bangor, A., Kortum, P.T., & Miller, J.T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction, 24*(6), 574-594. doi: 10.1080/10447310802205776

Baum, W.B. (2004). *Understanding behaviorism*. (2nd ed.). Wiley-Blackwell.

Bishop, M.R. & Kenzer, A.L (2012). Teaching behavioral therapists to conduct brief preference assessments during therapy sessions. *Research in Autism Spectrum Disorders, 6*, 450-457. doi:10.1016/j.rasd.2011.07.005

Brodhead, M.T., Abel, E.A., Al-Dubayan, M.N., Brouwers, L., Abston, G.W., & Rispoli, M.J. (2016). An evaluation of a brief multiple-stimulus without replacement preference assessment conducted in an electronic pictorial format. *Journal of Behavior Education, 25*, 417-430. doi: 10.1007/s10864-016-9254-3

Brodhead, M.T, Abston, G.W., Mates, M., & Abel, E.A. (2017). Further refinement of video-based brief multiple-stimulus without replacement preference assessments. *Journal of Applied Behavior Analysis, 50*(1), 170-175. doi: 10.1002/jaba.358

Butler, C., Kobylarz, A., Zaki-Scarpa, C., LaRue, R.H., Manente, C., & Kahng, S. (2021). Differential reinforcement to decrease stereotypy exhibited by an adult with autism spectrum disorder. *Behavioral Interventions, 36*(1), 261-270.

Call, N.A., Trosclair-Lasserre, N.M., Findley, A.J., & Reavis, A.R. (2012). Correspondence between single versus daily preference assessment outcomes and reinforcer efficacy under progressive-ratio schedules. *Journal of Applied Behavior Analysis, 45*(4), 763-777. doi: 10.1901/jaba.2012.45-763

Cannella-Malone, H.I., O'Reilly, M.F., & Lancioni, G. (2005). Choice and preference assessment research with people with severe to profound developmental disabilities: A review of the literature. *Research in Developmental Disabilities, 26*, 1-15. doi: 10.1016/j.ridd.2004.01.006

Cannella-Malone, H.I., Sabielny, L.M., Jimenez, E.D., & Miller, M.M. (2013). Pick one! Conducting preference assessments with students with significant disabilities. *Teaching Exceptional Children, 45*(6), 16-23.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33*(3), 353-357. doi: 10.1901/jaba.2000.33-353.

Carter, L., Harper, J.M., & Luiselli, J.K. (2019). Dental desensitization for students with autism spectrum disorder through graduated exposure, reinforcement, and reinforcement-fading. *Journal of Developmental & Physical Disabilities, 31*(2), 161-170. doi:10.1007/s10882-018-9635-8

Chappell, N., Graff, R.B., Libby, M.E., & Ahearn, W.H. (2009). Further evaluation of the effects of motivating operations on preference assessment outcomes. *Research in Autism Spectrum Disorders, 3*, 660-669. doi:10.1016/j.rasd.2009.01.002

Chazin, K.T. & Ledford, J.R. (2016). *Paired stimulus preference assessment*. Evidence-based instructional practices for young children with autism and other disabilities. http://ebip.vkcsites.org/paired-stimulus

Cooper, J.O., Heron, T.E., & Heward, W.L. (2020). *Applied behavior analysis*. Pearson Education Limited.

Conine, D.E., Morris, S.L., Kronfli, F.R., Slanzi, C.M., Petronelli, A.K., Kalick, L., & Vollmer, T.R. (2021). Comparing the results of one-session, two-sessions, and three-session MSWO preference assessments. *Journal of Applied Behavior Analysis, 54*(2), 700-712. doi: 10.1002/jaba.808

Cote, C.A., Thompson, R.H., Hanley, G.P., & McKerchar, P.M. (2007). Teacher report and direct assessment of preferences for identifying reinforcers for young children. *Journal of Applied Behavior Analysis, 40*(1), 157-166. doi: 10.1901/jaba.2007.177-05. doi: 10.1901/jaba.2007.177-05

Creswell, J.W., & Creswell, J.D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications, Inc.

Davis, C.J., Brock, M.D., McNulty, K., Rosswurm, M.L., Bruneau, B., & Zane, T. (2010). Efficacy of forced choice preference assessment: Comparing multiple presentation techniques. *Behavior Analyst Today, 10*(3-4), 440-445.

DeLeon, I.G., Fisher, W.W., Rodriguez-Catter, V., Maglieri, K., Herman, K., & Marhefka, J.M. (2001). Examination of relative reinforcement effects of stimuli identified through pretreatment and daily brief preference assessments. *Journal of Applied Behavior Analysis, 34*(4), 463-473. doi: 10.1901/jaba.2001.34-463

DeLeon, I.G., Frank, M.A., & Gregory, M.K. (2009). On the correspondence between preference assessment outcomes and progressive-ratio schedule assessments of stimulus value. *Journal of Applied Behavior Analysis, 42*(3), 729-733. doi: 10.1901/jaba.2009.42-729

DeLeon, I.G. & Iwata, B.A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*(4), 519-533. doi: 10.1901/jaba.1996.29-519

Didden, R., & de Moor, J. (2004). Preference assessment in toddlers with mild developmental and physical disabilities: A comparative study. *Journal of Developmental and Physical Disabilities, 16*(1), 107-116. doi:10.1023/B:JODD.0000010042.07391.08

Drmic, I.E., Szatmari, P., & Volkmar, F. (2018). Life course health development in Autism Spectrum Disorders. In N. Halfon, C.B. Forrest, R.M. Lerner, & E.M. Faustman (Eds.), *Handbook of life course health development* (pp. 237-274). Springer.

Fisher, W.W., Piazza, C.C., Bowman, L.G., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment to enhance reinforcer identification. *American Journal on Mental Retardation, 101,* 15-25.

Fisher, W., Piazza, C.C., Bowman, L.G., Hagopian, L.P., Owens, J.C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*(2), 491-498.

Fritz, J.N., Roath, C.T., Shoemaker, P.T., Edwards, A.B., Hussein, L.A., Vilante, N.K., Langlinais, C.A., & Rettig, L.A. (2020). Validity of the multiple-stimulus without replacement preference assessment for edible items. *Journal of Applied Behavior Analysis, 53*(3), 1688-1701. doi: 10.1002/jaba.703

Fuhrman, A.M., Fisher, W.W., & Greer, B.D. (2016). A preliminary investigation on improving functional communication training by mitigating resurgence of destructive behavior. *Journal of Applied Behavior Analysis, 49*(4), 884-899. doi: 10.1002/jaba.338

Gillon, G., Hyter, Y., Fernandes, F.D., Ferman, S., Hus, Y., Petinou, K., Segal, O., Tumanova, T., Vogindroukas, I., Westby, C., Westerveld, M. (2017). International survey of speech-language pathologists' practices in working with children with autism spectrum disorder. *Folia Phoniatrica et Logopaedica, 69*, 8-19. doi: 10.1159/000479063.

Glover, A.C., Roane, H.S., Kadey, H.J., & Grow, L.L. (2008). Preference for reinforcers under progressive and fixed-ratio schedules: A comparison of single and concurrent arrangements. *Journal of Applied Behavior Analysis, 41*(2), 163-176. doi: 10.1901/jaba.2008.41-163

Gottschalk, J.M., Libby, M.E., & Graff, R.B. (2000). The effects of establishing operations on preference assessment outcomes. *Journal of Applied Behavior Analysis, 33*(1), 85-88. doi: 10.1901/jaba.2000.33-85.

Graff, R.B. & Ciccone, F.J. (2002). A post hoc analysis of multiple-stimulus preference assessment results. *Behavioral Interventions, 17*, 85-92. doi:10.1002/bin.107

Graff, R.B., & Karsten, A.M. (2012a). Assessing preferences of individuals with developmental disabilities: A survey of current practices. *Behavior Analysis in Practice, 5*(2), 37-48. doi: 10.1007/BF03391822.

Graff, R.B., & Karsten, A.M. (2012b). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis, 45*(1), 69-82. doi: 10.1901/jaba.2012.45-69

Griffen, B., Lorah, E.R., Caldwell, N.S., Hantula, D., Nosek, J., Tincani, M., & Lemley, S. (accepted, pending revisions). The effects of intelligent agent technology on implementors' fidelity of instructional strategies during handwashing acquisition in children with autism. *Journal of Physical and Developmental Disorders.*

Hanley, G.P., Iwata, B.A., Roscoe, E.M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39*(2), 189-202. doi: 10.1901/jaba.2006.163-04

Hansard, C., & Kazemi, E. (2018). Evaluation of video self-instruction for implementing paired-stimulus preference assessments. *Journal of Applied Behavior Analysis, 51*(3), 675-680. doi: 10.1002/jaba.476

Higgins, W.J., Luczynski, K.C., Carroll, R.A., Fisher, W.W., & Mudford, O.C. (2017). Evaluation of a telehealth training package to remotely train staff to conduct a preference assessment. *Journal of Applied Behavior Analysis, 50*(2), 238-251. doi: 10.1002/jaba.370

Horner, R.H., Carr, E.G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-179.

Hu, X., Lee, G.T., Watkins, L., & Jiang, Y. (2021). Combining preferred activities with peer support to increase social interactions between preschoolers with ASD and typically developing peers. *Journal of Positive Behavior Interventions, 23*(4), 272-287. doi: 10.1177/1098300720987547

Hume, K., Steinbrenner, J.R., Odom, S.L., Morin, K.L., Nowell, S.W., Tomaszewski, B., Szendrey, S., McIntyre, N.S., Yucesoy-Ozkan, S., & Savage, M.N. (2020). Evidence-based practices for children, youth, and young adults with autism: Third generation review. *Journal of Autism and Developmental Disorders, 51*, 4013-4032. https://doi.org/10.1007/s10803-020-04844-2

Johnston, J. (2014). *Radical behaviorism for ABA practitioners*. Sloan Educational Publishing.

Kang, S., Lang, R.B., O'Reilly, M.F., Davis, T.N., Machalicek, W., Rispoli, M.J., & Chan, J.M. (2010). Problem behavior during preference assessments: An empirical analysis and practical recommendations. *Journal of Applied Behavior Analysis, 43*(1), 137-141. doi: 10.1901/jaba.2010.43-137

Kang, S., O'Reilly, M., Lancioni, G., Falcomata, T.S., Sigafoos, J., & Xu, Z. (2013). Comparison of the predictive validity and consistency among preference assessment procedures: A review of the literature. *Research in Developmental Disabilities, 34*, 1125-1133. https://doi.org/10.1016/j.ridd.2012.12.021

Karsten, A.M, Carr, J.E., & Lepper, T.L. (2011). Description of a practitioner model for identifying preferred stimuli with individuals with autism spectrum disorders. *Behavior Modification, 35*(4), 347-369. doi: 10.1177/0145445511405184

King, S.A. (2016). Multiple-stimulus without replacement preference assessment for students at risk for emotional disturbance. *Journal of Behavioral Education, 25*, 431-454. doi:10.1007/s10864-016-9256-1

Kirkpatrick, M., Akers, J., & Rivera, G. (2019). Use of behavioral skills training with teachers: A systematic review. *Journal of Behavioral Education, 28*, 344-361. https://doi.org/10.1007/s10864-019-09322-z

Kodak, T., Fisher, W.W., Kelley, M.E., & Kisamore, A. (2009). Comparing preference assessments: Selection-versus duration-based preference assessment procedures. *Research in Developmental Disabilities, 30*, 1068-1077. doi:10.1016/j.ridd.2009.02.010

Kretlow, A.G., & Bartholomew, C.C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education*, *33*(4), 279–299. doi:10.1177/0888406410371643

Kuhn, M.T. (2017). *A comparison of feedback provided through live video streaming and post video conferencing on the treatment integrity of individuals pursuing a board certification in behavior analysis*. [Doctoral dissertation, University of Texas].

Lanner, T., Nichols, B., Field, S., Hanson, J., & Zane, T. (2010). The clinical utility of two reinforcement preference assessment techniques: A comparison of duration of assessment and identification of functional reinforcers. *The Behavior Analyst Today, 10*(3-4), 456-466.

Lattimore, L.P., Parsons, M.B., & Reid, D.H. (2006). Enhancing job-site training of supported workers with autism: A reemphasis on simulation. *Journal of Applied Behavior Analysis, 39*(1), 91-102. DOI: 10.1901/jaba.2006.154-04

Ledford, J.R., & Gast, D.L. (2018). *Single case research methodology*. Routledge Taylor and Francis Group.

Lerman, D.C., Vorndran, C.M., Addison, L., & Kuhn, S.C. (2004). Preparing teachers in evidence-based practices for young children with autism. *School Psychology Review, 33*(4), 510-526.

Lim, N. & Hu, X. (2020). The effects of video modeling on staff implementation of behavioral procedures in China. *Behavioral Interventions, 35*, 704-716. doi: 10.1002/bin.1735

Lipschultz, J.L., Vladescu, J.C., Reeve, K.F., Reeve, S.A., & Dipsey, C.R. (2015). Using video modeling with voiceover instruction to train staff to conduct stimulus preference assessments. *Journal of Developmental and Physical Disabilities, 27*, 505-532. doi:10.1007/s10882-015-9434-4

Logan, K.R., & Gast, D.L. (2001). Conducting preference assessments and reinforcer testing for individuals with profound multiple disabilities: Issues and procedures. *Exceptionality, 9*(3), 123-134. https://doi.org/10.1207/S15327035EX0903_3

Luck, K.M., Lerman, D.C., Wu, W., Dupuis, D.L., & Hussein, L.A. (2018). A comparison of written, vocal, and video feedback when training teachers. *Journal of Behavioral Education, 27*, 124-144. https://doi.org/10.1007/s10864-017-9279-2

Mangum, A., Fredrick, L., Pabico, R., & Roane, H. (2012). The role of context in the evaluation of reinforcer efficacy: Implications for the preference assessment outcomes. *Research in Autism Spectrum Disorders, 6*, 158-167. doi:10.1016/j.rasd.2011.04.001

Mason, S.A., McGee, G.G., Farmer-Dougan, V., & Risley, T. (1989). A practical strategy for ongoing reinforcer assessment. *Journal of Applied Behavior Analysis, 22*(2), 171-179. doi: 10.1901/jaba.1989.22-171

McAdam, D.B., Klatt, K.P., Koffarnus, M., Dicesare, A., Solberg, K., Welch, C., & Murphy, S. (2005). The effects of establishing operations on preferences for tangible items. *Journal of Applied Behavior Analysis, 38*(1), 107-110. doi: 10.1901/jaba.2005.112-03

McDuff, E., Lanovaz, M.J., Morin, D., Vona, M., Kheloufi, Y., Giannakakos, A.R. (2019). Differential reinforcement of high rates of behavior to increase work productivity in adults with intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 32*(5), 1288-1293. doi: 10.1111/jar.12614

Miljkovic, M., Kaminski, L., & Wishnowski, L. (2015). Evaluation of video modelling and self-instructional manual to teach students to conduct a preference assessment. *Journal on Developmental Disabilities, 21*(1), 3-11.

Nelson, C., Paul, K., & Johnston, S. (2017). Use of a creative dance intervention package to increase social engagement and play complexity of young children with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities, 52*(2), 170-185.

Nergaard, S.K., & Couto, K.C. (2021). Effects of reinforcement and response-cost history on instructional control. *Journal of the Experimental Analysis of Behavior, 115*(3), 679-701. doi:10.1002/jeab.680

Nosek, J.T. (2017). *Using GAINS™ to overcome problems in ABA delivery* [White paper]. Guiding Technologies. https://app.hubspot.com/documents/3472846/view/13794223?accessId=c8e359

Nottingham, C.L., Vladescu, J.C., Giannakakos, A.R., Schnell, L.K., & Lipschultz, J.L. (2018). Reprint of: Using video modeling with voiceover instruction plus feedback to train implementation of stimulus preference assessments. *Learning and Motivation, 62*, 119-129. doi: 10.1016/j.lmot.2017.01.008

Pace, G.M., Ivancic, M.T., Edwards, G.L., Iwata, B.A., & Page, T.J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18*(3), 249-255. doi: 10.1901/jaba.1985.18-249.

Paclawskyj, T.R., & Vollmer, T.M. (1995). Reinforcer assessment for children with developmental disabilities and visual impairment. *Journal of Applied Behavior Analysis, 28*(2), 219-224. doi: 10.1901/jaba.1995.28-219

Parker, R.I., Vannest, K.J., & Davis, J.L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*(4), 303-322. doi: 10.1177/0145445511399147

Pellegrino, D. (2019). *Using a computerized application to teach behavior technicians how to conduct a multiple-stimulus without replacement preference assessment*. [Unpublished doctoral dissertation]. Chicago School of Professional Psychology

Pence, S.T., St. Peter, C.C., & Tetreault, A.S. (2012). Increasing accurate preference assessment implementation through pyramidal training. *Journal of Applied Behavior Analysis, 45*(2), 345-359. doi: 10.1901/jaba.2012.45-345

Phillips, C.L., Iannaccone, J.A., Rooker, G.W., & Hagopian, L.P. (2017). Noncontingent reinforcement for the treatment of severe problem behavior: An analysis of 27 consecutive applications. *Journal of Applied Behavior Analysis, 50*(2), 357-376. doi: 10.1002/jaba.376

Pino, O., & Dazzi, C. (2005). Evaluation of contextual variability in prediction of reinforcer effectiveness. *Education and Training in Developmental Disabilities, 40*(2). 117-130.

Ramon, D., Yu, C.T., Martin, G.L., & Martin, T. (2015). Evaluation of a self-instructional manual to teach multiple-stimulus without replacement preference assessments. *Journal of Behavioral Education, 24*, 289-303. doi: 10.1007/s10864-015-9222-3

Resetar, J.L. & Noell, G.H. (2008). Evaluating preference assessments for use in the general education population. *Journal of Applied Behavior Analysis, 41*(3), 447-451.

Rispoli, M., Neely, L., Lang, R., & Ganz, J. (2011). Training paraprofessionals to implement interventions for people autism spectrum disorders: A systematic review. *Developmental Neurorehabilitation, 14*(6), 378-388. doi: 10.3109/17518423.2011.620577

Roane, H.S., Vollmer, T. R., Ringdahl, J.E., & Marcus, B.A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis, 31*(4), 605-620. doi: 10.1901/jaba.1998.31-605

Rosales, R., Gongola, L., & Homlitas, C. (2015). An evaluation of video modeling with embedded instructions to teach implementation of stimulus preference assessments. *Journal of Applied Behavior Analysis, 48*(1), 209-214. doi: 10.1002/jaba.174

Roscoe, E.M. & Fisher, W.W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis, 41*(2), 249-254. doi: 10.1901/jaba.2008.41-249

Rossett, A. & Schafer, L. (2006). *Job aids and performance support: Moving from knowledge in the classroom to knowledge everywhere* (2nd ed.). Pfeiffer.

Shapiro, M., Kazemi, E., Pogosjana, M., Rios, D., & Mendoza, M. (2016). Preference assessment training via self-instruction: A replication and extension. *Journal of Applied Behavior Analysis, 49*(4), 794-808. doi: 10.1002/jaba.339

Steinbrenner, J. R., Hume, K., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yücesoy-Özkan, S., & Savage, M. N. (2020). *Evidence-*

based practices for children, youth, and young adults with autism. The University of North Carolina at Chapel Hill, Frank Porter Graham Child Development Institute, National Clearinghouse on Autism Evidence and Practice Review Team.

Sump, L.A., Richman, D.M., Schaefer, A.M., Grubb, L.M., & Brewer, A.T. (2018). Telehealth and in-person training outcomes for novice discrete trial training therapists. *Journal of Applied Behavior Analysis, 51*(3), 466-481. doi: 10.1002/jaba.461

Tereshko, L., Leaf, J.B., Weiss, M.J., Rich, A., & Pistorino, M. (2021). A systematic literature review of antecedent and reinforcement-based behavioral feeding interventions without the implementation of escape extinction. *Behavioral Interventions, 36*(2), 496-513.

Verschuur, R., Didden, R., Van der Meer, L., Achmadi, D., Kagohara, D., Green, V.A., Lang, R., & Lancioni, G.E. (2011). Investigating the validity of a structured interview protocol for assessing the preferences of children with autism spectrum disorders. *Developmental Neurorehabilitation, 14*(6), 366-371. doi: 10.3109/17518423.2011.606509

Virues-Ortega, J., Pritchard, K., Grant, R.L., North, S., Hurtado-Parrado, C., Lee, M.S.H., Temple, B., Julio, F., & Yu, C.T. (2014). Clinical decision making and preference assessment for individuals with intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities, 119*(2), 151-170. doi: 10.1352/1944-7558-119.2.151

Vollmer, R.T., & Iwata, B.A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis, 24*(2), 279-291. doi: 10.1901/jaba.1991.24-279

Wallace, M.D., Iwata, B.A., Hanley, G.P., Thompson, R.H., & Roscoe, E.M. (2013). Noncontingent reinforcement: A further examination of schedule effects during treatment. *Journal of Applied Behavior Analysis, 45*(4), 709-719. doi: 10.1901/jaba.2012.45-709.

Weeden, M., & Poling, A. (2011). Identifying reinforcers in skill acquisition studies involving participants with autism: Procedures reported from 2005 to 2009. *Research in Autism Spectrum Disorders, 5*, 388-391. doi:10.1016/j.rasd.2010.05.004

Weldy, C.R., Rapp, J.T., & Capocasa, K. (2014). Training staff to implement brief stimulus preference assessments. *Journal of Applied Behavior Analysis, 47*(1), 214-218. doi: 10.1002/jaba.98

Weston, R., Davis, T.N., Radhakrishnan, S., O'Guinn, N., & Rivera, G. (2020). Comparing performance feedback and video self-monitoring within a BST package to train pre-service behavior analysts to conduct preference assessments. *Journal of Behavioral Education, 29*, 571-583. https://doi.org/10.1007/s10864-019-09333-w

Wilkinson, L. (2017). *A best practice guide to assessment and intervention for autism spectrum disorder in schools* (2nd ed.). Jessica Kingsley Publishers.

Wishnowski, L.A., Yu, C.T., Pear, J., Chand, C., & Saltel, L. (2017). Effects of computer-aided instruction on the implementation of the MSWO stimulus preference assessment. *Behavioral Interventions, 33*, 56-68. doi: 10.1002/bin.1508

Windsor, J., Piche, L.M., & Locke, P.A. (1994). Preference testing: A comparison of two presentation methods. *Research in Developmental Disabilities, 15*(6), 439-455.

Zhou, L., Iwata, B.A., Goff, G.A., & Shore, B.A. (2001). Longitudinal analysis of leisure-item preferences. *Journal of Applied Behavior Analysis, 34*(2), 179-184. doi: 10.1901/jaba.2001.34-179

Appendix A

Participant Information

Name:______________________________________ Date of Birth:_________________________

Biological Sex (Circle One): Male Female Prefer Not to Say

Race/Ethnic Category (Circle One): American Indian or Alaska Native Asian

Black or African American Native Hawaiian or Pacific Islander White/Non-Hispanic or Latino

White, Hispanic or Latino Prefer Not to Say

Highest Level of Education Completed (Circle One): High School Associates Degree

Bachelor's Degree Graduate Degree

How long have you worked in the field of Special Education (Circle One):

Less than a year 1-3 years 3-5 years 5+ years

Have you ever had formal training on preference assessment procedures? Yes No Unsure

If yes, about how much? Less than one hour Half day Full day More than one day

Have you ever conducted a multiple stimulus without replacement preference assessment with a child with autism before this study? Yes No Unsure

Appendix B

Scripts Used by Simulated Learner

Script 1

Trial 1 – Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 2 – Select 1 item and then immediately select another item. If the items are represented, select only 1 item.

Trial 3 – Select 1 item approximately 10 seconds after items are presented.

Trial 4 – Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 5 — Select 1 item approximately 1 second after it is presented, interact with that item until it is removed from your hands.

Trial 6 - Select 2 items simultaneously. If this is blocked/items are immediately removed, select only one item.

Trial 7 - Select one item immediately after it is presented, begin playing with item and after approximately 5 seconds reach for a second item. If access to a second item is blocked continue to play with initial item. If it is not blocked, play with both items until they are removed.

Trial 8 - Select 1 item and then immediately select another item. If the items are represented, select only 1 item.

Trial 9 - Select 2 items simultaneously. If this is blocked/items are immediately removed, select only 1 item.

Trial 10 – Do NOT select the last item. If item is represented, do NOT select it.

Trial 11 - Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 12 - Select 1 item immediately after it is presented, begin playing with item and after approximately 5 seconds reach for a second item. If access to a second item is blocked continue to play with initial item. If it is not blocked, play with both items until they are removed.

Trial 13 - Select 2 items simultaneously. If this is blocked/items are immediately removed, again select 2 items simultaneously.

Trial 14 - Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 15- Select last item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Appendix B (Cont.)

Script 2

Trial 1 - Select 1 item approximately 1 second after it is presented, interact with that item until it is removed from your hands.

Trial 2 - Select 2 items simultaneously. If this is blocked/items are immediately removed, select only one item.

Trial 3 - Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 4 - Select 1 item and then immediately select another item. If the items are represented, select only 1 item.

Trial 5 – Do NOT select last item. If item is represented, select the item.

Trial 6 - Select 1 item and then immediately select another item. If the items are represented, select only 1 item.

Trial 7 – Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 8 - Select 1 item approximately 10 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 9 - Select 2 items simultaneously. If this is blocked/items are immediately removed, select only one item.

Trial 10 - Select 1 item approximately 10 seconds after items are presented.

Trial 11 - Select 1 item approximately 10 seconds after items are presented.

Trial 12 - Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 13 - Select 1 item immediately after it is presented, begin playing with item and after approximately 5 seconds reach for a second item. If access to a second item is blocked continue to play with initial item. If it is not blocked, play with both items until they are removed.

Trial 14- Do NOT select an item. If items are represented, select one item.

Trial 15 - Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Appendix B (Cont.)

Script 3

Trial 1 - Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 2 - Select 2 items simultaneously. If this is blocked/items are immediately removed, select only one item.

Trial 3 - Select 1 item approximately 10 seconds after items are presented.

Trial 4 - Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 5 - Select 1 item approximately 10 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 6 - Select 1 item immediately after it is presented, begin playing with item and after approximately 5 seconds reach for a second item. If access to a second item is blocked continue to play with initial item. If it is not blocked, play with both items until they are removed.

Trial 7 – Do NOT select an item. If items are represented, select one item.

Trial 8 - Select an item that is not in the stimulus array (i.e., a pen, paper, etc.). If items are re-presented, select 1 appropriate item.

Trial 9 - Select 1 item immediately after it is presented, interact with that item until it is removed from your hands.

Trial 10 - Select 1 item approximately 10 seconds after it is presented.

Trial 11 - Select 1 item and then immediately select another item. If the items are represented, select only 1 item.

Trial 12- Select 1 item approximately 5 seconds after it is presented, interact with that item until it is removed from your hands.

Trial 13 - Select one item immediately after it is presented, begin playing with item and after approximately 5 seconds reach for a second item. If access to a second item is blocked continue to play with initial item. If it is not blocked, play with both items until they are removed.

Trial 14 - Select 1 item immediately after it is presented, interact with that item until it is removed.

Trial 15- Do NOT select last item. If item is represented, do NOT select it.

Appendix C

Child Participant Information

Name:________________________________ **Date of Birth:**________________________

Biological Sex (Circle One): Male Female Prefer Not to Say

Race/Ethnic Category (Circle One): American Indian or Alaska Native Asian

Black or African American Native Hawaiian or Pacific Islander White/Non-Hispanic or Latino

White, Hispanic or Latino Prefer Not to Say

Appendix D

IRB Approval Letter

UNIVERSITY OF ARKANSAS

To:	Brenna R Griffen
From:	Douglas J Adams, Chair IRB Expedited Review
Date:	02/03/2023
Action:	**Expedited Approval**
Action Date:	02/03/2023
Protocol #:	2208415588
Study Title:	Evaluating a Technological Tool for Conducting Preference Assessments
Expiration Date:	01/12/2024
Last Approval Date:	

The above-referenced protocol has been approved following expedited review by the IRB Committee that oversees research with human subjects.

If the research involves collaboration with another institution then the research cannot commence until the Committee receives written notification of approval from the collaborating institution's IRB.

It is the Principal Investigator's responsibility to obtain review and continued approval before the expiration date.

Protocols are approved for a maximum period of one year. You may not continue any research activity beyond the expiration date without Committee approval. Please submit continuation requests early enough to allow sufficient time for review. Failure to receive approval for continuation before the expiration date will result in the automatic suspension of the approval of this protocol. Information collected following suspension is unapproved research and cannot be reported or published as research data. If you do not wish continued approval, please notify the Committee of the study closure.

Adverse Events: Any serious or unexpected adverse event must be reported to the IRB Committee within 48 hours. All other adverse events should be reported within 10 working days.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, study personnel, or number of participants, please submit an amendment to the IRB. All changes must be approved by the IRB Committee before they can be initiated.

You must maintain a research file for at least 3 years after completion of the study. This file should include all correspondence with the IRB Committee, original signed consent forms, and study data.

cc: Elizabeth R Lorah, Investigator

Appendix E

Adult Participant Consent Form

Peabody 215; Fayetteville, Arkansas 72701; (479) 575-4209; (479) 575-6676 (Fax)
Department of Curriculum & Instruction, College of Education & Health Professions

PARTICIPANT CONSENT FORM

Title: Evaluating a Technological Tool for Conducting Preference Assessments

Principal Investigator: Brenna Griffen, Graduate Assistant and Doctoral Student, Curriculum & Instruction, University of Arkansas, 501-388-3265

Faculty Supervisor: Dr. Elizabeth R. Lorah, Associate Professor, Curriculum and Instruction, University of Arkansas, 479-575-5498

INTRODUCTION

This study will evaluate the use of an application to conduct preference assessments with young children. You have been selected to participate in this study because you currently work or in the future will work with children who require preferred items to be identified and used as a part of their instructional program. Therefore, it is appropriate to conduct preference assessments with them to identify these preferred items.

During this study, you will use both traditional means (i.e., paper and pencil directions and data sheet) and GAINS technology (an application on a tablet) to conduct and take data on a preference assessment for a child. The GAINS application will provide you with auditory and visual prompts that provide the steps to conducting the preference assessment, while also collecting data on child selections of preferred items. For example, the application may say to you (via Bluetooth headphone) “Place 5 items on the table in front of the child in a straight line. Say ‘pick one.’” It may show a visual of the order to place the items in, and then you may click on which item the child selects.

Fidelity data will be collected on your implementation of the preference assessment protocol using both methods. Data will also be collected on the duration of time it takes to conduct and

Appendix E (Cont.)

score both methods. Additionally, researchers will request basic demographic data on yourself (e.g., name, age, education level, etc.). Videos of each session will be recorded on the GAINS® application using tablet-based devices and will be used to verify the data collected. These videos will be uploaded to a secure server and can only be viewed through the GAINS reporting website, which is only accessible to approved users with verified passwords. All data will be kept confidential to the extent allowed by law and University of Arkansas policy.

This study is predicted to last one month and may require approximately 2 hour of assessment for 1-2 days.

This study will help to inform the use of GAINS technology for students with disabilities. There are no risks for participation in this study.

If I have any questions about the research, I may contact Brenna Griffen at (501)388-3265 or brgriffe@uark.edu. If I have any questions or concerns about my rights as a research participant, I may contact Ro Windwalker, the University of Arkansas's IRB Compliance Coordinator at (479) 575-2208 or irb@uark.edu.

I understand that participation in this research project is voluntary, and that refusal to participate in this research will involve no loss or penalty to benefits to which I am entitled. I also understand that I may withdraw my consent at any time and discontinue my involvement in the research study, without loss or penalty to benefits to which I am entitled.

<u>CONSENT</u>

I hereby consent to participate in this study.

______________________________ _________________________

Participant Name (please print) Date

____

Participant Signature

______________________________ _________________________

Principal Investigator's Name Date

____

Principal Investigator's Signature

Appendix F

Child Participant Consent Form

Peabody 215 ◆ Fayetteville, Arkansas 72701 ◆ (479) 575-4209 ◆ (479) 575-6676 (Fax)
Department of Curriculum & Instruction, College of Education & Health Professions

PARTICIPANT CONSENT FORM

Title: Evaluating a Technological Tool for Conducting Preference Assessments

Principal Investigator: Brenna Griffen, Graduate Assistant and Doctoral Student, Curriculum & Instruction, University of Arkansas, 501-388-3265

Faculty Supervisor: Dr. Elizabeth R. Lorah, Associate Professor, Curriculum and Instruction, University of Arkansas, 479-575-5498

INTRODUCTION

This study will evaluate the use of an application to conduct preference assessments with young children. You have been selected to participate in this study because your child's preferred items are typically included in their educational program. Therefore, it is appropriate to conduct preference assessments with them to identify these preferred items.

During this study, adult undergraduate students will use both traditional means (i.e., paper and pencil directions and data sheet) and GAINS technology (an application on a tablet) to conduct and take data on a preference assessment for your child. The GAINS application will provide the adult with auditory and visual prompts that provide the steps to conducting the preference assessment, while also collecting data on your child's selections of preferred items.

Fidelity data will be collected on the adults' implementation of the preference assessment protocol using both methods. Data will also be collected on the duration of time it takes to conduct and score both methods. Additionally, researchers will request basic demographic data on your child (e.g., name, age, education level, etc.). Videos of each session will be recorded on the GAINS® application using tablet-based devices and will be used to verify the data collected. These videos will be uploaded to a secure server and can only be viewed through the GAINS reporting website, which is only accessible to approved users with verified passwords. All data

Appendix F (Cont.)

will be kept confidential to the extent allowed by law and University of Arkansas policy. Following the completion of the study, the primary investigator will ask your child a few questions on if they enjoyed the preference assessments. This will take no more than ten minutes.

This study is predicted to last no more than six weeks and may require no more than four hours of assessment per day for 3-4 days a week.

This study will help to inform the use of GAINS technology for students with disabilities. There are no risks for participation in this study.

If I have any questions about the research, I may contact Brenna Griffen at (501)388-3265 or brgriffe@uark.edu. If I have any questions or concerns about my child's rights as a research participant, I may contact Ro Windwalker, the University of Arkansas's IRB Compliance Coordinator at (479) 575-2208 or irb@uark.edu.

I understand that my child's participation in this research project is voluntary, and that refusal to participate or allow my child to participate in this research will involve no loss or penalty to benefits to which my child is entitled. I also understand that I may withdraw my consent for my child to participate at any time and discontinue my child's involvement in the research study, without loss or penalty to benefits to which my child is entitled.

CONSENT

I hereby consent for my child ________________________________ to participate in this study.

______________________________ ______________________________

Participant Name (please print) Date

____

Parent/ Guardian Name (please print)

____

Parent/ Guardian Signature

______________________________ ______________________________

Principal Investigator's Name Date

____

Principal Investigator's Signature

Appendix G

Pen and Paper Written Instructions

Multiple Stimulus Without Replacement (MSWO; adapted from DeLeon & Iwata, 1996)

Procedure

Prior to each session, participants were instructed or prompted to sit in one of the chairs; the experimenter sat in the other. Five items per participant were chosen for presentation during each assessment. A selection response was recorded when the participant made physical contact with one of the presented items. The experimenter also served as an observer and recorded selections on data sheets that were customized for the MSWO procedure. The primary dependent variable consisted of a percentage score indicating the number of times an item was selected over the number of trials during which the item was presented. This score was then used to rank the items from most-to-least preferred.

Prior to the beginning of the first session, participants were given 30-s access to each of the leisure items. Subsequently, participants were exposed to one or two assessment sessions per day. Each session began with all items sequenced randomly in a straight line on the table about 5 inches apart. While a participant was seated at the table approximately ½ -to- 1 ft. from the stimulus array, the experimenter instructed the participant to "pick one." The participant had 10 s to select an item. If the participant selected an item, the participant received 30-s access to the item. After the participant received access to the item, the item was removed from the immediate area. Prior to the next trial, the sequencing of the remaining items was rotated by taking the item at the left end of the line and moving it to the right end, then shifting the other items so that they were again equally spaced on the table. The second trial then followed immediately. This procedure continued until all items were selected or until a participant made no selection within 10 s from the beginning of a trial. In the latter case, the session ended (all remaining stimuli were removed), remaining stimuli were recorded as "not selected," and a new session was initiated. If the participant made contact with more than one item, the participant was allowed access to the first item contacted, and that item was recorded as the selection. If two or more items were simultaneously selected, participants were not allowed access to the items and the trial was reinitiated (i.e., the stimuli were rotated and they were again asked to "pick one"). If the participant simultaneously selected two items a second time, the session ended. If the participant grabbed an item that was not in the array, the grabbed item was removed, and the therapist continued with the current trial.

Appendix H

MSWO Data Sheet

Learner Name: ______________________

Therapist Initials: ______

Stimuli

1.	3.	5.
2.	4.	

Trial 1 ______ **Initial Order:** 1-2-3-4-5		**Date:**	**Comments**
Trial		**Item Selected**	
1	1 2 3 4 5		
2	X X X X		
3	X X X		
4	X X		
5	X		

Trial 2 _______ **Initial Order:** 2-3-4-5-1		**Date:**	**Comments**
Trial	**Circle position**	**Item Selected**	
1	2 3 4 5 1		
2	X X X X		
3	X X X		
4	X X		
5	X		

Trial 3 ______ **Initial Order:** 3-4-5-6-7-1-2		**Date:** **DC:** **Prim /** **Reli**	**Comments**
Trial	**Circle position**	**Item Selected**	
1	3 4 5 1 2		
2	X X X X		
3	X X X		
4	X X		
5	X		

Appendix H (Cont.)

Scoring MSWO:

Calculate the percentage of selection for an item by taking the number of times the item was selected divided by the number of times the item was presented and multiplying by 100.

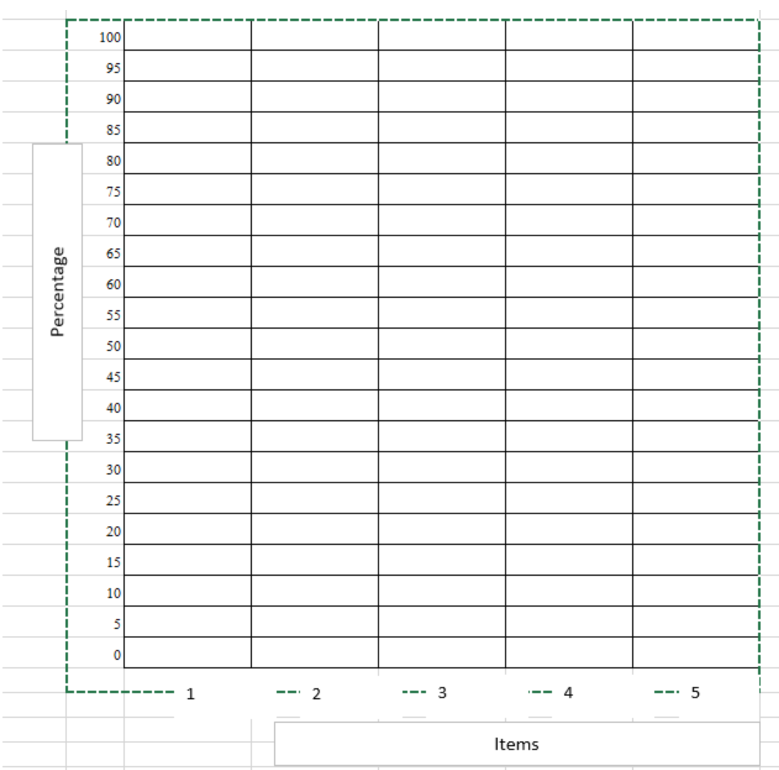
(# Times selected / # times presented) X 100

For example, if skittle was selected on the first trial for two sessions (presented 2 times; selected 2 times), the second trial for one session (presented 2 times; selected 1 time), the third trial for one session (presented 3 times; selected 1 time), and the fifth trial for one session (presented 5 times; selected 1 time), the percentage selected would be 41.67%.

(5 / 12) X 100 = 41.67 %

Below calculate the percentage each item was selected.

Stimulus	% Selected
1.	
2.	
3.	
4.	
5.	



Appendix I

MSWO Fidelity Checklist

Implementor: ___________________________ Date and Session: ____________________

Data Collector: ____________________________ Is this IOA : YES NO

1. Data on Paper or App	Items and info entered correctly.	+ - N/A	N/A	N/A
		Trial 1	Trial 2	Trial 3
2. Item Arrangement (Trial 1)	Items placed in front of learner within arm's reach.	+ - N/A	+ - N/A	+ - N/A
	Spaced evenly (5 inches apart) in a straight line.	+ - N/A	+ - N/A	+ - N/A
3. Say "pick one"	Correct instruction given.	+ - N/A	+ - N/A	+ - N/A
4. Item Selection	Learner allowed up to 10s for selection.	+ - N/A	+ - N/A	+ - N/A
	Blocked if more than one selection.	+ - N/A	+ - N/A	+ - N/A
	Data collected.	+ - N/A	+ - N/A	+ - N/A
5. Access to Item	Learner allowed access for 30s or until consumed.	+ - N/A	+ - N/A	+ - N/A
	Correct instruction for item removal and prompted as needed.	+ - N/A	+ - N/A	+ - N/A
	Item placed out of sight.	+ - N/A	+ - N/A	+ - N/A
6. Rotate Items	Items rotated in the correct order.	+ - N/A	+ - N/A	+ - N/A
7. Item Arrangement (Trial 2)	Items placed in front of learner within arm's reach.	+ - N/A	+ - N/A	+ - N/A
	Spaced evenly (5 inches apart) in a straight line.	+ - N/A	+ - N/A	+ - N/A
8. Say "pick one"	Correct instruction given.	+ - N/A	+ - N/A	+ - N/A
9. Item Selection	Learner allowed up to 10s for selection.	+ - N/A	+ - N/A	+ - N/A
	Blocked if more than one selection.	+ - N/A	+ - N/A	+ - N/A
	Data collected.	+ - N/A	+ - N/A	+ - N/A

Appendix I (Cont.)

10. Access to Item	Learner allowed access for 30s or until consumed.	+	-	N/A	+	-	N/A	+	-	N/A
	Correct instruction for item removal and prompted as needed.	+	-	N/A	+	-	N/A	+	-	N/A
	Item placed out of sight.	+	-	N/A	+	-	N/A	+	-	N/A
11. Rotate Items	Items rotated in correct order.	+	-	N/A	+	-	N/A	+	-	N/A
12. Item Arrangement (Trial 3)	Items placed in front of learner within arm's reach.	+	-	N/A	+	-	N/A	+	-	N/A
	Spaced evenly (5 inches apart) in a straight line.	+	-	N/A	+	-	N/A	+	-	N/A
13. Say "pick one"	Correct instruction given.	+	-	N/A	+	-	N/A	+	-	N/A
14. Item Selection	Learner allowed up to 10s for selection.	+	-	N/A	+	-	N/A	+	-	N/A
	Blocked if more than one selection.	+	-	N/A	+	-	N/A	+	-	N/A
	Data collected.	+	-	N/A	+	-	N/A	+	-	N/A
15. Access to Item	Learner allowed access for 30s or until consumed.	+	-	N/A	+	-	N/A	+	-	N/A
	Correct instruction for item removal and prompted as needed.	+	-	N/A	+	-	N/A	+	-	N/A
	Item placed out of sight.	+	-	N/A	+	-	N/A	+	-	N/A
16. Rotate Items	Items rotated in correct order.	+	-	N/A	+	-	N/A	+	-	N/A
17. Item Arrangement (Trial 4)	Items placed in front of learner within arm's reach.	+	-	N/A	+	-	N/A	+	-	N/A
	Spaced evenly (5 inches apart) in a straight line.	+	-	N/A	+	-	N/A	+	-	N/A
18. Say "pick one"	Correct instruction given.	+	-	N/A	+	-	N/A	+	-	N/A
19. Item Selection	Learner allowed up to 10s for selection.	+	-	N/A	+	-	N/A	+	-	N/A
	Blocked if more than one selection.	+	-	N/A	+	-	N/A	+	-	N/A
	Data collected.	+	-	N/A	+	-	N/A	+	-	N/A

Appendix I (Cont.)

20. Access to Item	Learner allowed access for 30s or until consumed.	+	-	N/A	+	-	N/A	+	-	N/A
	Correct instruction for item removal and prompted as needed.	+	-	N/A	+	-	N/A	+	-	N/A
	Item placed out of sight.	+	-	N/A	+	-	N/A	+	-	N/A
21. Access to Final Item	Correct instruction given.	+	-	N/A	+	-	N/A	+	-	N/A
	Learner given up to 10s to begin interaction.	+	-	N/A	+	-	N/A	+	-	N/A
	Learner allowed access for 30s or until consumed.	+	-	N/A	+	-	N/A	+	-	N/A
	Correct instruction for item removal and prompted as needed.	+	-	N/A	+	-	N/A	+	-	N/A
	Data collected.	+	-	N/A	+	-	N/A	+	-	N/A
22. Calculate and Rank	Correct percentages.	N/A			N/A			+	-	N/A
	Correct ranking.	N/A			N/A			+	-	N/A
	Total per trial:	/45			/44			/48		

Total Percentage for Session: __________/137 X 100 =_______________

Total Duration: ____

Appendix J

GAINS Usability Survey

Participant (circle): 1 2 3 4 **Date: _____________**

We are interested in your thoughts about GAINS. As a professional who has used GAINS with a consumer, we value your input in helping us improve GAINS.

Please place an x in the cell for your response.

First, we have some questions about your experience with GAINS.	**Question Responses**				
	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
The use of GAINS can enhance my job performance in helping consumers learn new skills.					
The use of GAINS can make me more effective in helping consumers learn new skills.					
The use of GAINS in helping consumers learn new skills enhances my productivity.					
Generally, I consider GAINS can be useful to me in helping consumers learn new skills.					
Learning to use GAINS would be easy for me.					
I find it easy to interact with GAINS.					
Interaction with GAINS is clear and easy to understand for me.					
Generally, I consider GAINS easy to use.					
If I have a choice, I intend to use GAINS in helping consumers learn new skills.					
I predict I will use GAINS in helping consumers learn new skills.					
If I have a choice, I plan to use GAINS in helping consumers learn new skills.					
I am able to integrate GAINS in helping consumers learn new skills.					
I can use GAINS even if there is no one to help me.					

Appendix J (Cont.)

Now, we have questions about some features of GAINS.	**Question Responses**				
	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
It is easy to login into GAINS.					
It is easy to start instruction with the consumer.					
I can hear the audio assistance provided.					
It is easy to know what to do next.					
Audio assistance is easy to follow.					
Audio assistance is useful.					
The display is easy to read.					
Audio is enough. It is not necessary to read the display.					
It is easier than pen and paper to input data.					
Data recording is more accurate than pen and paper.					
It is useful to be provided error correction for a step.					
It is easy to tell what the error correction is for a step.					
It is useful that GAINS tracks preference of each item.					
GAINS makes it easy to track preference of each item.					
Choosing Guidance Type (detailed, brief, etc.) is easy.					
Choosing Guidance Type (detailed, brief, etc.) is useful.					

Now, we have questions about GAINS and the Preference Assessment	**Question Responses**				
	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
GAINS improves Preference Assessment administration.					
It is important for consumers to identify preferences.					
Using Preference Assessment to identify preferences is useful.					
Learning how to conduct Preference Assessments is useful.					

Thank You!

Appendix K

Social Validity Questionnaire for Children

1. Did you like having choices of things to (play with) or (to eat)? YES I DON'T CARE NO

2. Did you like it when the teacher used the paper? YES I DON'T CARE NO

3. Did you like it when the teacher used the tablet? YES I DON'T CARE NO

4. Which one did you like more the paper or the tablet or it doesn't matter? Paper Tablet I DON'T CARE

5. Did you have fun here today? YES I DON'T CARE NO

Appendix L

Procedural Fidelity Questionnaire

1. Did the primary investigator provide instructions before beginning the MSWO?
   YES NO N/A

2. Did the primary investigator provide commentary during the MSWO?
   YES NO N/A

3. Did the primary investigator provide feedback after the MSWO?
   YES NO N/A

4. Did the primary investigator help interpret the results of the MSWO (i.e., help calculate percentages of choice, assist in developing preference hierarchy, etc.)?
   YES NO N/A

5. Did the adult simulated learner follow the steps of the chosen script accurately and completely?
   YES NO N/A

Table 1

Adult Participant Demographic Information

Name	Age (in years)	Biological Sex	Race/ Ethnicity Category	Years Working in Field	Previous Training with SPA	Previous experience conducting SPA	Year
Jill	21	Female	White/ Non-Hispanic	Less than one	No	No	Junior
Molly	20	Female	White/ Non-Hispanic	None	No	No	Sophomore
Yazmina	20	Female	White, Hispanic or Latino	None	No	No	Junior
Shawna	22	Female	White/ Non-Hispanic	None	No	No	First Year Graduate Student
Adriana	20	Female	Hispanic	None	No	No	Junior

Table 2

Social Validity- Acceptability

Question	Mean and Range
Learning to use GAINS would be easy for me.	5
I find it easy to interact with GAINS.	5
Interaction with GAINS is clear and easy to understand for me.	4.8 (range 4-5)
Generally, I consider GAINS easy to use.	4.8 (range 4-5)
If I have a choice, I intend to use GAINS in helping consumers learn new skills.	4.6 (range 4-5)
I predict I will use GAINS in helping consumers learn new skills.	4.2 (range 3-5)
If I have a choice, I plan to use GAINS in helping consumers learn new skills.	4.4 (range 4-5)
I am able to integrate GAINS in helping consumers learn new skills.	4.6 (range 4-5)
I can use GAINS even if there is no one to help me.	5
It is easy to login into GAINS.	4.6 (range 3-5)
It is easy to start instruction with the consumer.	5
I can hear the audio assistance provided.	4.8 (range 4-5)
It is easy to know what to do next.	5
Audio assistance is easy to follow.	4.2 (range 3-5)
Audio assistance is useful.	3.8 (range 3-5)
The display is easy to read.	5
Audio is enough. It is not necessary to read the display.	2.6 (range 1-5)
It is easier than pen and paper to input data.	4.4 (range 2-5)
Data recording is more accurate than pen and paper.	4.6 (range 3-5)
It is useful to be provided error correction for a step.	4.4 (range 4-5)
It is easy to tell what the error correction is for a step.	4.4 (range 3-5)
It is useful that GAINS tracks preference of each item.	5
Choosing Guidance Type (detailed, brief, etc.) is easy.	4.4 (range 3-5)
Choosing Guidance Type (detailed, brief, etc.) is useful.	4.4 (range 3-5)

Table 3

Social Validity- Social Significance

Question	Mean and Range
The use of GAINS can enhance my job performance in helping consumers learn new skills.	5
The use of GAINS can make me more effective in helping consumers learn new skills.	4.6 (range 4-5)
The use of GAINS in helping consumers learn new skills enhances my productivity.	4.6 (range 4-5)
Generally, I consider GAINS can be useful to me in helping consumers learn new skills.	4.6 (range 4-5)
GAINS makes it easy to track preference of each item.	4.8 (range 4-5)
GAINS improves Preference Assessment administration.	5
It is important for consumers to identify preferences.	5
Using Preference Assessment to identify preferences is useful.	5
Learning how to conduct Preference Assessments is useful.	5

Table 4

Social Validity Responses for Child Participants

Question	Percentage Responding "Yes" or Green Smiley Face	Percentage Responding "I Don't Care" or Yellow Face	Percentage Responding "No" or Red Frowning Face
Did you like having choices of things to (play with) or (to eat)?	100%	0%	0%
Did you like it when the teacher used the paper?	75%	0%	25%
Did you like it when the teacher used the tablet?	75%	25%	0%
Did you have fun today?	100%	0%	0%

Question	Percentage Responding "Tablet"	Percentage Responding "It Doesn't Matter"	Percentage Responding "Paper"
Which one did you like more the paper or the tablet or it doesn't matter?	75%	25%	0%

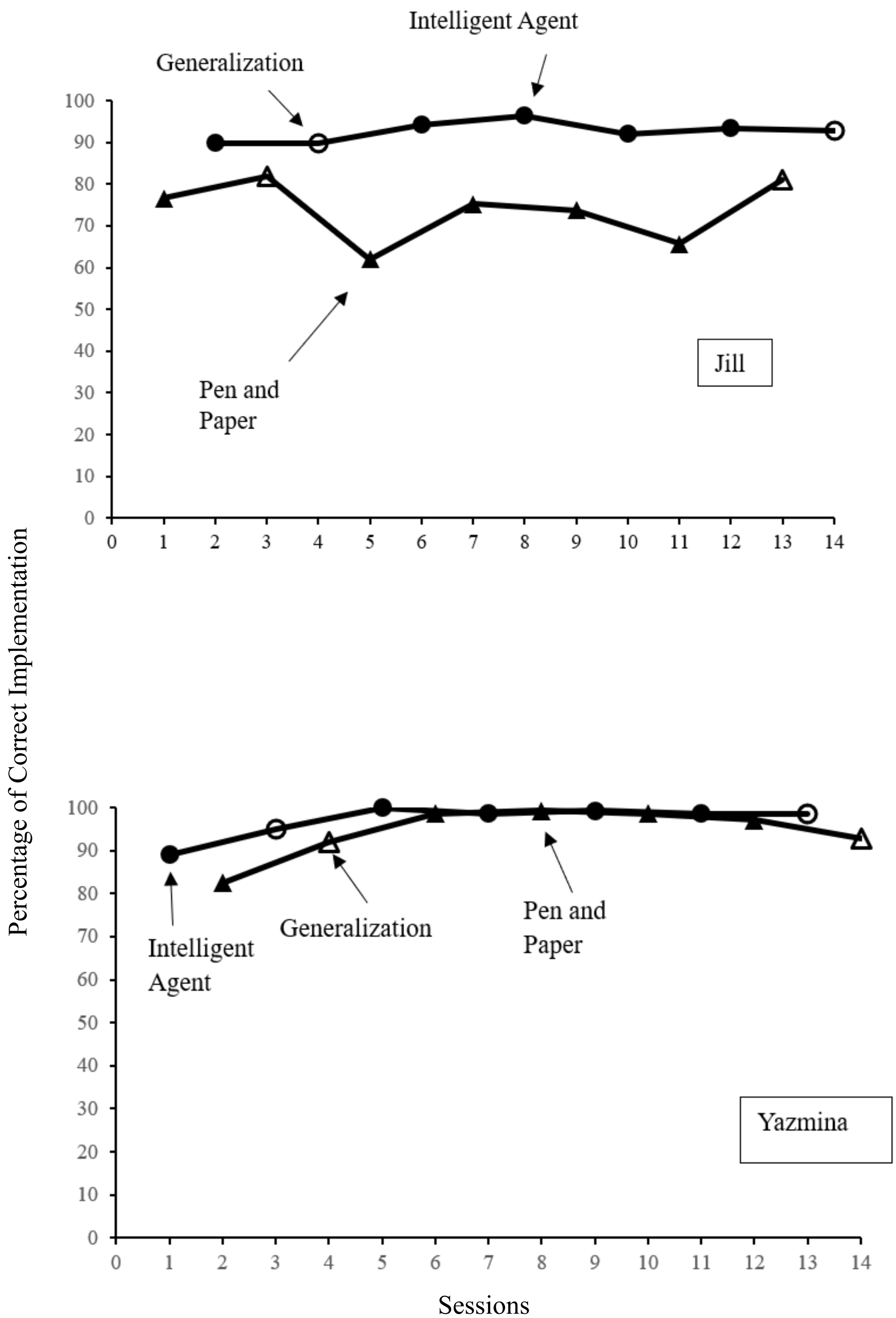


Figure 1. Percentage of Correct Implementation for Jill and Yazmina. This figure depicts the percentage of steps in the MSWO correctly completed by Jill and Yazmina.

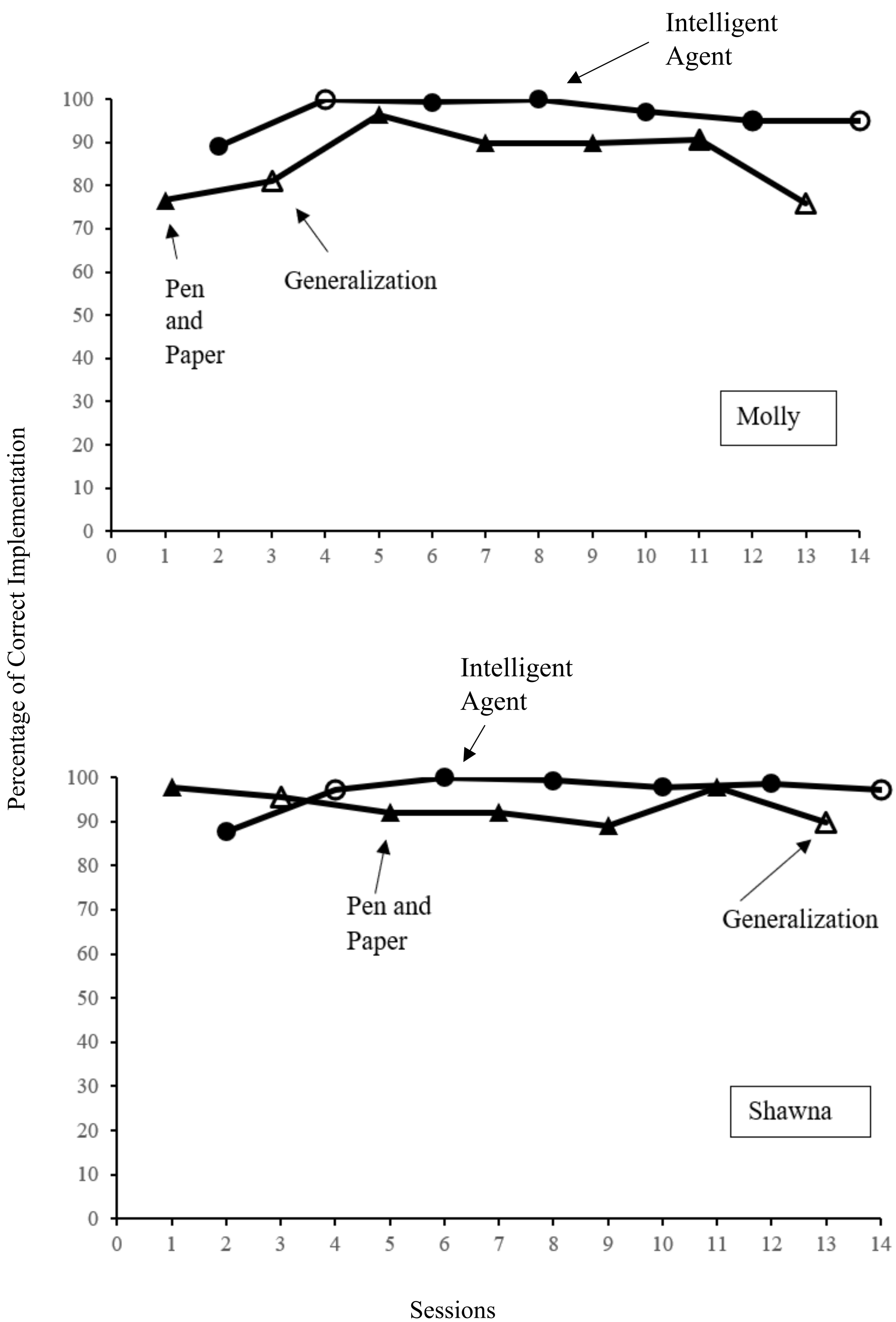


Figure 2. Percentage of Correct Implementation for Molly and Shawna. This figure depicts the percentage of steps in the MSWO correctly completed by Molly and Shawna.

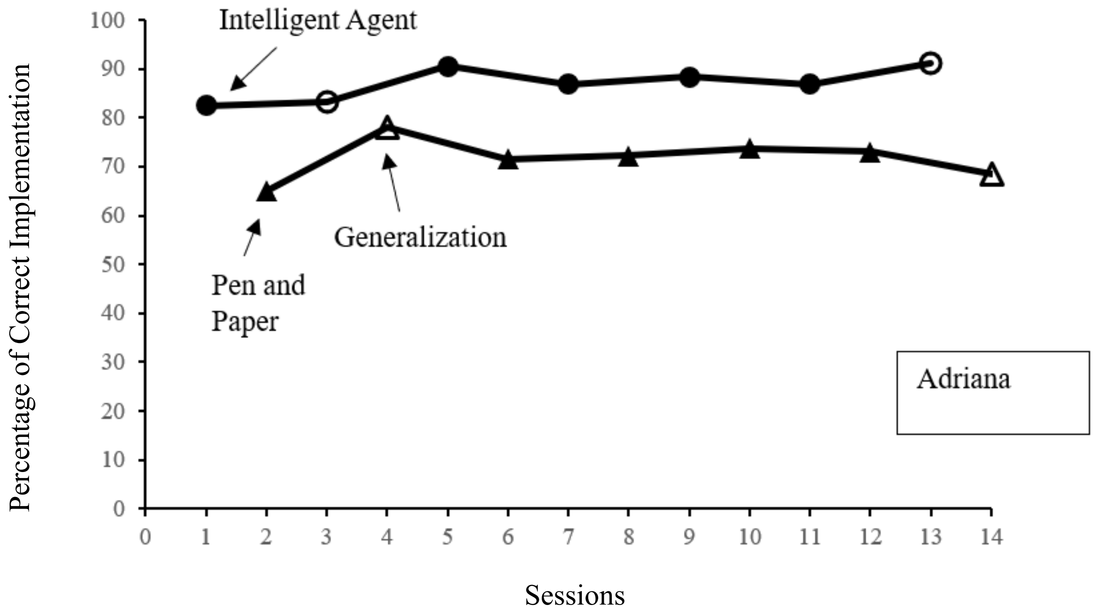


Figure 3. Percentage of Correct Implementation for Adriana. This figure depicts the percentage of steps in the MSWO correctly completed by Adriana.

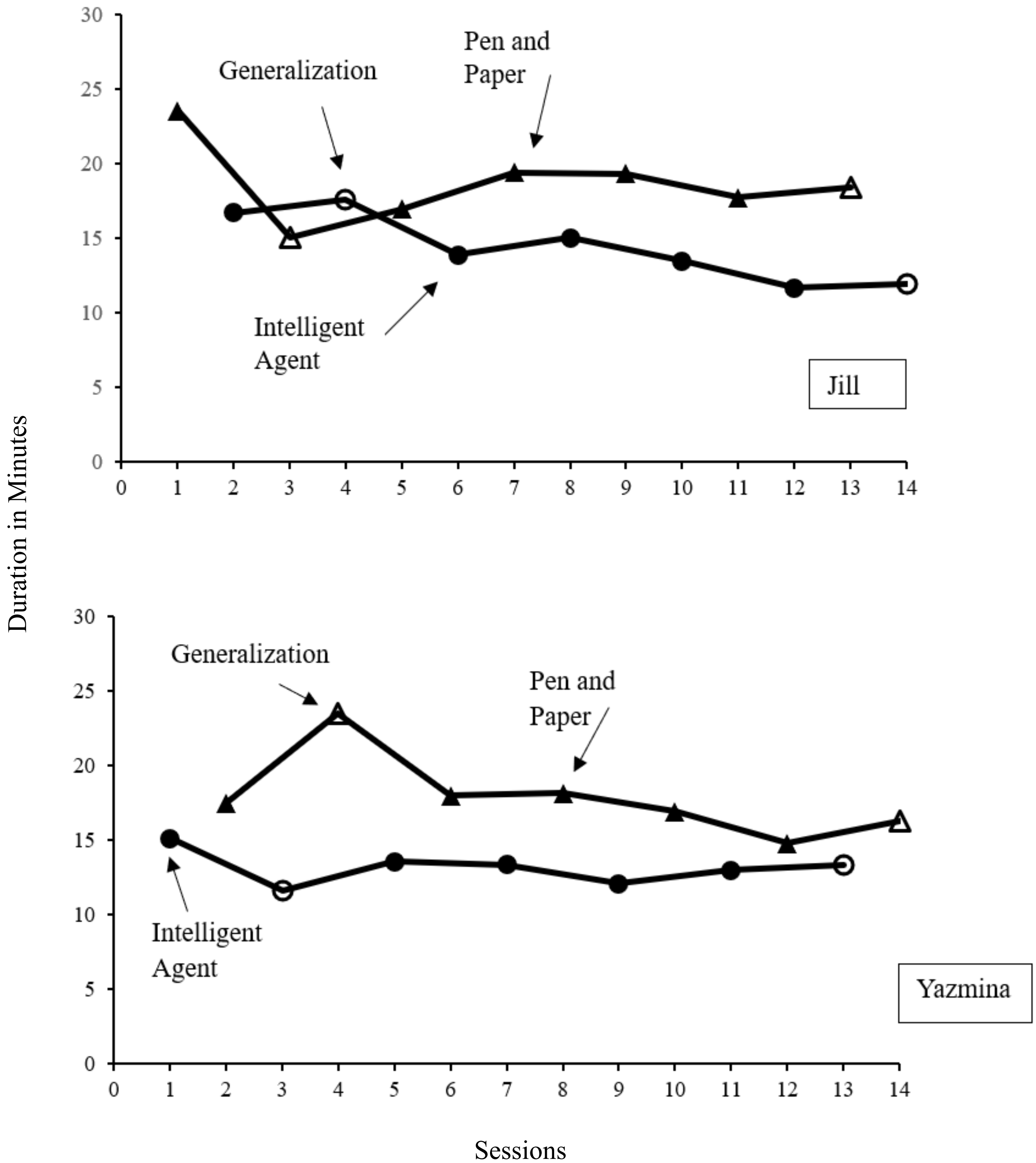


Figure 4. Duration of MSWO for Jill and Yazmina. This figure depicts the duration of each MSWO session for Jill and Yazmina.

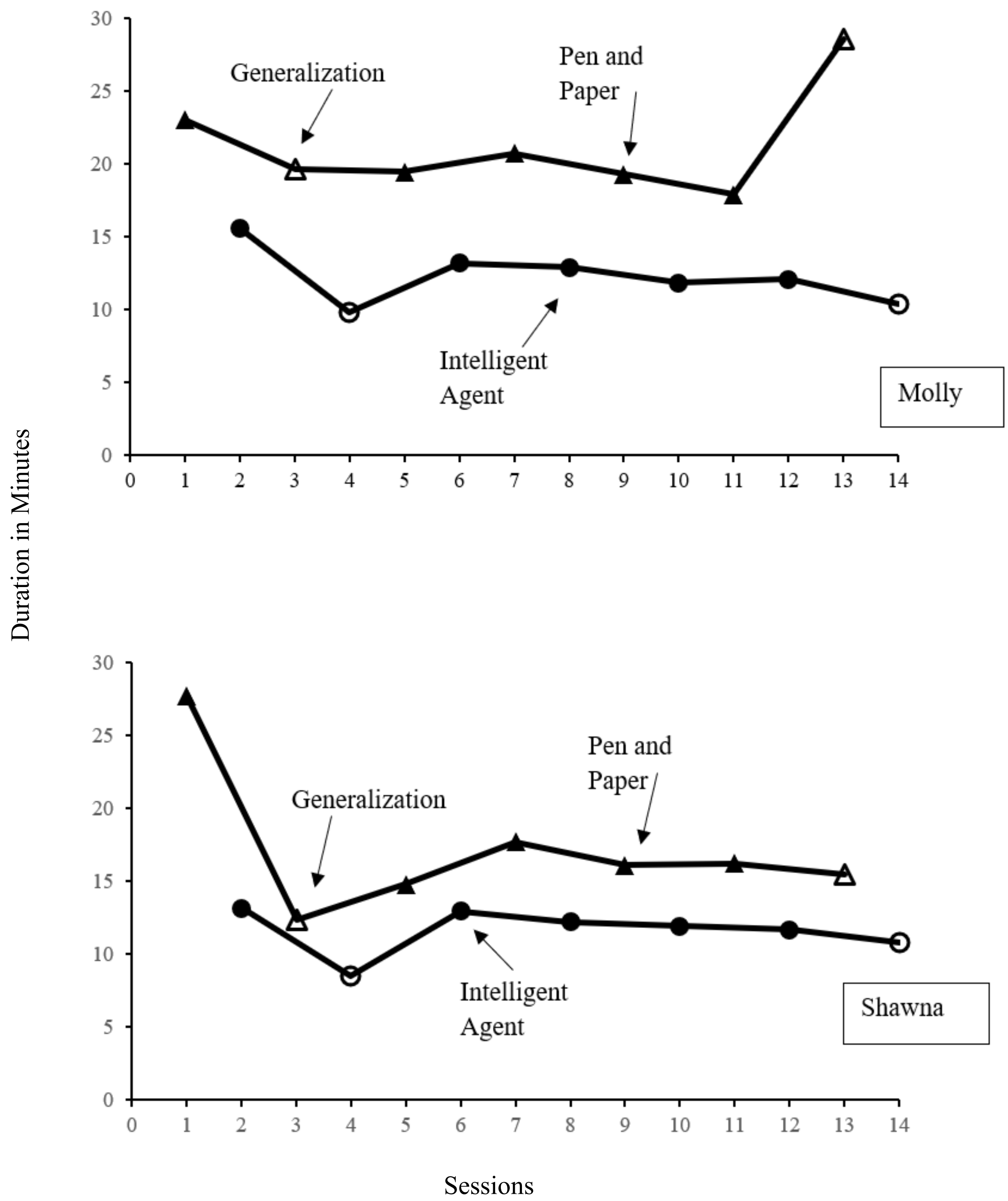


Figure 5. Duration of MSWO for Molly and Shawna. This figure depicts the duration of each MSWO session for Molly and Shawna.

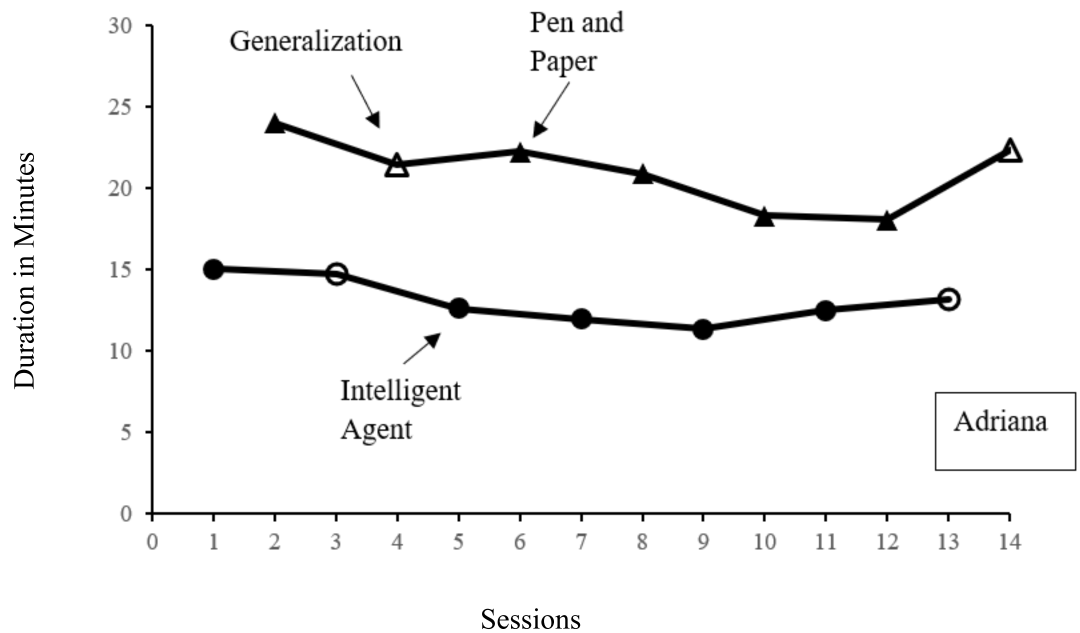


Figure 6. Duration of MSWO for Adriana. This figure depicts the duration of each MSWO session for Adriana.