TARGET ACQUISITION AND RELATEDNESS TO EMBEDDED PLAY

by

Grace O. Feiner

Thesis

Submitted to the Faculty of

Peabody College of Vanderbilt University

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF EDUCATION

in

Special Education

May 2023

JR Ledford	03/21/2023
Major Professor	Date
Ann P. Kaiser	3/21/2023
Second Reader	Date
	3/21/2023
Department Chair	Date
	4/21/23
Dean of Peabody College	Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Teachers have multiple ways to teach young children, including embedded, distributed, and massed trial instruction. Massed trial instruction consists of rapid repeated trials, and distributed instruction includes arranging trials throughout the day. Embedded direct instruction involves incorporating learning opportunities into classroom activities and routines, such as centers, free play, small group instruction, and transitions.

Instructional Strategies

While all teaching arrangements are effective for some children, direct instruction through massed trials is resource-intensive and may not be feasible in classrooms with a low ratio of teachers to students. Furthermore, distributed trials in naturally occurring learning opportunities may not provide sufficient opportunities to respond for some children to learn certain skills. There is a growing body of literature demonstrating that embedded direct instruction is an effective and efficient alternative to target high priority learning objectives without the use of massed trial instruction or distributed trials (Grisham-Brown et al., 2002; Ledford et al., 2017; Shepley & Grisham-Brown, 2019). Using embedded strategies, direct instruction can be built into daily routines and activities to provide more structured or programmed learning opportunities.

Research demonstrates that for some children, embedded direct instruction is an effective teaching method for changing child behavior (Ledford et al., 2017). This strategy can be used to meet a variety of academic, social, and life skills goals (Venn et al., 1993; Werts et al., 1992; Wolery et al., 2002). Skills that have been taught through embedded direct instruction include sight words, vocabulary definitions, counting, multiplication problems, shape and color identification, sharing, and peer imitation (Daughtery et al., 2001; Lane et al., 2015; McDonnell et al., 2006; Venn et al., 1993; Werts et al., 1992; Wolery et al., 2002). Additionally, teachers have reported positive perceptions and increased use of embedded teaching procedures following various coaching interventions along with corresponding increases in child acquisition of learning objectives (Horn et al., 2000).

Instructional Arrangements

Given massed trial instruction, research supports small group instruction with massed trials because provides more learning opportunities for students with disabilities and their peers in comparison to other types of instruction, such as large group art activities (Venn et al., 1993), transitions, (Werts et al., 1992), and circle time (Wolery et al., 2002). Small group instruction has led to rapid acquisition of letter name and sound identification as well as increases in phonological awareness fluency (Olszewski et al., 2017). Additionally, McDonnell and colleagues (2006) conducted a comparison of 1:1 embedded instruction and small group instruction in teaching vocabulary words to middle school students with developmental disabilities. Results show that small group instruction with two additional peers was equally or more effective in promoting acquisition and generalization of the target skills in comparison to 1:1 instruction (McDonnell et al., 2006). Furthermore, in a review of 190 studies using small group instruction to teach discrete skills, Ledford and colleagues (2012) found that nearly all participants met mastery criteria (Ledford et al., 2012).

Small group instruction is also beneficial as it allows for students to acquire information through targets delivered to their peers via observational learning, in addition to their own targets (Ledford et al., 2012). Out of a selection of 33 studies that measured acquisition of peers' target skills, participants learned an average of 63.3% of their peers' targets (Ledford et al., 2012). For example, direct instruction of sight words provided to dyads of preschool students resulted in acquisition of targeted sight words as well as up to 100% of their peers' sight words (Lane et al., 2015). Similarly, preschool students in small group triads consisting of two typically developing peers and one with a diagnosed developmental or learning disability acquired individualized target academic skills, as well as up to 100% of the academic behaviors taught to their peers (Ledford & Wolery, 2013).

However, research on embedded instruction has largely focused on individual instruction rather than instruction that occurs in groups of children. That is, although multiple children may be present during an activity such as free play, instruction was assessed with only one child (Daughtery et al., 2001; Venn et al., 1993; Werts et al., 1992; Wolery et al., 2002). Most research on small group instruction suggests that it is helpful and efficient, as programmed opportunities can be used to target skills for multiple students simultaneously, (McDonnell et al., 2006; Olszewski et al., 2017), but little is known about the effectiveness of embedded instruction in small groups.

Considerations for Small Group Embedded Instruction

Embedded instruction can be defined and implemented in a variety of ways. Researchers have defined embedded instruction as instruction provided through trials inserted into ongoing routines and activities (McDonnell et al., 2006; Wolery et al., 2002). Some utilize more didactic strategies when implementing embedded instruction, while others are more natural and child-led trials (Daugherty et al., 2001). For example, some researchers utilize child-initiated instruction and teach through naturally occurring learning opportunities (e.g., a teacher begins using a prompting procedure to prompt a child to share with a peer after acquiring two toys), while others begin teacher-directed trials with a task direction (Daugherty et al., 2001; Lane et al., 2015; McDonnell et al., 2006; Werts et al., 1992). Furthermore, some view embedded instruction as that which occurs within one activity, while others prefer to distribute trials between or across activities (McDonnell et al., 2006; Werts et al., 1992). Regardless of the type of embedded strategy, embedded instruction should, to some extent, expand, modify, or adapt an activity to provide additional opportunities for learning (Daugherty et al., 2001).

Procedure

One consideration for embedded instruction is the strategy that should be used for teaching. Constant time delay (CTD) is an effective and feasible procedure for implementing embedded instruction. With CTD, an implementer presents a task direction, waits for a pre-determined interval to allow the participant to respond, then provides a controlling prompt that supplies the learner with the correct answer. Initial sessions will have a 0 s wait interval before the controlling prompt, resulting in near errorless responding and decreased learner frustration. After mastery criteria is reached with a 0 s delay, subsequent sessions will have a terminal delay interval, such as 3 s, before provision of the controlling prompt to allow for independent responding (Ledford et al., 2019). Ledford and colleagues identified CTD as the most common prompting procedure utilized in small group instruction to teach discrete skills (2012). It is simple to implement, and reasonable for students. CTD is particularly suited for learners with discrete skillsets and the ability to wait for a prompt. In addition to near errorless learning opportunities, CTD produces high response rates (Ault et al., 1988b; Doyle et al., 1990). CTD has been used in embedded instruction to teach counting (e.g., Daugherty et al., 2001), shapes (e.g., Werts et al., 1992), and multiplication problems (e.g., Wolery et al., 2002). It has also been used in small group instruction to teach number words and Roman numerals (e.g., Holocombe et al., 1993), sight words (e.g., Lane et al., 2015), and vocabulary definitions (e.g., McDonnell et al., 2006).

Another prompting procedure is progressive time delay (PTD), in which the wait time between the provision of the task direction and the controlling prompt gradually and systematically increases across sessions according to pre-determined criteria for student performance. During the initial sessions using PTD, an implementer will provide a task direction and immediately provide the controlling prompt, demonstrating a 0 s wait interval. After the student reaches criteria to progress to the next delay interval, the implementer will wait 1 s to allow for independent responding before providing the controlling prompt. This process continues with increasing the delay interval in small increments (typically 1 s) until the student reaches criteria for a terminal delay, such as 3 s, before prompt delivery (Ledford et al., 2012; Collins et al., 2018). This procedure is similar to CTD with the exception of the incremental increases in wait time rather than progressing directly from a 0 s delay to a terminal delay (Ledford et al., 2012). It may be argued that CTD is easier to implement as it does not require instructors to remember the correct wait interval across sessions and participants. However, PTD may be more appropriate for learners who have difficulty waiting for a prompt (Collins et al., 2018).

PTD has been used in both small group and embedded instruction, but not both small group and embedded instruction simultaneously. PTD has been used in small group instruction to teach peer imitation (Sweeney et al., 2018), social problem solving (Korba et al., 2021), sharing and saying “thank you” as well as word reading and naming colors (Ledford & Wolery, 2013), sight word reading and commenting toward peers (Urlacher et al., 2016), word reading, shape naming, sharing, and providing social feedback (Ledford & Wehby, 2014), and reading sight words (Winstead et al., 2019). PTD has been used in embedded instruction to teach peer imitation (Francis et al., 2020), letter sounds, letter names, subtraction, word naming, shape names, and peer names (Ledford et al., 2017), and manding using Siri (Calzi, 2020).

Instructive Feedback

Instructive feedback, which is sometimes used with direct instruction, is additional nontarget information typically provided alongside a consequence or reinforcement statement after the target response (Werts et al., 1995). For example, an implementer may provide the task direction, “What color is this?” while presenting a visual cue displaying a blue square. If the participant responds, “Blue”, the implementer will reply, “Great job, this is blue. And squares have 4 sides”. The extent that feedback is related to the target information falls into the following three categories: feedback can be (a) parallel to the target stimulus (target and feedback stimuli differ but produce similar responses), (b) an expansion of the target stimulus (target and feedback stimuli require different responses but the two stimuli are conceptually related), or (c) novel (target and feedback stimuli are not related or from the same skill domain and do not produce a similar response; Ledford et al., 2019); Werts et al., 1995). Regardless of the type of instructive feedback (e.g., parallel, expansion, or novel), students acquire skills with a high level of success (Albarran & Sandbank, 2019).

Relatedness of Target Stimuli

Relatedness can be defined as the presence of a shared physical dimension between target and nontarget components of direct instruction that is not required for instruction or reinforcement but that can be used to support acquisition. Instructive feedback provides one context in which the relatedness of target stimuli has been studied, but current knowledge is limited. One existing study directly compares acquisition of related and unrelated target and feedback stimuli. Results show that students learned more of the unrelated (novel) feedback stimuli than related (parallel) stimuli (Werts et al., 1993). However, when the academic domains of each stimuli group (novel and parallel) were reversed, students learned an equal or greater amount of the related stimuli than unrelated stimuli. These findings suggest that the academic domain of the stimuli had a greater influence on target acquisition than the relatedness of the feedback (Werts et al., 1993). There is, therefore, no known effect of the relatedness of target skills and feedback information on learning efficiency.

While instructive feedback provides evidence of acquisition of skills that are related and unrelated to targets, the concept of relatedness of target stimuli has rarely been applied to the context of activities with embedded discrete trials.That is, only one study was identified that explicitly compared the efficiency of target skill acquisition embedded into an activity that is related to the target skills with one that is unrelated to the target skills (Botts et al., 2014).

Botts and colleagues (2014) attempted to do so by comparing acquisition of phonological awareness skills during embedded direct instruction with targets that were unrelated to the activity and an activity-based intervention with targets that were related to the activity. In the embedded direct instruction condition, the implementer used CTD and scripted antecedents and consequences during an adult-directed activity. In the activity-based intervention condition, trials occurred during a child-directed activity, and the implementer did not use a systematic prompting procedure. Rather, naturally occurring trials were initiated following logical antecedents (such as a peer model) and consequences that were connected to the desired response (Botts et al., 2014). Botts and colleagues found that embedded direct instruction (unrelated target stimuli) was more effective and efficient in the acquisition of phonological awareness skills in comparison to the activity-based intervention (related target stimuli). However, the results are difficult to interpret due to differences in procedures across conditions (e.g., presence and absence of a prompting procedure) (Botts et al., 2014).

While researchers have attempted to identify connections between learning and relatedness of targets, little is known about the efficiency of skill acquisition in embedded instruction when targets are related and unrelated to an ongoing activity. It might be that if learners are presented with one set of target stimuli that are related to an embedded activity, and another set of target stimuli that are unrelated to an embedded activity, learners will demonstrate more efficient acquisition of the related targets. This is in part because targets will incorporate concepts pertaining to an activity in which they are already interested and engaged. The Division of Early Childhood (DEC) also recommends that instruction is provided through contextually relevant learning opportunities, which is more descriptive of related stimuli than unrelated stimuli (2014). In addition, DEC supports the use of evidence-based practices and monitoring and evolving such practices to meet the everchanging needs of students, teachers, and practitioners.

Research Questions

As research continues to support the importance of small group instruction as well as embedded instruction, it becomes increasingly critical that teachers have access to the most feasible and efficient strategies for teaching instructional targets for one or more children. Therefore, the purpose of the current study is to evaluate discrete academic skill acquisition when adult-directed trials are embedded into a small-group activity with multiple target students, each with their own target skills. Specifically, this study will compare the efficiency of skill acquisition when target stimuli are embedded into an activity that is related to the target skills as opposed to an unrelated activity.

The following research questions will be addressed: (a) Do children learn targeted behaviors via small group embedded instruction? (b) Do learners acquire instructional targets more efficiently in small group embedded instruction if targets are related to play, rather than unrelated to play?; (c) Are skills embedded in a small group instructional context retained in a classroom environment when not embedded into an activity?; (d) Are skills embedded in a small group instructional context retained with a novel implementer?; (e) Are skills embedded in a small group instructional context retained in the same environment after 4 weeks without direct instruction?; (f) Do learners prefer to engage in sessions in which trials are embedded into an activity that is related to the targets, rather than unrelated to targets?; and (g) Do learners demonstrate increased engagement in play sessions without trials, rather than those with trials?

CHAPTER II

METHOD

Participants

Participants include students in a university-affiliated research-based preschool in southeastern US. Inclusion and exclusion criteria are shown in Table 1.

Table 1

Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Age between 24 and 60 months	Inability to attend in a small group activity for approximately 5 min
Caregiver or teacher identified a discrete academic skill that the child needs to acquire and that is confirmed in screening sessions	Absence of vocal-verbal communication
Ability to stay within proximity of peers in a small group free play setting	Engages in frequent peer-directed challenging behavior
Ability to stay engaged for the duration of a 5 min activity given teacher proximity	Engages in challenging behavior when adult attention is being provided to a peer (i.e., diverted attention)

The researcher recruited 4 participants after observations and interactions with students as well as discussions with classroom teachers. The researcher provided caregivers with a consent form detailing the purpose and expectations of the study, as well as risks and benefits of participation. The researcher also asked the caregiver to describe any desired academic skill to target during intervention. Caregivers provided either perceived areas of need for further instruction, or areas that would serve as an academic challenge for their child. In addition to parent consent, child assent was assessed daily by asking “Do you want to play with me today” (or a similar question).

Participants were between the ages of 45 and 51 months at the onset of the study and included two children who received early intervention services from birth to three but did not have a diagnosis, one typically developing child and one child diagnosed with autism. The researcher recruited pairs of participants within the same age group and classroom. Dyad 1 consisted of participants Elsa and Anna. Elsa and Anna were twin sisters who were 51 months old and received early intervention services. Elsa received 55 minutes of physical therapy each week, and Anna received 55 minutes of speech therapy twice per week in addition to 55 minutes of occupational therapy every other week. Their parent reported their race/ethnicity as “two or more races”. Dyad 2 consisted of participants Minnie and Mickey. Minnie was a 45-month-old typically developing White female. Mickey was a 48-month-old Autistic Black male. He had emerging oral language skills largely consisting of 4–5-word sentences and utterances. He received physical therapy and occupational therapy each once per week for 55 minutes, and speech therapy twice per week for 55 minutes.

The primary implementer was a White, non-Hispanic female completing a graduate degree in special education and accruing hours for behavior analysis certification under the supervision of a doctoral level BCBA who was also a White woman. The primary implementer conducted all sessions except for massed trial sessions for two participants (Anna and Elsa) after 5 weeks without instruction and generalization sessions for Mickey and Minnie. The implementer for these sessions were two White female graduate students enrolled in the same special education program as the primary implementer. The data collectors included the primary and one secondary implementer as well as one Indian female graduate student who was also enrolled in the same special education program. All implementers served as research assistants in a Preschool and Early Elementary Learning Lab.

Settings

Screening, baseline, control, and intervention sessions occurred in a small resource room outside of the participants' typical classroom within a university-based preschool. Two participants and an implementer sat on the floor with an activity placed on a carpet. Generalization and maintenance sessions occurred in the participants' classrooms. A secondary implementer conducted generalization sessions in a center. Approximately 12 non-participating peers and two non-participating adults were present in the classroom. Two or three peers were engaged in each center around the classroom, and adults facilitated activities in centers. No non-participating peers were engaged in an activity within the same center as the participants during generalization and maintenance sessions.

Materials

Throughout all conditions, the implementer collected skill acquisition data in Excel using a data collection sheet (see Appendix A) as well as engagement data through observational software using ProCoderDV (Tapp, 2003; see Appendix B). The implementer recorded all sessions using a video camera. Materials used during sessions varied based on the participants, their identified targets, and the condition (related and unrelated to targets) but consisted of items that would typically be found in centers in a classroom, as well as a game created by the researcher. Items included wooden blocks, toy cars, sand, kinetic sand, Magnatiles, Legos, Play-doh, pretend food, rubber balls, ping-pong balls, buckets, and animal figurines. Pom-poms or plastic tokens were used for subitizing in unrelated and control conditions for Anna in Dyad 1. Visual cues for color and letter targets for participants Mickey and Minnie, respectively, were displayed on laminated index cards or commercial cards in unrelated and control conditions. Colored 3x5 carpets (red, green, and black), were also included and corresponded with the experimental condition (Table 2).

Table 2

Materials and Targets Across Dyads and Conditions

		Related	Unrelated	Control
Dyad 1	Carpet	Red	Green	Black
	Toys	Magnatiles Play-doh	Wooden blocks Kinetic sand	Bristle blocks Sand
	Trial-Based Additives	Sea animals Dinosaurs	Farm animals Insects Pom-poms Plastic tokens	Dogs Zoo animals Pom-poms Plastic tokens
	Letter Sounds (Elsa)	D H	R M	T B
	Subitizing (Anna)	6 9	5 8	7 10
Dyad 2	Carpet	Red	Green	Black
	Toys	Play-doh Buckets Ping-pong balls	Magnatiles Wooden blocks Cars	Sand Legos
	Trial-based Additives	Letter stamps Post-its	Commercial color cards Letter flash cards	Commercial color cards Letter flash cards
	Letters (Minnie)	E H	L T	F Y
	Colors (Mickey)	Pink White	Teal Brown	Tan Gray

Response Definitions and Measurement Systems

Target Skill Acquisition

The primary dependent variable in this study was the efficiency of target skill acquisition. Efficiency was determined by the number of sessions to mastery of targets. Each session consisted of two targets presented three times each for a total of six trials per participant and 12 trials per session. Mastery criterion was met when the learner produced an unprompted correct response in 5 out of 6 of the trials for three consecutive sessions in each experimental condition. To measure mastery of target acquisition, the researcher collected data in Excel using event recording to record whether the participant produced (a) an unprompted correct response, (b) a prompted correct response, (c) an unprompted error, or (d) a prompted error. The researcher then calculated the percentage of unprompted correct responses produced in each session. Experimental decisions were made based on progress monitoring of this measurement.

Unprompted correct responses were defined as vocally producing the correct response before a vocal model was presented within 10 s after the entire task direction was stated, and the visual cue was presented, if applicable. The implementer began counting 10 s at the offset of the task direction. Examples included producing the response “blue” within 10 s of being given the task direction of “What color is this?” with a blue visual cue and producing the response “Four” within 10 s of being given the task direction, “How many are there?” and presentation of four stimuli. Non-examples included producing the correct response after the controlling prompt or imitating the task direction. A non-example of unprompted correct responding included producing the response, “blue” after the controlling prompt, “blue” is provided following the task direction, “What color is this?” with a blue visual cue, or imitating the task direction (e.g., responding “What color is this?” when presented with the task direction, “What color is this?”).

Prompted correct responses were defined as vocally producing the correct response within 10 s following the provision of the vocal model. Examples included producing the correct response, “Rojo”, 10 s after the implementer provides the task direction, “How do you say ‘red’ in Spanish?” and models “Rojo” or producing the response, “Dog”, 10 s after the implementer presents a visual cue of a picture of a dog, provides the task direction “What is this?”, and models “Dog”. Non-examples include producing a correct response without a model or an incorrect response after the model. Non-examples of prompted correct responding may include producing a correct response, “Dog” after provision of the task direction, “What is this?” with a visual cue of a picture of a dog, but before the implementer models the response, “Dog”, or producing an incorrect response (e.g., “Cat”) after the implementer provides a model (e.g., “Dog”).

Unprompted errors were defined as producing an incorrect response after delivery of the task direction and before provision of the model prompt. Examples included producing an incorrect response (e.g., “A”) when the implementer presents a visual cue (e.g., displaying the letter ‘B’) and provides the task direction (e.g., “What letter is this?”) without a model. Non-examples include producing an incorrect response following a model or producing a correct response with or without a model. Non-examples of an unprompted error may include producing the incorrect response, “Cat”, when the implementer provides a visual cue of an image of a dog and the task direction, “What animal is this?" as well as the model prompt or producing a correct response before or after the model prompt (e.g., responding "Yellow" when the implementer displays a visual cue of a yellow square and presents the task direction, "What color is this?").

Prompted errors were defined as producing an incorrect response after provision of the task direction and model prompt or producing no vocal response following complete provision of the task direction and/or model prompt. Examples include producing the correct response with or without the model prompt (e.g., responding, "Three", when presented with a visual cue of the numeral 3 and the task direction, "What number is this?" after a model). Non-examples include producing an incorrect response without a model (e.g., responding "B" when presented with a visual cue with the letter D and the task direction "What letter is this?" without a model), producing the correct response within 5 seconds after imitating the task direction, or producing a correct response with or without a model (e.g., responding "five" when presented with five objects and the task direction, "How many are there?" with or without a model).

Engagement

The secondary dependent variable was the percentage of engagement across sessions. Engagement data were collected through observational coding (ProCoderDV; Tapp, 2003) using momentary time sampling with 10 s intervals. At the end of each 10 s interval, the observer coded each participant's behavior as not engaged, engaged in play, engaged with instruction, or offscreen. The observer then calculated the percentage of total engagement by dividing the percentage of intervals in which the participants were engaged with instruction or engaged in play by the total number of intervals. Intervals in which a participant was coded as offscreen were not included in this calculation.

Engaged in play was defined as visually attending to, talking about, and/or physically interacting with or manipulating toys included in the embedded activity or looking at the implementer or peer while they are playing with and/or talking about the toys. Examples include commenting on play by saying “I’m going to build a house” when provided with blocks, or actively engaging in manual manipulation of Play-doh, as well as scanning, watching the peer and implementer play. Non-examples include visually attending to the floor, wall, or ceiling, or posing statements or questions unrelated to the activity or targets (e.g., “Is it raining today?”).

Engaged with instruction was defined as visually attending to the implementer, target stimuli, or visual cue while the task direction is being delivered and producing a vocal response within the same response class as the target response or repeating or commenting on a target after the corresponding task direction was initially given. Examples included looking at the visual cue “Blue” and responding with a color when presented with the task direction, “What color is this?” or looking at the implementer and responding “Verde” when presented with the task direction, “How do you say ‘green’ in Spanish?”, as well as making a comment related to a peer’s target (e.g., “I like shoes” following target of “How do you say ‘shoes’ in Spanish?” is delivered to peer, or repeating a response to a target more than 15 seconds after an initial response. Non-examples include engaging in unrelated conversation, such as, “I’m going to the park this weekend”, after being presented with the task direction, “What letter is this?” and a visual cue of “A” or looking at the toys and producing no vocal response within 10 s of being presented with a task direction or controlling prompt. Additionally, receiving or attending to a reinforcer was not considered to be an example of engaged with instruction.

Engaged with play and engaged with instruction were coded such that they were mutually exclusive. Therefore, if a child was manipulating play materials while attending to a trial, the coder selected the most salient behavior. For example, if a child was engaged in play and responding to a task direction, the coder selected “engaged with instruction”, as the child had to be attending to the task direction in order to provide a response. However, if a task direction was given and the participant did not respond, or responded with a comment unrelated to the task direction, then the coder selected “engaged in play”, as the child was likely ignoring the delivery of instruction.

Not engaged was defined as visually attending to the implementer, peer, or non-play materials without engaging in appropriate play or conversation; not giving a response to a task direction within 5 s of the offset of the task direction or producing a response that is not within the response class of the target response. Examples include asking, “What did you have for dinner last night?” while looking at the implementer following provision of the task direction, “What number is this?”, repeating the task direction and not giving a response, staring at the implementer without engaging in play or conversation or watching their play behaviors, crying or tantruming, receiving or attending to reinforcers, and visually attending to the ceiling without actively manipulating the toys placed on the table. Non-examples include producing a vocal response of “Red” within 5 s of the task direction or controlling prompt, manually manipulating toys such as blocks by building a tower and commenting “My tower is going to be so tall!” or talking about the toys while offscreen.

Offscreen was defined as not visible from the perspective of the camera and not vocalizing or producing noise to the effect that an observer could not draw a conclusion about the behaviors in which the child was engaged. Examples include complete silence and absence from the camera view. Non-examples include talking about the toys while outside the view of the camera (e.g., “I’m making an elevator!” while out of view).

The researcher assessed generalization across settings and implementers. The researcher collected data pertaining to target skill acquisition and engagement using the same measurement procedures as baseline and intervention sessions but with a novel implementer. Maintenance sessions occurred after 4 weeks without direct instruction and occurred in the classroom with the primary implementer. Measurement procedures were the same across all conditions.

Interobserver Agreement

A secondary observer collected interobserver agreement (IOA) for 33% of sessions in all conditions and for each participant. A weekly meeting was held between the primary and secondary observers to discuss all discrepancies. IOA for target skill acquisition was collected from video using the same data collection sheet used for primary coding, and IOA for engagement was collected from video using observational software (ProCoderDV; Tapp, 2003). All IOA calculations were performed in Excel. To collect IOA for target skill acquisition, the observer recorded whether the participant produced a prompted correct response, unprompted correct response, prompted error, or unprompted error. Point-by-point agreement was calculated for each trial and participant in a session by dividing the total number of agreements across all trials for each participant by the total number of trials for each participant to calculate a percentage of agreement. IOA was calculated separately for participants within a dyad. A minimum of 90% agreement within a session with 12 trials was considered acceptable for target skill acquisition as these are discrete, trial-based academic skills with little to no subjectivity. To collect IOA for engagement, the secondary observer used momentary time sampling with 10 s intervals. The total number of intervals in which the primary and secondary data collectors agreed was divided by the total number of intervals in the session to generate a percentage of agreement. A minimum of 80% agreement was considered acceptable for engagement due to the subjective nature of the variable.

Observers were trained by the primary researcher prior to data collection by performing consensus coding and practicing coding independently. The primary researcher also provided explicit teaching of the operational definitions for each dependent variable. Observers performed consensus coding on two sessions with the primary researcher (approximately 10 min each in length). Observers were then assigned two videos to practice coding independently. For these sessions, observers had to reach at least 90% agreement for target skill acquisition and at least 80% agreement for engagement. The researcher met with secondary observers to discuss all discrepancies. If IOA fell below the acceptable criterion, the primary researchers and secondary observers were to perform consensus coding for another session and discuss potential changes that may need to be made to the operational definitions to make them clearer and more distinct.

Experimental Design

Adapted Alternating Treatments Design

An adapted alternating treatments design was used to assess and compare the efficiency of target skill acquisition between trials embedded into a related activity and unrelated activity. AATDs are typically used to compare instructional practices for teaching non-reversible behaviors, such as discrete academic skills (Ledford & Gast, 2018). In an AATD, sessions alternate between two distinct interventions, and a control condition. Each condition is applied to a different, but functionally equivalent, behavior set (Ledford & Gast, 2018**).** The behavior sets each consisted of six brief discrete responses of the same topography (e.g., expressive vocal verbal response) and level of difficulty. Target skills included in the behavior sets varied across participants, but included initial letter sound identification, subitizing, letter identification, and color identification. The level of difficulty was assessed through logical analysis of the responses and discriminations required for the response to be correct (Ledford & Gast, 2018). Logical analysis included appropriate dimensions in accordance with the target skill. For example, if a student were to learn to name colors, the behavior sets were be evaluated for the number of syllables, initial and final consonants, repeated use of letters across words, and the participant's current ability to recognize and say each word (Ledford & Gast, 2018).

An AATD was an appropriate design for this study because it allowed for rapid comparisons of two distinct interventions and their impact on the efficiency of skill acquisition, which is a non-reversible behavior. Additionally, this design enabled analysis of skill acquisition in a small group context, as multiple participants experienced the same intervention simultaneously, but with their own individualized targets. Neither intervention was withdrawn at any time, and reversibility of the behavior was not required. Additionally, a social validity measure of child preference was included in this design by allowing the child to report their preferred condition in randomized sessions throughout the study. This design helps to answer comparison research questions without the risk of separation of treatments, because each intervention is applied to a different set of behaviors.

Internal Validity

Common threats to internal validity, such as testing, instrumentation, and attrition, are less likely to occur in comparison to other designs due to the relatively short duration of an AATD. The inclusion of intermittent control sessions also provides opportunities to detect history and maturation effects as it shows whether participants may have learned behaviors in contexts outside of the study (Ledford & Gast, 2018). If participants show evidence of learning behaviors in the control set, then it is likely that non-study exposure to behaviors in the intervention sets have also occurred.

The researcher minimized the threat of procedural infidelity by conducting all baseline and intervention sessions and having a data collector record trial-based behaviors and play-based behaviors to calculate procedural fidelity on a minimum of 33% of sessions across each participant and condition. The researcher minimized this threat in generalization sessions by training implementers across all conditions and providing easily accessible reference sheets for procedures. Additionally, the researcher assessed procedural fidelity formatively across sessions and conducted ongoing training for implementers if needed. The procedure of this design minimized the threat of multi-treatment interference as the implementer conducted sessions of alternating treatments across days, rather than within the same day, to decrease the chance of one intervention influencing learning in another intervention. Finally, applying each intervention to a different set of behaviors minimized the risk of separation of treatments, and logical analysis and random assignment of behaviors to conditions minimized the risk of inequality in the target skill difficulty.

Visual Analysis

To conduct visual analysis, the researcher evaluated pairwise comparisons of each intervention in the comparison condition. Only within participant data was compared between the two interventions to determine the rapidity of target acquisition. The percentage of unprompted correct responses from a participant in sessions of the related condition was compared to the percentage of unprompted correct responses from the same participant in sessions of the unrelated condition. More specifically, the researcher assessed changes in the level and trend of unprompted correct responses for each participant as they progressed through sessions of each intervention (Gast & Spriggs, 2019). The primary researcher used this data formatively to determine if any procedural modifications were necessary and appropriate, and to determine whether participants were making adequate progress. For example, when Elsa and Anna stopped progressing in skill acquisition and exhibited lower levels of engagement due to tantruming and arguing, the implementer introduced new reinforcers. Additionally, the primary researcher used skill acquisition data to guide decisions on instructional procedures and whether massed trials would be a more efficient strategy. The condition in which a participant reached mastery criterion in fewer sessions was deemed the most efficient intervention for that participant.

Procedures

Screening Condition

The researcher used caregiver request and teacher report to select a socially valid target behavior for each participant. The researcher then conducted screening sessions to determine whether the target skill was in the participant's repertoire and confirm that the student could attend to a task direction while engaged in an activity alongside a peer and maintain engagement for 10 minutes. If the participant did not attend to a task direction while engaging in an activity or engaged in peer-directed aggression while working in a small group, then the participant was screened out of the study (this never occurred). If the participant attended to the activity alongside a peer but the implementer determined that a participant was already familiar with the targets, new targets were identified and tested following the same procedures.

During these sessions, the implementer directly tested for student knowledge of 12 – 104 target items using massed trials. The implementer conducted trials in rapid succession without prompting and alternated between participants within a dyad. The researcher presented each question (e.g., What sound does R make?) one at a time and, for some target items (e.g., color identification, letter identification), showed a visual cue while stating the task direction vocally (e.g., “What is this?”). Following completion of trials, the implementer and students engaged in a child-directed play activity that was assigned at random. The researcher and child were seated on the floor in a resource room at the school used for outside-of-classroom activities (e.g., special events, after care, individual therapy). Trials and the activity were completed on the floor.

Baseline Condition

During baseline, all behavior sets for each participant were assessed for a minimum of three sessions. Baseline sessions served as the initial probe sessions to measure knowledge of targets before instruction. Baseline sessions consisted of the implementer and two participants sitting on the floor in the art room within the preschool. The implementer presented all 6 targets (i.e., two targets each assigned to related, unrelated, and control conditions) for each participant using intermixed massed trials and did not provide any prompting. To maintain engagement for some participants (Elsa and Mickey), the implementer interspersed familiar targets within the novel behavior sets. The implementer provided a brief task direction and visual cue where appropriate (e.g., “What letter is this?” with a picture of the letter “a”) and gave the child 5 s to respond before delivering the next task direction. The implementer alternated between the two participants and their respective targets. Correct responding was reinforced with a brief praise statement to not unintentionally suppress correct responses. Across all baseline sessions, both members of each dyad participated simultaneously, the same implementer conducted sessions and provided the same number of opportunities to respond, and all sessions were approximately 5 – 10 min in duration.

Comparison Condition

Prior to the comparison phase, behavior sets of 2 different targets from the pool of 6 were assigned to each condition (control, time delay with *related* stimuli, and time delay with *unrelated* stimuli). During the comparison phase, both interventions were applied to their respective behavior sets and control sessions were probed intermittently. Session order was determined randomly, using block randomization (e.g., one intervention type was randomly selected, and then the other type was conducted next before repeating the next random selection). To aid in discrimination, play materials were placed on a red carpet for the related condition, a green carpet for the unrelated condition, and a black carpet for control sessions.

Control Sessions

During the comparison phase, control sessions occurred intermittently and included procedures identical to baseline sessions, with one exception. Instead of presenting targets from all behavior sets, the implementer only presented the 2 targets randomly assigned to the control behavior set. Each target was presented three times, for a total of 6 trials per participant. The implementer was seated on a black carpet on the floor with two participants and delivered task directions alternating between each participant. Task directions were delivered in rapid succession, as there was no prompting, and targets were not embedded into an activity. Massed trials occurred first, and then implementer provided an activity for the participants to engage in after completing the entire behavior set. The activity was randomly selected from those that were assigned to the control condition. The length of these sessions remained at approximately 10 min, with 5 min allotted for massed trials and 5 min allotted for engaging in the activity.

Intervention: Related Stimuli Sessions

The targets assigned to the *related* condition were embedded into an activity in a way that was related to the activity itself. For example, during related sessions, if the targets were colors, then one set of colors were taught during an activity such as Play-doh, with the corresponding colors of Play-doh made available (e.g., if a child was playing with blue play-doh, the teacher points to the Play-doh and says “I love that you’re making a horse—what color is it?”). All aspects of the baseline condition were kept the same during *related* sessions, except for the provision of prompting, and embedding targets into an activity. During intervention sessions, the implementer taught the 2 targets that were assigned to that condition. Targets were presented three times each, for a total of 6 trials per participant in every intervention session.

During sessions, the implementer and two participants engaged in an activity on the floor and the implementer provided a task direction and visual cue (where appropriate) approximately every 30 s (about once per min per child). To begin each trial, the implementer ensured the child was attending, presented a stimulus or drew attention to the item of interest (i.e., the toy related to the a trial; e.g., “Cool car!”), and then provided the task direction (e.g., “What color is it?”).

The implementer used a progressive time delay procedure beginning with a 0 s delay and immediately providing the controlling prompt (vocal model) of the target response. The implementer waited 5 s for a response before re-engaging in the activity.

The delay between provision of the task direction and the controlling prompt increased to a 1 s delay after participants reach pre-determined mastery criteria for 0 s delay trials, which required 3 consecutive sessions at 100% prompted correct responding. The same criteria were used to progress to a 2 s and 3 s terminal delay. Intervention sessions continued with a 3 s delay until participants reached a priori mastery criteria of 3 consecutive sessions at 100% unprompted correct responding. Progression through delay intervals was determined individually; if one participant reached criteria for the next delay interval before the other participant, the implementer would use the appropriate delay interval for each participant rather than waiting for the other participant to reach the same criteria. Generalization sessions began after both participants within a dyad reached mastery criterion.

Intervention: Unrelated Stimuli Sessions

The targets assigned to the unrelated condition were embedded into an activity in a way that was unrelated to the activity. For example, a set of color words were taught while playing with neutral-colored wooden blocks; a set of nouns or adjectives were used to teach initial letter sounds without describing or commenting on play or play materials (e.g., a child is playing with a horse, and a teacher holds up a notecard and says "what color is this"). All other procedures were identical to the related sessions.

Generalization and Maintenance

Two generalization sessions occurred immediately after the comparison phase ended, to assess participant knowledge of the targets in their classroom setting with a novel implementer. Procedures for generalization were identical to those of baseline, except for the location and implementer. A novel implementer and the participants were seated in a center within the participants' classroom. The implementer presented all 4 targets from the intervention condition three times each in rapid succession for each participant. Before and after presenting 12 trials for each participant, the implementer engaged in play with the participants. During sessions, non-participating peers in the classroom were engaged in centers, though no non-participating peers were engaged in an activity within the same center as the participants. The two adults in the classroom were facilitating activities in other centers.

Maintenance sessions occurred after four weeks without direct instruction. Sessions followed procedures identical to those of baseline but occurred in the classroom instead of a separate room. Maintenance sessions functioned as the final probe condition, in which all behavior sets, including the control set, were assessed in an intermixed manner without prompting. Before and after presenting 18 trials for each participant (3 trials for each of the 6 targets), the implementer engaged in play with the participants.

Procedural Fidelity

Procedural fidelity was measured by assessing adherence to expected trial-based and play behaviors and differentiation across conditions. Trial-based behaviors in baseline, generalization, and maintenance sessions included ensuring that the participant is attending, presenting all 6 trials for each participant, providing no prompting, reinforcing correct responses, giving a neutral statement in response to incorrect or nonresponses, and including an inter-trial interval of approximately 5 to 10 s.

Additionally, the sequence of targets should be randomized prior to each session (see Appendix E). The same behaviors were expected in the control condition, except that only two targets were provided to each participant. Trial-based behaviors in the comparison condition included ensuring that the participant was attending, presenting all 6 trials for each participant, providing the appropriate wait interval before prompting, giving the appropriate controlling prompt, reinforcing correct responses, giving a corrective statement for prompted errors, and giving a corrective and wait reminder for unprompted errors, maintaining an inter-trial interval of approximately 30-90 s, and randomizing targets and play materials before each session (see Appendix C).

Play-based behaviors included the adult's behavior when not embedding a trial (see Appendix D). These applied to the play session following control, generalization, and maintenance trials, and the embedded play activity in intervention sessions. Play-based behaviors included engaging in parallel play, responding to a child's comment, and commenting on a child's play. The implementer was expected to engage in at least one of these behaviors for at least 80% of each session, estimated using 10 s intervals, via momentary time sampling. For intervention sessions, coders also indicated if the implementer was actively delivering a task direction, reinforcing a response, or providing a corrective statement, as these were acceptable behaviors during sessions with embedded trials.

Each of these behaviors was measured using direct systematic observational recording. Trial-based behaviors were recorded from video using a data collection form in Excel, and play-based behaviors were coded from video using observational software (ProCoderDV; Tapp, 2003). The observer recorded whether the implementer correctly performed each trial-based behavior in the manner in which it was intended, and as often as it was intended, for each of the 12 trials across the two participants. Play-based behaviors were assessed using 10 s momentary time sampling. Procedural fidelity was collected for a minimum of 33% of sessions across all conditions, implementers, and participants. Data analysis occurred formatively and separately for trial-based and play-based behaviors across each session. To analyze procedural fidelity, the researcher calculated the percentage of correct trial-based behaviors by dividing the total number of correct behaviors by the total number of opportunities to perform the behavior (e.g., the total number of trials within the session). The researcher also calculated the percentage of intervals of correct play behaviors by dividing the total number of intervals in which the implementer was engaging in an appropriate behavior by the total number of intervals within the session. The researcher then compared these percentages across sessions. An acceptable level of procedural fidelity was a minimum of 90%. If fidelity fell below this criterion, the implementer would have re-trained by reviewing procedures, updating a reference sheet, and role-playing with another implementer, but this never occurred.

Social Validity

Child preference is a valuable measure of social validity as children are the consumers of this intervention and their acceptance of the procedures is necessary for it to be used effectively. If children do not approve of the procedures, then the intervention may be deemed aversive or ineffective and require revisions. Furthermore, embedding choices is an evidence-based practice to increase child motivation, engagement, and independence in the classroom (Barnett, 2018). Social validity was measured by evaluating student preference between the two intervention conditions. Throughout the comparison phase, the implementer assessed student preference by intermittently asking if the participants preferred the related or unrelated condition, distinguished by the color of the carpet. To do so, the implementer prepared materials and corresponding carpets for the control condition and either the related or unrelated condition. Before beginning the session, the implementer asked each participant, "Do you like to answer the red carpet questions or the green carpet questions?". After 14 sessions of each intervention condition, the implementer began assessing child preference between embedded instruction and massed trials for Anna and Elsa only. This was done by asking, "Would you rather answer questions before we play or while we play?". The implementer would then conduct the session according to the participants' choice.

Modifications

Several modifications were made to accommodate for both anticipated and unanticipated patterns in responding.

Baseline Condition

During baseline, the implementer did not alternate between participants consistently. The implementer presented multiple targets in a row to a single participant to maintain engagement and attendance to the task. Additionally, for participants Elsa and Anna, baseline sessions were divided into two segments, each with half the number of prepared targets. This was due to Elsa's disengagement from the task and Anna's dissenting during the first baseline session. Because of the low engagement during baseline sessions with Elsa and Anna, the implementer offered one sticker at the beginning of the session, and one sticker at the end of the session, rather than offering reinforcement only at the end of sessions. The implementer also offered Elsa and Anna a choice of playing with blocks or coloring at the end of baseline sessions. Play materials used following baseline sessions were not included in intervention sessions. Finally, given Elsa's inconsistent responding when presented with words beginning with /t/ during baseline, /t/ was switched from the related condition behavior set to the control condition behavior set to ensure equal difficulty between instructional sets. This occurred prior to any sessions of the related or control conditions. Following this change, the implementer began conducting intervention sessions with Elsa and Anna despite Elsa's inconsistent responding during baseline.

Comparison Condition

Several modifications were made for Elsa and Anna during intervention sessions. It was observed that during unrelated sessions involving wooden blocks, Elsa and Anna expressed significant interest in playing with the bag in which the blocks were carried. They frequently engaged in pretend play while sitting in the bag. The implementer therefore began considering interactions with bags and containers as engaged in play.

Secondly, the implementer increased the salience of reinforcement for prompted correct and unprompted correct responding for Elsa and Anna after 12 sessions of each intervention condition. Following every prompted or unprompted correct response, the implementer presented an array of two stamps, and allowed the participant to choose one to put on their hand or leg. The implementer considered attendance toward stamps as not engaged, as the participants were not engaged in play or instruction at that time.

Thirdly, the implementer began providing explicit instruction prior to the session initiation for Anna after 12 sessions of each intervention condition. Prior to the contingency review, the implementer arranged the items as they would be during the session, described the arrangement, and named the quantity while pointing to the items and directing Anna's attention to them. Following 14 sessions of each intervention condition, the implementer began asking Anna and Elsa prior to each session if they would prefer to do massed trials or embedded instruction. The implementer did so by asking, "Do you want to answer questions before we play, or after we play?". This was due to a pattern of disengagement and expressed disapproval of instructional tasks during embedded instructional sessions. However, both participants expressed a preference for embedded instruction each time they were provided with this choice. After the 20th session of embedded instruction in each intervention condition, the implementer began conducting sessions using massed trials followed by a play session without embedded instruction. This modification was made as the participants were considered to be making inadequate progress in target skill acquisition. The implementer conducted two related and two unrelated sessions using a massed trials format before temporarily terminating instruction with the dyad. This decision was made as the participants were engaging in frequent arguing during sessions, which was ultimately resulting in high levels of disengagement. Generalization and maintenance sessions therefore did not occur as originally intended.

After 5 weeks without instruction, a novel implementer conducted a single probe session for Elsa and Anna, independently of one another. The implementer used massed trials and presented each participant with two trials for each of their four targets (two related targets and two unrelated targets) in rapid succession. The implementer did not provide any prompting. Elsa provided an unprompted correct response for one (12.5%) trial, indicating that she inconsistently responded correctly when presented with one target from the related condition. Anna provided an unprompted correct response for two (25%) of trials, indicating that she maintained one target from the unrelated condition.

After 6 weeks without instruction, the same implementer began conducting massed trials sessions with Anna and Elsa at a 0 s delay. The researcher randomized each participant's targets and selected two targets (one from the unrelated condition and one from the related condition) to use for initial massed trials sessions. For Elsa, the implementer presented a visual along with the task direction, "What is the first sound in the word *x*?". Each of the two targets were presented 6 times for a total of 12 trials per session. For Anna, the implementer used a match-to-sample procedure and provided a card with a printed arrangement of dots, as well as an array of two samples. The implementer provided the task direction of "Match", and following the correct matching of the sample, asked, "How many are there?". The implementer conducted two sessions at a 0 s delay for each participant before progressing to a 1 s delay. This rapid progression was used in order to prevent the participants from becoming reliant on waiting for a prompt before responding. This continued for 10 sessions for Elsa, and 11 sessions for Anna, before both participants transferred schools and thus data collection was terminated.

For Mickey and Minnie, two modifications were made during the comparison condition, both pertaining to selected play materials. After one unrelated session using sand, the implementer switched the unrelated materials with the control materials, so that sand would be used for forthcoming control sessions, and Magnatiles would be used for forthcoming unrelated sessions. This was due to Mickey's low levels of engagement with sand and the need to have approximately equal preference materials in instructional conditions. Additionally, after two unrelated sessions using wooden blocks, the implementer added cars to the play materials. Cars and wooden blocks were used in conjunction for subsequent sessions, as Mickey demonstrated low levels of engagement with the wooden blocks alone. Similarly, after three related sessions using Play-doh, a set of pretend food was added to the play materials for subsequent Play-doh sessions. This was also due to Mickey's low levels of engagement with Play-doh. Mickey's play engagement increased significantly following the modifications of these materials.

CHAPTER III

RESULTS

Data Analysis

Changing Conditions

Data entry and graphing occurred daily following sessions of every condition. Conditions changed from baseline to comparison after three baseline sessions with stable data from both participants within a dyad. Within the comparison condition, sessions alternated each day between related and unrelated sessions, with intermittent control sessions. Within a dyad, each participant progressed through delay intervals based on individual progress. If one participant reached criteria to progress from a 0 s delay to a 1 s delay, that participant would continue sessions at a 1 s delay while the other participant remained at a 0 s delay until they independently reached criteria to progress to a 1 s delay. The implementer reviewed data prior to each session to determine the correct delay interval for each participant.

When one participant attained mastery criteria to progress from the comparison phase to generalization sessions before the other participant, intervention sessions continued until both participants completed 1.5 times the number of sessions required for the first participant to reach mastery. For example, when Mickey attained mastery criteria in 10 sessions of the unrelated condition, the implementer conducted 5 additional unrelated sessions with both participants in the dyad to allow Minnie an opportunity to master her targets as well. This same rule was applied when Mickey attained mastery of targets in the unrelated condition before the related condition. Since he reached mastery criteria in 10 unrelated sessions, the implementer conducted an additional 5 related sessions to allow for continued acquisition of the related targets. Continuing with a limited number of sessions provides a stronger comparison between conditions, as it allows the implementer to determine if one intervention was effective at all, or if it was just less efficient than the other. It also provides an opportunity for learners with lower acquisition rates to have more exposure before considering the targets to be unmastered. After this rule was applied and sessions were conducted accordingly, the implementer determined through formative data analysis that it would be appropriate to continue sessions until both participants reached mastery criterion in both intervention conditions.

Generalization sessions occurred daily for two consecutive days immediately following completion of the comparison condition. Maintenance sessions occurred after 4 weeks without instruction, excluding generalization sessions.

Characteristics and Tools for Analysis

To conduct visual analysis, the primary researcher used Microsoft Excel. The researcher formatively analyzed pairwise comparisons with differentiation between the two interventions and compared each intervention to the control set. Only within participant data was assessed using horizontal analysis. However, the researcher also considered patterns of responding across participants within the same dyad to evaluate the effectiveness of each intervention within a small group.

Elsa

Elsa's skill acquisition data are shown in Figure 1.

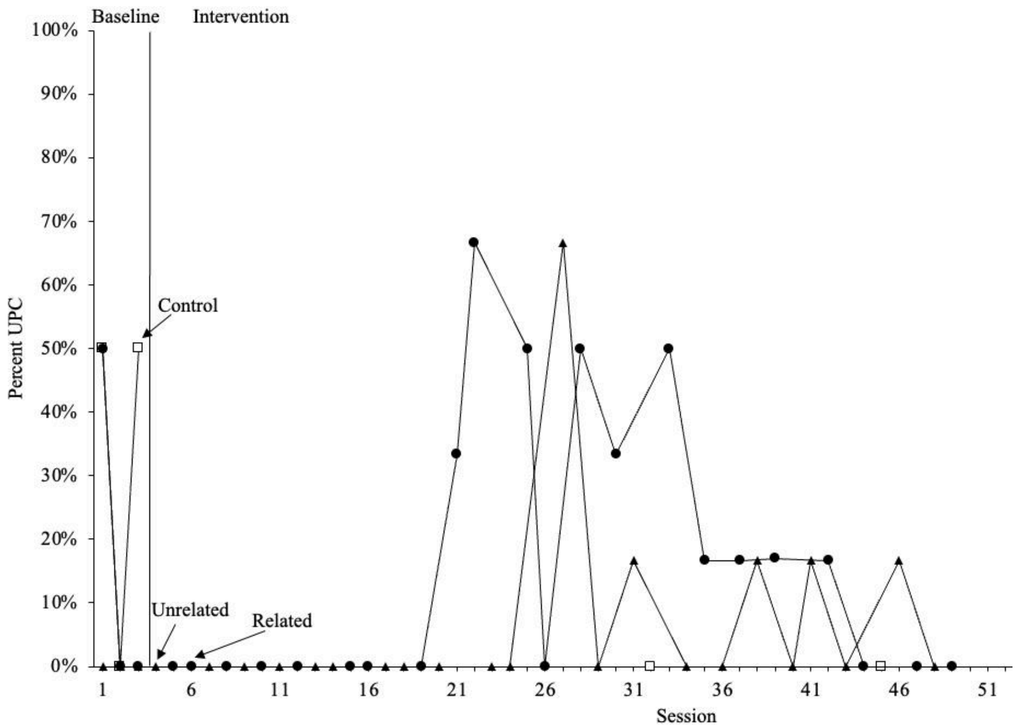

Figure 1. Skill acquisition data for Elsa using embedded instruction.

During baseline, Elsa demonstrated an inconsistent pattern of responding with one target initially assigned to the related condition, /t/, and one target initially assigned to the control condition, /b/. She provided unprompted correct responses for /b/ in both presentations of this letter sound during the first and third baseline session, and for both presentations of /t/ in the first baseline session. For this reason, /b/ and /t/ were assigned to the control condition for all sessions during the comparison phase. Unprompted correct responding was at 0% for the remaining control and related targets, as well as for the unrelated targets.

During the related sessions of the comparison phase, Elsa progressed to a 1 s delay following 8 related sessions. In the related condition, Elsa demonstrated an increasing trend of unprompted correct responding immediately following the onset of 1 s delay sessions. However, this trend only lasted for two sessions, reaching a maximum of 66% unprompted correct responding before decreasing. The decreasing trend began with the thirteenth session, which is the same session in which the implementer introduced stamps as reinforcers for both prompted and unprompted correct responding. Elsa produced 0% unprompted correct responses in one session before increasing once again to 50%, back to 33%, and up to 50% once more before decreasing to 17% in the sixteenth related session, and 0% by the twentieth related session. Unprompted correct responding in this condition remained at 0% for subsequent massed trials sessions.

In the unrelated condition, Elsa progressed to a 1 s delay following 11 unrelated sessions. In the first session at a 1 s delay, she provided 66% unprompted correct responses. Following this peak in her level of independent responding, Elsa's unprompted correct responses immediately decreased and stabilized at a consistent level between 0% and 17%. This pattern continued during massed trials sessions. Elsa did not progress to a 2 s delay in either intervention condition, and her targets were considered not mastered.

In the first massed trial session at a 0 s delay with the novel implementer, Elsa provided a prompted correct response for 100% of trials (Figure 9).

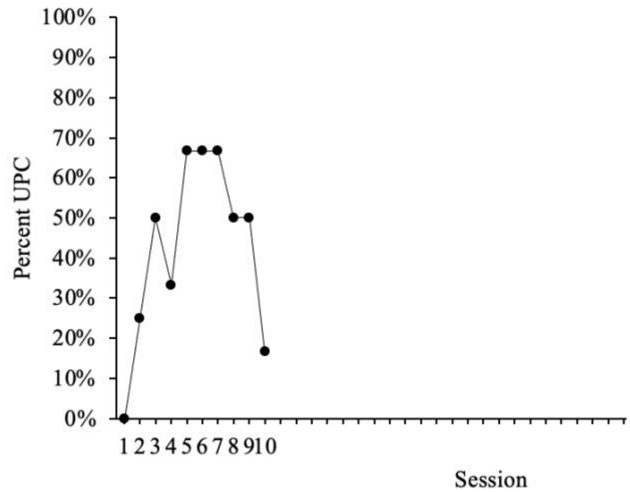

Figure 9. Skill acquisition data for Elsa using massed trials.

The implementer increased the wait interval to 1 s after the first session. For the next three 1 s delay sessions, Elsa produced 25%, 50%, and 33% unprompted correct responses, respectively. In the following three sessions, Elsa provided a correct response without a prompt for 67% of trials, followed by two sessions at 50% unprompted correct responding and one session at 17%. Massed trial sessions in this format ceased following Elsa's transition to a new school.

Anna

As shown in Figure 2, during baseline, Anna provided unprompted correct responses for one trial of a target assigned to each condition (control, related, and unrelated). However, these instances of correct responding were due to repeated guesses of the same response, resulting in at least one correct response. It was determined that she had not already acquired any of the targets before beginning intervention.

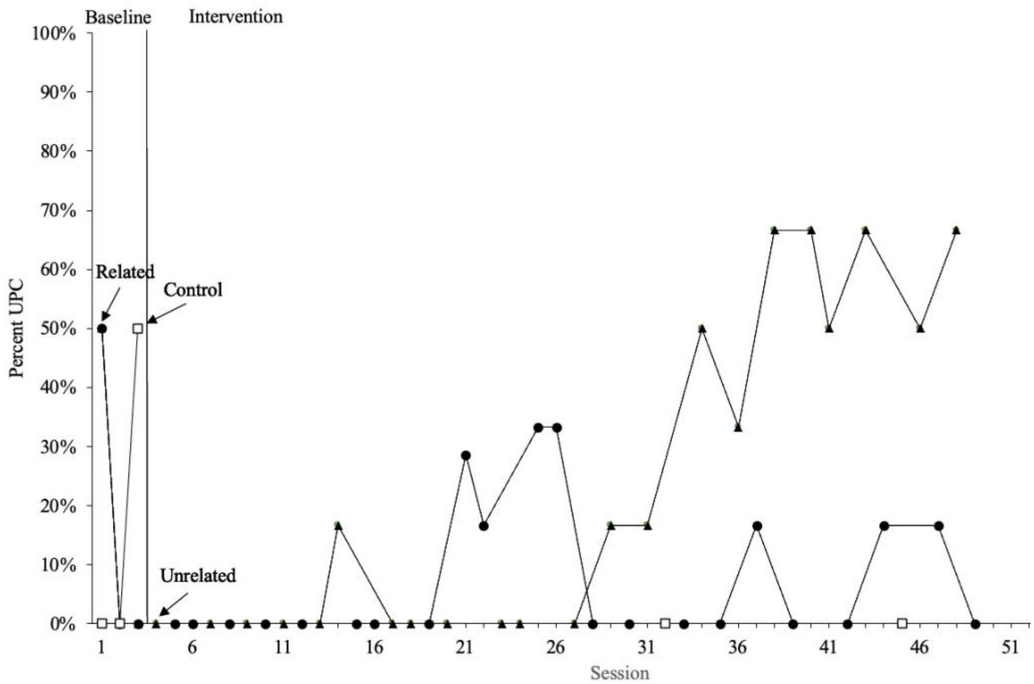

Figure 2. Skill acquisition data for Anna using embedded instruction.

During the related sessions of the comparison phase, Anna progressed to a 1 s delay following three related sessions. Unprompted correct responding remained at 0% for the first five related sessions at a 1 s delay. By the ninth related session, Anna progressed to a 2 s delay. At this time, Anna's unprompted correct responding demonstrated a slight increasing trend, with four consecutive related sessions at 17% to 33% unprompted correct responding. However, after the twelfth related session, the implementer began providing stamps to reinforce prompted and unprompted correct responding. At this time, Anna's unprompted correct responding dropped to 0%, and remained between 0% and 17% for the duration of the study, including massed trials sessions.

During the unrelated sessions of the comparison phase, Anna progressed to a 1 s delay after five unrelated sessions. She showed a brief increase in unprompted correct responding during her first session at a 1 s delay but returned to 0% in the next session. Her level of unprompted correct responding remained at 0% until the thirteenth unrelated session. At this time, the implementer introduced stamps as a reinforcer contingent on prompted or unprompted correct responding. Beginning with the thirteenth session, Anna's unprompted correct responding established a consistent increasing trend through the final session, reaching maximum levels of 66% unprompted correct responses. Anna did not progress to a 2 s delay in this condition, and her targets in both conditions were considered not mastered.

In the first massed trial session at a 0 s delay with the novel implementer, Anna provided a prompted correct response for 100% of trials (Figure 10).

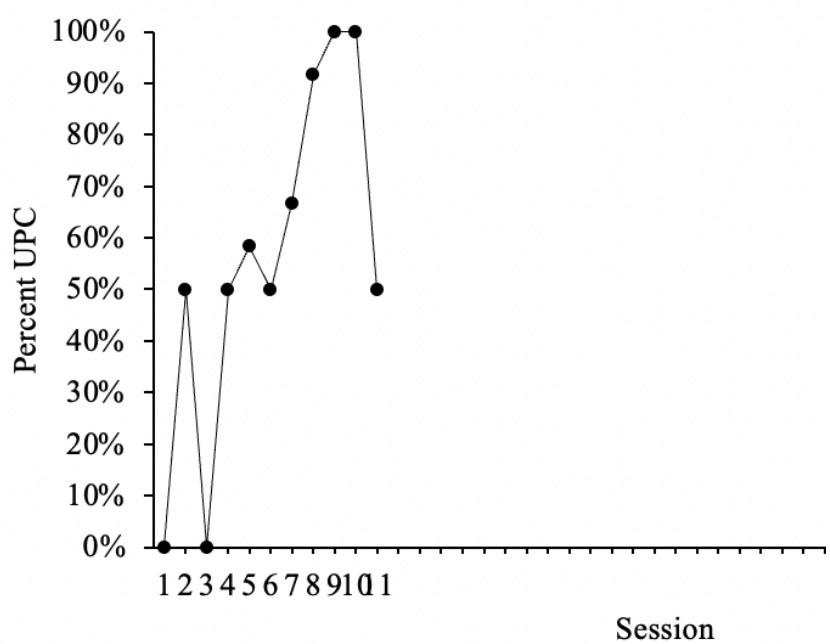

Figure 10. Skill acquisition data for Anna using massed trials.

When the wait interval increased to 1 s, Anna produced 50%, 0%, 50%, 58%, 50%, and 67% unprompted correct responses, respectively. At this time, the delay interval was increased to 2 s, and Anna provided correct responses without a prompt in 92%, 100%, 100%, and 50% of trials, respectively. Massed trial sessions in this format continued until she transferred schools and data collection was terminated.

Mickey

Mickey's skill acquisition data are shown in Figure 3.

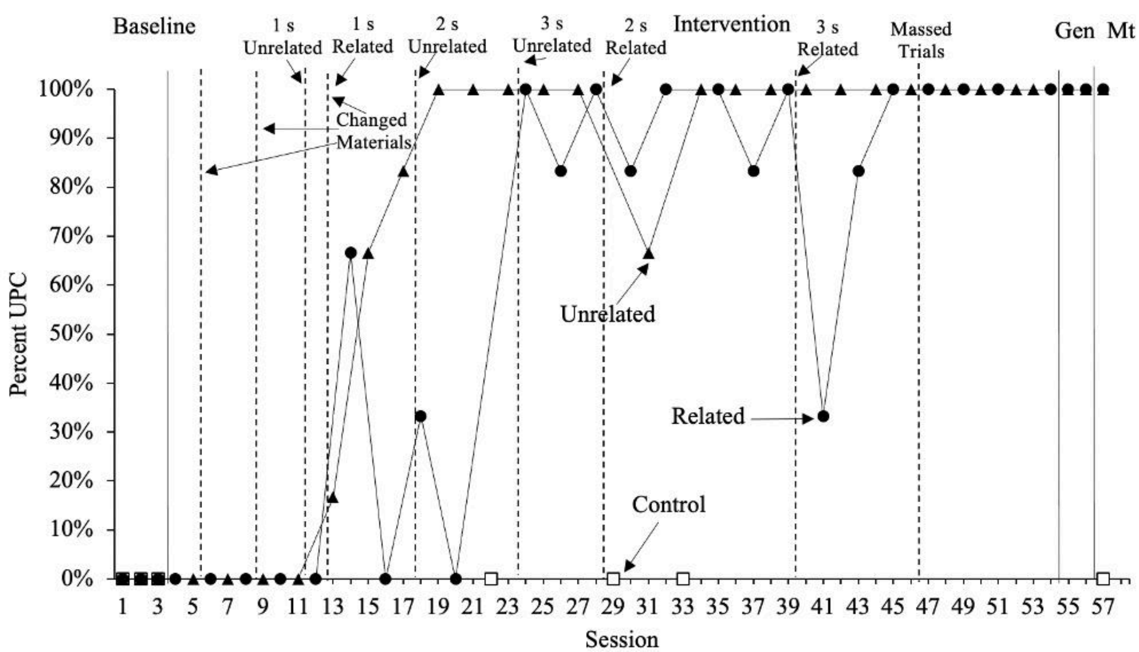

Figure 3. Skill acquisition data for Mickey.

Mickey progressed to a 1 s delay after four unrelated sessions, and five related sessions. Mickey's unprompted correct responding showed an increasing trend in both conditions immediately following his progression to a 1 s wait interval. In the unrelated condition, his level of unprompted correct responding steadily increased from 17% to 83% in the first three sessions at a 1 s delay, leading to his progression to a 2 s delay. In his first 2 s delay session, Mickey responded correctly without a prompt for 100% of trials for three consecutive sessions, reaching mastery criterion in a total of 10 unrelated sessions. He progressed to a 3 s delay and continued to provide 100% unprompted correct responses for 93% of the additional sessions that followed attainment of mastery criteria. He provided correct responses without a prompt for 100% of trials in all generalization and maintenance sessions.

In the related condition, Mickey's unprompted correct responding immediately increased to 66% following the introduction of a 1 s delay. His level of responding was more variable for three subsequent sessions, returning to 0%, 33%, and 0%, respectively. He reached 100% unprompted correct responding in the tenth related session. Unprompted correct responding alternated between 83% and 100% for the next four related sessions. He then maintained 100% unprompted correct responding for two consecutive related sessions before decreasing to 83% once more and returning to 100%. Mickey responded correctly without a prompt for 33% of trials in the following session, which was likely due to decreased interest in the play materials, as he expressed vocally that he wanted to play with something different. Unprompted correct responding then returned to 83% and 100% in the next two sessions before beginning massed trials. Mickey responded correctly without a prompt for the remaining four massed trial sessions, reaching mastery criterion. He answered correctly without a prompt for 100% of trials in all generalization and maintenance sessions. When comparing intervention and control conditions, a functional relation existed such that PTD was related to increases in acquisition regardless of whether it was related or unrelated. However, when comparing the intervention conditions to each other, a functional relation existed such that the presentation of unrelated targets was associated with a steeper slope of acquisition and faster mastery than the use of related targets.

Minnie

Figure 4 displays Minnie's skill acquisition data.

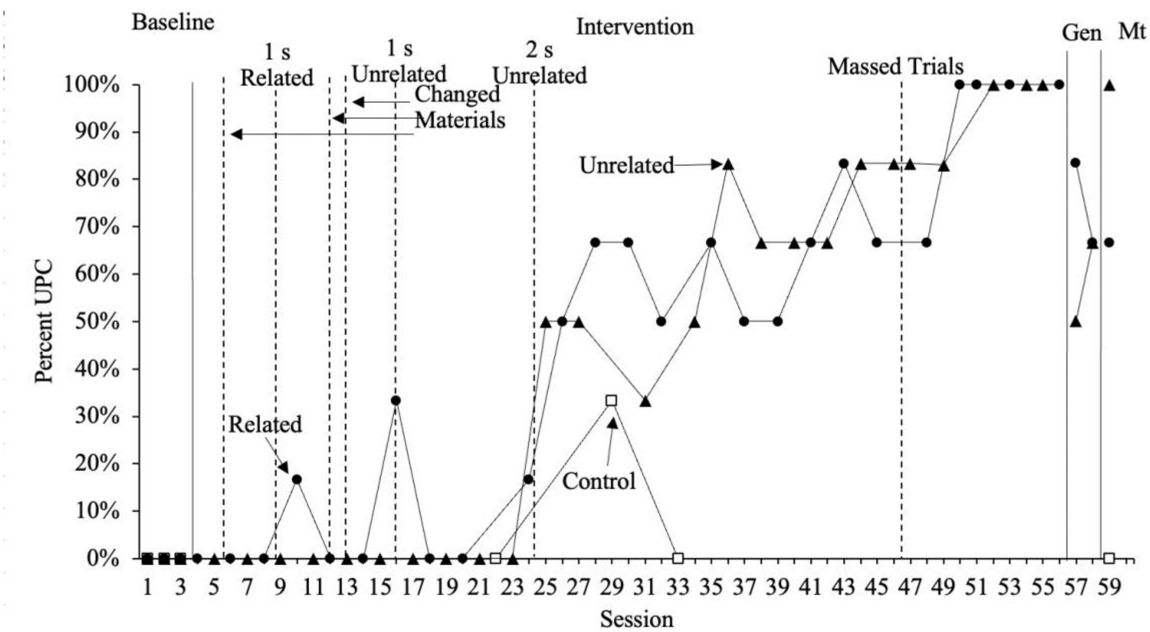

Figure 4. Skill acquisition data for Minnie.

In the related condition, Minnie progressed to a 1 s delay interval after three sessions in which she produced 100% prompted correct responses. She did not progress to a 2 s delay and showed inconsistent and variable levels of unprompted correct responding ranging from 0% to 17% in the next six related sessions. After the ninth related session, Minnie's level of responding began to show an increasing trend from 17% to 66% over the course of 3 consecutive related sessions. Unprompted correct responding then alternated between 66% and 50% for the following three related sessions, before establishing an increasing trend from 50% to 83% over the course of three consecutive related sessions. Minnie's level of unprompted correct responding then decreased to 50% once more before beginning massed trial sessions. Minnie provided a correct response without a prompt for 50% of trials in the first massed trial session, followed by three consecutive sessions at 100%, reaching mastery criterion. In the two generalization sessions that followed, Minnie provided a correct response without a prompt for 83% and 50% of related targets. One maintenance session occurred after 4 weeks without instruction, and Minnie answered correctly without a prompt in 67% of trials for related targets.

In the unrelated condition, Minnie progressed to a 1 s delay after six sessions. Unprompted correct responding remained at a consistent level of 0% for the following four sessions after beginning a 1 s delay. However, after progressing to a 2s delay for the eleventh unrelated session, Minnie's level of unprompted correct responding suddenly increased to 50%. Unprompted correct responding then alternated between 50% and 33% for the following 3 unrelated sessions, then increased to 83% before falling back to 66% for three consecutive unrelated sessions. Minnie's level of unprompted correct responding returned to 83% in the next two unrelated sessions before beginning massed trials. Minnie's level of unprompted correct responding for the unrelated targets remained at 83% in the first two massed trial sessions, before completing three consecutive sessions at 100%, reaching mastery criterion. In the two generalization sessions that followed, Minnie provided a correct response without a prompt for 50% and 67% of unrelated targets. During maintenance, Minnie answered correctly without a prompt for

100% of trials for unrelated targets. When comparing intervention and control conditions, a functional relation existed such that PTD was related to increases in acquisition regardless of whether it was related or unrelated

Engagement

Elsa

Elsa demonstrated similar levels of engagement across all intervention conditions (Figure 5).

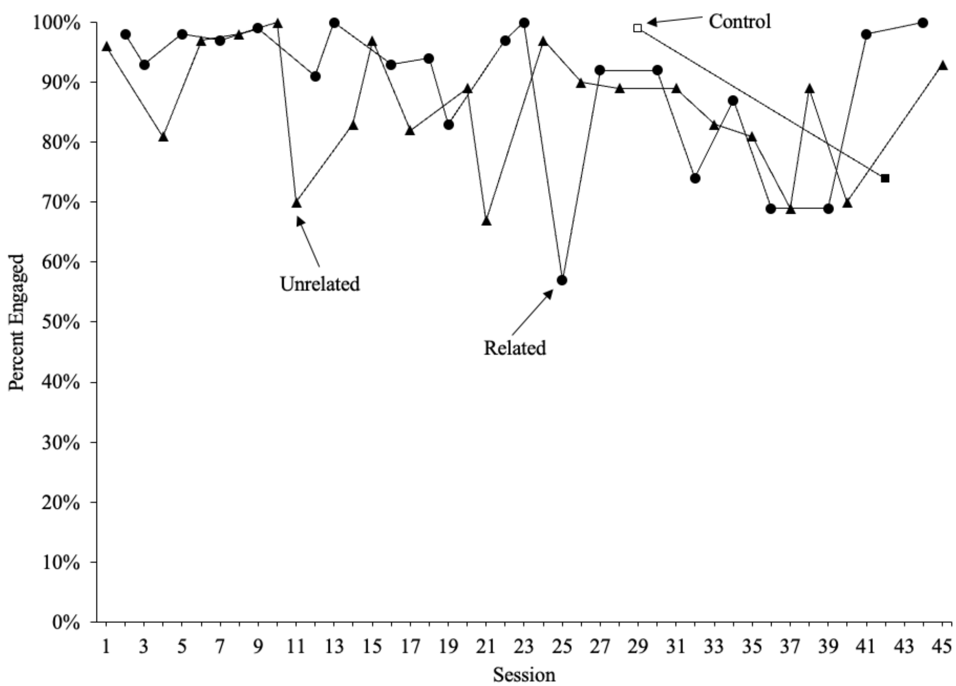

Figure 5. Engagement data for Elsa.

In control sessions, during which the play session followed massed trials instruction, Elsa's engagement ranged was an average of 86.5% (74-99%). In the related condition, Elsa maintained an average level of engagement of 88% (50-100%). Her level of engagement remained above 80% for all but five sessions. In the unrelated condition, Elsa maintained an average level of engagement of 86% (67-100%). There were no observable differences in engagement between sessions of varying conditions. It should be noted that sessions with lower levels of engagement were typically due to arguing between participants which often lead to tantrums or interacting with peers and adults walking by the room in which sessions took place. Specifically, there were many instances in which the participants' mother walked by the room, and both Anna and Elsa disengaged from the activity and instruction for a brief period of time to interact with her. While there were no distinct differences in level of engagement between the two interventions, when comparing the intervention condition to the control condition, the data show that the provision of trials during intervention corresponded with a slight decrease in overall level of engagement.

Anna

As shown in Figure 6, Anna demonstrated similar levels of engagement across all intervention conditions. In control sessions, during which play sessions were conducted without embedding trials, Anna's play engagement was an average of 95% (93-97%). In related sessions, Anna showed an average level of engagement of 88% (64-100%). In unrelated sessions, Anna had an average level of engagement of 87% (44-100%). There were no notable differences between her engagement in unrelated, related, and control sessions. Similarly to Elsa, Anna's lower levels of engagement were frequently due to arguments with Elsa or interactions with their mother and peers walking by the door.

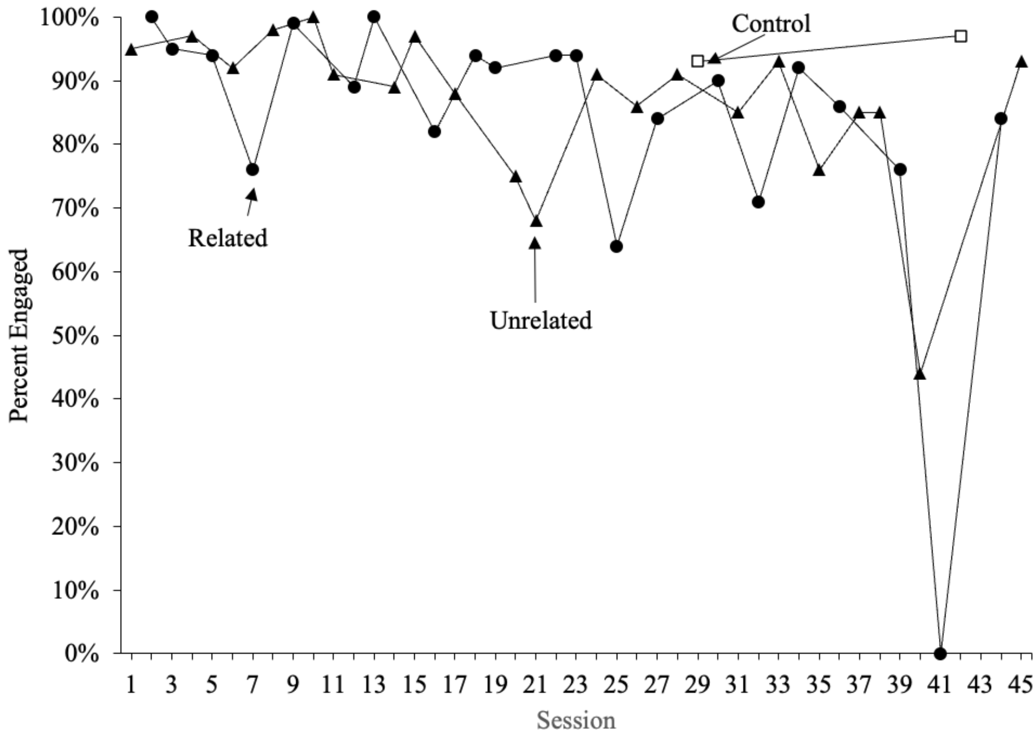

Figure 6. Engagement data for Anna.

Based on comparisons between the control condition and intervention conditions, there is no evidence that the embedding of trials into play disrupted Anna's play engagement.

Mickey

Mickey's engagement data are shown in Figure 7. During control sessions, Mickey showed an average engagement level of 100%. Throughout the first nine sessions of the comparison condition, Mickey showed highly variable levels of engagement in both the unrelated and related condition. Engagement during these sessions ranged from 49% to 97%. After four unrelated sessions (two of which involved wooden blocks), and five related sessions (three of which involved Play-doh), the implementer added cars to the wooden blocks, and pretend food to the Play-doh sets. Following these changes in

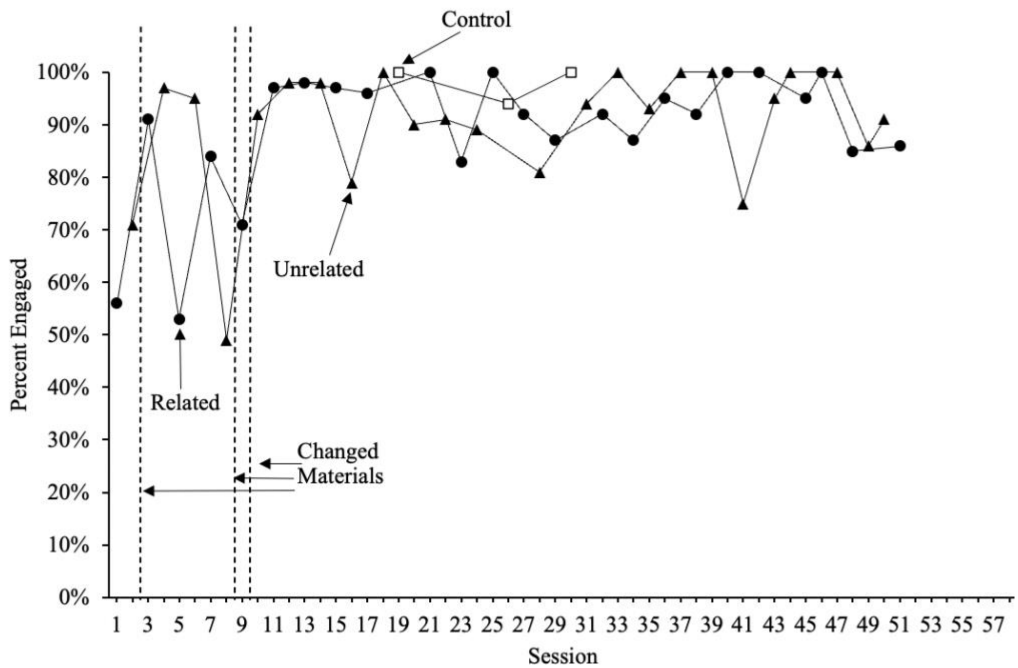

Figure 7. Engagement data for Mickey.

play materials, Mickey showed higher and more consistent levels of engagement in both conditions. In the related condition (after provision of new materials), Mickey had an average level of engagement of 94% (83-100%). In the unrelated condition (after the provision of new materials), Mickey had an average level of engagement of 92% (75-100%). Sessions in which Mickey's level of engagement dropped below 80% typically occurred during a reinforcement period provided by a classroom-based contingency for successfully using the bathroom. His reinforcement period consisted of 3 minutes of access to a preferred car ramp toy. In these sessions, Mickey was considered to be not engaged while interacting with the car ramp, as this was not included in play materials provided by the implementer. There were no significant differences in his level of engagement between either intervention condition. However, when comparing the intervention condition to the control condition, the data demonstrate a slight decrease in overall engagement with the presentation of trials during intervention.

Minnie

Figure 8 displays Minnie's engagement data across intervention conditions.

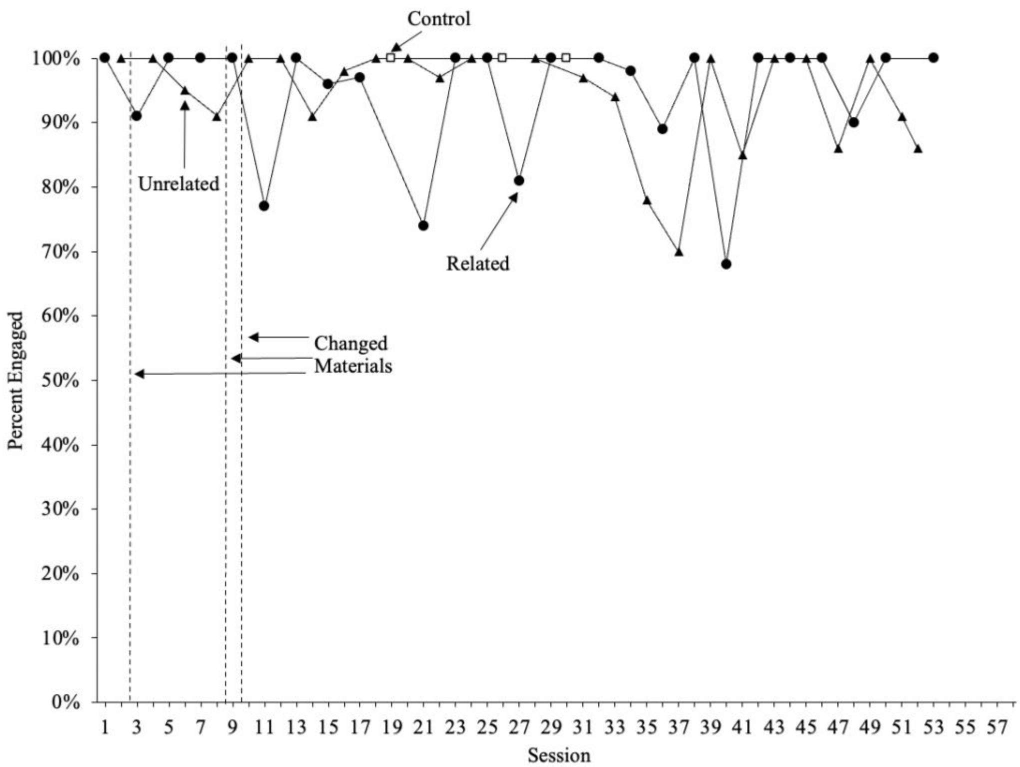

Figure 8. Engagement data for Minnie.

Minnie displayed a consistent high level of engagement throughout all sessions of the study. During control sessions, she showed an average engagement level of 100%. During related sessions, she maintained an average engagement level of 94% (68-100%), and during unrelated sessions, she maintained an average level of 95% (70-100%). There were only three related sessions and two unrelated sessions in which Minnie's engagement dropped below 80%. During these sessions, Minnie was typically engaging with other materials in the room or initiating conversations with the implementer and peer that were not related to the materials or instructional content. Therefore, while the data are not demonstrative of a significant difference in engagement levels between related and unrelated sessions, there was a slight overall decrease in engagement that corresponded with the provision of trials during intervention.

Child Preference

Child preference for related and unrelated embedded instruction was assessed on three occasions for Elsa and Anna, and two occasions for Mickey and Minnie. The implementer did so by asking, “Do you like to answer questions on red carpet days, or green carpet days?” Elsa reported a preference for related instruction on all three occasions, while Anna reported a preference for unrelated instruction. In the second dyad, Mickey reported a preference for unrelated instruction, while Minnie preferred related instruction. It should be noted that in the second dyad, Minnie explained her preference by expressing that she liked the play materials better in the related condition. Her preference therefore may not be due to the method for instruction. Additionally, it was not clear if Mickey associated the carpet color with the method of instruction when he expressed his preference. His teacher reported that green is a preferred color, and therefore may have influenced his expressed preference for the unrelated condition which corresponds with the green carpet.

Reliability

Inter-observer agreement data were collected for 45.7% of sessions across conditions for Elsa and Anna, and 40.0% of sessions across conditions for Mickey and Minnie. A secondary observer recorded data independently from the primary observer using behavioral video coding through Digital ProCoderDV. IOA was calculated using point-by-point analysis for all variables pertaining to skill acquisition (assessed per trial) and engagement (assessed per interval using 10 s MTS). IOA was calculated individually for each participant. Skill acquisition data were in agreement if both the primary and secondary observer indicated the same response type from the participant (e.g., PC, UPC, PE, UPE). Engagement data were in agreement if both the primary and secondary observer indicated that the participant was engaged in play, engaged in instruction, not engaged, or offscreen at the end of each 10 s interval. IOA data were monitored formatively and discrepancy discussions were held for every session with secondary data. Across all conditions, skill acquisition IOA was an average of 88.7% (33.3-100%) for Elsa, 92.6% (66.6-100%) for Anna, 96.1% (66.6-100%) for Mickey, and 93.1% (66.6-100%) for Minnie. Across all conditions, engagement IOA was an average of 92.5% (61,1-100%) for Elsa, 88.3% (44.4-96.6%) for Anna, 84.3% (64.5-96.8%) for Mickey, and 91.6% (71-100%) for Minnie.

Procedural Fidelity

Procedural fidelity was assessed for 45.7% of sessions across conditions for Elsa and Anna, and 40.0% of sessions across conditions for Mickey and Minnie to ensure that the implementer adhered to procedures and differentiated between conditions. The same secondary observer that collected IOA data performed procedural fidelity coding for trial-based and play-based behaviors. During instructional trials, the secondary observer indicated the frequency at which the implementer presented the correct materials and contingency review, ensured participant attending, provided the appropriate wait interval and controlling prompt, provided the correct feedback following a participant response, provided the correct number of trials and separated trials by 60 s +/- 30 s, and maintained the relatedness or unrelatedness of the target by strategically embedding trials. To assess play-based behaviors, the secondary observer indicated whether the implementer was engaged in an appropriate behavior at the end of each 10 s interval during all non-trial segments of each session. These included engaging in parallel play, commenting on participant play, or responding to a comment. For sessions with Elsa and Anna, trial-based procedural fidelity was an average of 96.1% (90.7-100%) and play-based procedural fidelity was an average of 88.2% (79.0-96.8%). For sessions with Mickey and Minnie, trial-based procedural fidelity was an average of 99.7% (97.4-100%) and play-based procedural fidelity was an average of 96.9% (90.0-100%).

CHAPTER IV

DISCUSSION

Summary of Findings

This study provides insight into some of the considerations that should be made prior to selecting and implementing a teaching procedure for small group instruction. Firstly, the results of this study support previous findings that embedded instruction is effective for some children, but not all children (Ledford et al., 2017). Elsa and Anna did not reach mastery after 22 sessions of each condition. Mickey reached mastery after 10 sessions in the unrelated condition. While he did not attain mastery criterion after 19 related sessions, he inconsistently responded correctly without a prompt for 100% of trials beginning in the tenth related session, and reached mastery within the first two massed trial sessions. Mickey maintained and generalized all targets across conditions. Minnie did not reach mastery after 19 sessions in either condition but showed significant improvements in her skill acquisition, and reached mastery criterion in both conditions within the first five massed trial sessions. Minnie maintained 67% of unrelated targets and 100% of related targets, and generalized an average of 58% of unrelated targets and 75% of related targets.

Additionally, it should be acknowledged that providing embedded direct instruction in small groups can become increasingly complicated if children are not compatible play partners. Incompatibility can be determined by the likelihood that children will engage in arguing or tantruming that may disrupt instruction as well as play engagement. Furthermore, it should be noted that for all participants except Anna, the inclusion of trials affected engagement in such a way that it slightly decreased engagement across sessions of both intervention conditions.

Child preference data showed that Elsa and Minnie preferred related instruction, while Anna and Mickey preferred unrelated. Anna verbally expressed that she did not like to use play materials for instructional trials, demonstrating a consistent preference for a distinction between play and instructional materials. Mickey and Minnie, however, both cited a preference for the play materials corresponding with each condition as the reasoning for their preference of the condition. Therefore, it is possible that these report measures for preference were not valid for Mickey and Minnie, instead performing as a measure of preference for *toys* rather than *instruction type*. One potential hypothesis for Mickey's preference and efficiency in acquiring unrelated targets is that he may have perceived more salience in the trials using commercial cards rather than play materials. Mickey's data also supports Botts and colleagues (2014) who found that embedded direct instruction with unrelated target stimuli was more effective and efficient in the acquisition of phonological awareness skills in comparison to an activity-based intervention which closely reflected related embedded instruction.

It is then critical to consider the recommendations of the Division of Early Childhood (2014), which include that "Practitioners, with the family, identify each child's strengths, preferences, and interests to engage the child in active learning" and "Practitioners embed instruction within and across routines, activities, and environments to provide contextually relevant learning opportunities". The examples of these practitioner behaviors provided by the DEC include that "Team members identify logical and appropriate opportunities for the child to practice and learn targeted skills during routine, planned, and child-initiated activities that occur in the classroom". While naturalistic opportunities embedded into routines is documented as a recommended practice, it is possible that for some students, child-initiated activities are not an appropriate or preferred time to embed instruction. Some students may find embedding trials during child-initiated activities to be aversive or disruptive, hindering both learning and engagement.

Perhaps a more useful qualitative observation is that children who demonstrated more intense engagement in play expressed a stronger disapproval for disruptions in play. This is relevant for both typically developing and autistic children, as many young children demonstrate high levels of play engagement, and some autistic children demonstrate fixations on preferred objects. Data from this study suggest that during activities that lend itself to particularly high levels of engagement or fixation, disrupting play with the inclusion of trials may not be beneficial for the child's learning.

Limitations

There are several limitations to this study that are of notable importance. Firstly, low interobserver agreement was collected on multiple occasions during which participants were not on screen. The implementer attempted to narrate the child's behaviors during sessions in which children removed themselves from the area visible to the camera. However, there were sessions in which the secondary observer had difficulty hearing and could not conclude what the child was doing while off screen. Secondly, the implementer had low procedural fidelity for play-based behaviors during multiple sessions with Anna and Elsa. This was largely due to time spent de-escalating participants from arguing or tantruming as well as redirecting them from engaging with other materials in the room. Thirdly, there was some inconsistency in the use of play materials and targets, as they were switched after baseline for Elsa and after a few intervention sessions for Mickey due to inconsistent baseline responding (Elsa) and a lack of play engagement (Mickey). Additionally, the implementer added a component of explicit instruction prior to the start of sessions for only one participant (Anna). This was introduced following a series of sessions in which Anna did not visually attend to the stimuli presented during trials. Finally, the researcher did not use an adequate method for measuring child preference. Perhaps in future studies, it would be worthwhile to conduct a preference assessment before beginning intervention sessions to ensure that materials are moderately preferred across conditions. However, this would be difficult to maintain when multiple participants with differing play preferences are being served simultaneously.

Implications

The findings of this study have implications for the use of embedded instruction in both research and in practice. This study suggests it is possible to embed disparate trials for multiple children into a single play activity. However, the results of this study also showed that some children do not prefer embedded instruction to massed trial instruction. This reflects the findings of Ledford and colleagues (2017) and Heal and Hanley (2007). Ledford and colleagues demonstrated that most children preferred the type of instruction that resulted in the most efficient learning. The instruction that resulted in the most efficient learning, and consequently was the preferred type of instruction, varied across participants (Ledford et al., 2017). Heal and Hanley (2007) found that children preferred high quality reinforcers following teacher-led instruction to high quality materials in embedded instruction. This is critical for teachers and practitioners as instruction cannot be delivered effectively in the same way for all students. Additionally, it is important to utilize a pre-determined criteria for adequate progress and closely monitor student progress in order to identify the point at which a different instructional approach is necessary. Students who participate in small group embedded instruction for extended periods of time without making adequate progress are not necessarily incapable of acquiring the material. Rather, they may need the material to be presented differently. This applies for both typically developing students and students with support needs in classroom and research settings.

It is also important to note that small group embedded instruction becomes increasingly challenging if students are not compatible play partners. This has implications for small group arrangements in classroom and research settings, as students may not learn effectively if they are learning alongside students with incompatible play behaviors. Small groups in classrooms should be selected intentionally to optimize peer compatibility, and participants in research settings should be recruited with existing evidence that the participants will play well together.

An anecdotal finding of this study is that rather large differences existed between conditions in terms of requirements for the implementer. Embedding related targets in play required significantly more pre-session preparation and in-situ decision-making on the part of the implementer. This may be important for practitioners to consider if they have a limited amount of time during which to embed instruction, or if they are completing multiple tasks simultaneously, such as collecting data and supervising children in other activities while providing instruction. This also has implications when translating research into practice. Researchers must acknowledge the mental effort required to provide this type of instruction and consider its impact on the feasibility of a related embedded intervention.

Finally, when selecting a teaching strategy, researchers and practitioners should consider the salience of trials and child preferences surrounding disruptions to play. Some students may learn effectively if the embedded instruction involves play materials (e.g., making a letter with play doh), while others may require instructional targets to be presented through a different modality. It is possible that instructional targets lose salience and become more abstract when presented in the context of play and while using only play materials. Furthermore, some students may not object to interruptions in play to respond to instructional trials. Others, however, may find disruptions to play aversive. This is critical as these preferences may not only influence a student's likelihood to attend to trials and acquire instructional targets, but consistent pairing of an aversive disruption and academic target may have a countertherapeutic effect and result in a more generalized aversion to instruction.

Suggestions for Future Research and Practice

Future researchers and practitioners interested in using small group embedded instruction should consider several factors before recruiting participants, selecting targets, or beginning instruction. These factors include child competence in play, baseline levels of play engagement and fixation, and the likelihood that the children will be compatible play partners. Further research is needed for more conclusive evidence of the qualities that determine whether embedded instruction will be effective or efficient, and the qualities that may influence whether related or unrelated stimuli will be more preferred. More research may also provide insight into whether it is the relatedness of targets to an engaging activity or the salience of instructional materials that is a stronger predictor of skill acquisition. Future findings can be used to inform and substantiate DEC recommendations for embedded instruction. It may be helpful to have more guidance and clearer suggestions for practitioners attempting to embed targets for multiple students across routines and activities. It is important to determine if such recommendations are feasible, efficient, preferred, and effective.

Additional components that may be added to future research might include a measure of observational learning and an improved measure of child preference. To do so, researchers and practitioners may collect baseline data for a participant's responses to a peer's targets and include the peer's targets in intermittent control sessions to gauge learning. An improved measure of child preference would require distinctive correlated stimuli for each condition, as well as explaining the correlated stimuli and phrasing questions pertaining to student preference in a way that is comprehensible to the participants. This might involve additional preference assessments to ensure moderately preferred play materials will be used in both conditions, as the participants in the current study misconstrued the correlated stimuli as corresponding with play materials rather than the relatedness of instruction. These added measures may provide useful information about observational learning and clarifying child preferences in a small group embedded instructional context.

Conclusions

The results of this study support the conclusion that embedded instruction works for some, but not all, students. Additionally, quantitative and anecdotal data suggest that the effectiveness of embedded instruction may be related to peer compatibility in play settings and preferences pertaining to disruptions in play as well as the salience of embedded trials. This study demonstrates that functional relations exist such that PTD is related to increases in skill acquisition across three participants, and for one participant, acquisition of unrelated targets was more efficient than acquisition of related targets. Finally, while child preferences pertaining to relatedness of targets in embedded instruction exist, more accurate methods of collecting child preference data are needed. This study provides evidence that embedded small group instruction is feasible, but there are several critical factors to consider that may impact its effectiveness.

REFERENCES

Albarran, S.A. & Sandbank, M.P. (2019). Teaching non-target information to children with disabilities: An examination of instructive feedback literature. *Journal of Behavioral Education, 28*(1), 107-140. Doi: 10.1007/s10864-018-9301-3

Ault, M.J., Gast, D.L., & Wolery, M. (1988b). Comparison of progressive and constant time delay procedures in teaching community-sign word reading. *American Journal of Mental Retardation, 93*(1), 44-56.Barnett, J.H. (2018). Three evidence-based strategies that support social skills and play among young children with autism spectrum disorders. *Early Childhood Education Journal, 46*, 665-672. Doi: 10.1007/s10643-018-0911-0

Botts, D.C., Losardo, A.S., Tillery, C.Y., & Werts, M.G. (2014). A comparison of activity-based intervention and embedded direct instruction when teaching emergent literacy skills. *The Journal of Special Education, 48*(2), 120-134. Doi: 10.1177/0022466912449652

Calzi, M.N. (2020). Progressive time delay to teach high school students with intellectual disability to initiate manding Siri for unknown information. *Theses and Dissertations- Early Childhood, Special Education, and Counselor Education, 82.* doi: 10.13023/etd.2020.076

Collins, B.C., Lo, Y., Park, G., & Haughney, K. (2018). Response prompting as an ABA-based instructional approach for teaching students with disabilities. *TEACHING Exceptional Children, 50*(6), 343-355. doi: 10.1177/0040059918774920

Daugherty, S., Grisham-Brown, J.G., & Hemmeter, M.L. (2001). The effects of embedded skill instruction on the acquisition of target and nontarget skills in preschoolers with developmental delays. *Topics in Early Childhood Special Education, 21*(4), 213-221. Doi: 10.1177/027112140102100402

Division of Early Childhood for the Council of Exceptional Education. (2014). *DEC Recommended practices*. https://divisionearlychildhood.egnyte.com/dl/7urLPWCt5U

Doyle, P.M., Wolery, M., Gast, D.L., Ault, M.J., Wiley, K. (1990). Comparison of time delay and system of least prompts in teaching preschoolers with developmental delays. *Research in Developmental Disabilities, 11*(1), 1-22. Doi: 10.1016/0891-4222(90)90002-

Francis, R., Winchester, C., Barton, E.E., Ledford, J.R., & Velez, M. (2020). Using progressive time delay to increase levels of peer imitation during play with preschoolers with disabilities. *American Journal of Intellectual and Developmental Disabilities, 125*(3), 186-199. doi: 10.1352/1944-7558-125.3.186

Gast, D.L., & Spriggs, A.D. (2019). Visual analysis of graphic data. In D.L. Gast & J.R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp.176-210). Routledge.

Grisham-Brown, J., Pretti-Frontczak, K., Hemmeter, M.L., & Ridgley, R. (2002). Teaching IEP goals and objectives. *Young Exceptional Children, 6*(1), 18-27.

Heal, N.A., & Hanley, G.P. (2007). Evaluating preschool children's preferences for motivational systems during instruction. *Journal of Applied Behavior Analysis, 40*(2), 249-261. doi: 10.1901/jaba.2007.59-05.

Holocombe, A., Wolery, M., Gessler Werts, M., & Hrenkevich, P. (1993). Effects of instructive feedback on future learning. *Journal of Behavioral Education, 3*(3), 259-285. Doi: 10.1007/BF00961555

Horn, E., Lieber, J., Li, S., Sandall, S., & Schwartz, I. (2000). Supporting young children's IEP goals in inclusive settings through embedded learning opportunities. *Topics in Early Childhood Special Education, 20*(4), 208-223. Doi: 10.1177/027112140002000402

Korba, J., Hemmeter, M.L, Golden, A.K., & Nuhring, K. (2021). The effects of progressive time delay to teach social problem-solving to preschoolers. *Journal of Early Intervention, 44*(3), 268-288. doi: 10.1177/10538151211052004

Lane, J.D., Gast, D.L., Shepley, C., & Ledford, J.R. (2015). Including social opportunities during small group instruction of preschool children with social-communication delays. *Journal of Early Intervention, 37*(1), 3-22. Doi: 10.1177/1053815115588828

Ledford, J.R., Lane, J.D., Elam, K.L., & Wolery, M. (2012). Using response-prompting procedures during small-group direct instruction: Outcomes and procedural variations. *American Journal on Intellectual and Developmental Disabilities, 117*(5), 413-434. Doi: 10.1352/1944-7558-117.5.413

Ledford, J.R., & Wolery, M. (2013). Peer modeling of academic and social behaviors during small-group direct instruction. *Council for Exceptional Children, 79*(4), 439-458. Doi: 10.1177/001440291307900404

Ledford, J.R., & Wehby, J.H. (2014). Teaching children with autism in small groups with students who are at-risk for academic problems: Effects on academic and social behaviors. *Journal of Autism and Developmental Disorders, 45*, 1624-1635. doi: 10.1007/s/10803-014-2317-1

Ledford, J.R., Chazin, K.T., Harbin, E.R., & Ward, S.E. (2017). Massed trials versus trials embedded into game play: Child outcomes and preference. *Topics in Early Childhood Special Education, 37*(2). 107-120. doi: 10.1177/0271121417704703

Ledford, J.R., & Gast, D.L. (2018*). Single case research methodology: Applications in special education and behavioral sciences*. New York, NY: Routledge.

Ledford, J.R., Lane, J.D., & Barton, E.E. (2019). *Methods for teaching in early education*. Mahwah, NJ: Routledge.

McDonnell, J., Johnson, J.W., Polychronis, S., Riesen, T., Jameson, M., & Kercher, K. (2006). Comparison of one-to-one embedded instruction in general education classes with small group instruction in special education classes. *Education and Training in Developmental Disabilities, 41*(2), 125-138.

Olszewski, A., Soto, X., & Goldstein, H. (2017). Modeling alphabet skills as instructive feedback within a phonological awareness intervention. *American Journal of Speech-Language Pathology, 26*(3), 769-790. Doi: 10.1044/2017_AJSLP-16-0042

Shepley, C., Grisham-Brown, J. (2019). Multi-tiered systems of support for preschool-aged children: A review and meta-analysis. *Early Childhood Research Quarterly, 47*. doi: 10.1016/j.ecresq.2019.01.004

Sweeney, E., Barton, E.E., & Ledford, J.R. (2018). Using progressive time delay to increase levels of peer imitation during sculpting play. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s/10803-018-3638-2

Tapp, J. (2003). *ProcoderDV.* Nashville, TN: Vanderbilt Kennedy Center.

Urlacher, S., Wolery, M., & Ledford, J.R. (2016). Peer modeling of commenting during small group direct instruction for academic behaviors. *Journal of Early Intervention, 38*(1), 24-40. doi: 10.1177/1053815116636645

Venn, M.L., Wolery, M., Gessler Werts, M., Morris, A., DeCesare, L.D., Siegesmund Cuffs, M. (1993). Embedding instruction in art activities to teach preschoolers with disabilities to imitate their peers. *Early Childhood Research Quarterly, 8*(3), 277-294. Doi: 10.1016/S0885-2006(05)80068-7

Werts, M.G., Wolery, M., Holocombe-Ligon, A., Vassilaros, M.A., & Billings, S.S. (1992). Efficacy of transition-based teaching with instructive feedback. *Education and Treatment of Children, 15*(4), 320-334.

Werts, M.G., Wolery, M., Holocombe, A., & Frederick, C. (1993). Effects of instructive feedback related and unrelated to the target behaviors. *Exceptionality, 4*(2), 81-95. Doi: 10.1207/s15327035ex0402_2

Werts, M.G., Wolery, M., Holocombe, A., & Gast, D.L. (1995). Instructive feedback: Review of parameters and effects. *Journal of Behavioral Education, 5*(1), 55-75. Doi: 10.1007/BF02110214

Winstead, O., Lane, J.D. Spriggs, A.D., & Allday, R.A. (2019). Providing small group instruction to children with disabilities and same-age peers. *Journal of Early Intervention, 41*(3), 202-219. doi: 10.1177/1053815119832985

Wolery, M., Anthony, L., Caldwell, N.K., Snyder, E.D., & Morgante, J.D. (2002). Embedding and distributing constant time delay in circle time and transitions. *Topics in Early Childhood Special Education, 22*(1), 14-25. Doi: 10.1177/027112140202200102

APPENDIX A

DATA COLLECTION FORMS

Skill Acquisition Data Collection Form

<table>
<tr><td colspan="2">Date:</td><td colspan="2">Condition and Session #:</td></tr>
<tr><td colspan="2">Time Delay:</td><td colspan="2">Participants:</td></tr>
<tr><td colspan="2">Primary Data Collector:</td><td colspan="2">Secondary Data Collector:</td></tr>
<tr><td colspan="2">Participant 1</td><td colspan="2">Participant 2</td></tr>
<tr><td rowspan="2">Item</td><td>Response</td><td rowspan="2">Item</td><td>Response</td></tr>
<tr><td>(UPC, PC, UPE, PE)</td><td>(UPC, PC, UPE, PE)</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="4">UPC: correct response before model is given</td></tr>
<tr><td colspan="4">PC: correct response after model is given</td></tr>
<tr><td colspan="4">UPE: incorrect response before model is given</td></tr>
<tr><td colspan="4">PE: incorrect response after model is given</td></tr>
</table>

Engagement Data Collection Form

Condition:					**Session #:**				
Time Delay:					**Participants:**				
Primary Data Collector:					**Secondary Data Collector:**				
Participant 1:					**Participant 2:**				
Interval (s)	**Engaged (Play)**	**Engaged (Inst)**	**Not Engaged**	**Offscreen**	**Interval (s)**	**Engaged (Play)**	**Engaged (Inst)**	**Not Engaged**	**Offscreen**
0-10s					0-10s				
11-20s					11-20s				
21-30s					21-30s				
31-40s					31-40s				
41-50s					41-50s				
51-60s					51-60s				
61-70s					61-70s				
71-80s					71-80s				
81-90s					81-90s				
91-100s					91-100s				
101-110s					101-110s				
111-120s					111-120s				
121-130s					121-130s				
131-140s					131-140s				
141-150s					141-150s				
151-160s					151-160s				
161-170s					161-170s				
171-180s					171-180s				
181-190s					181-190s				
191-200s					191-200s				
200-210s					200-210s				
211-220s					211-220s				
221-230s					221-230s				
231-240s					231-240s				
241-250s					241-250s				
251-260s					251-260s				
261-270s					261-270s				
271-280s					271-280s				
281-290s					281-290s				
291-300s					291-300s				

Trial-Based Procedural Fidelity Data Collection Form

Date:	**Session #:**	**Time Delay:**
Participants:	**Implementer:**	**Data Collector:**

Identify the participant and target then indicate Yes, No, or NA for each trial

			Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Trial 11	Trial 12
		Participant												
		Target												
		Ensure attending												
		Appropriate wait interval												
		Appropriate controlling prompt												
Dyad 2	One of these columns should be Y or N for every trial; the other two rows should be NA	If UPC or PC: brief praise statement												
		If UPE: wait reminder and error correction (repeat task direction)												
		If PE: error correction (repeat task direction)												
Dyad 1	One of these columns should be Y or N for every trial; the other two rows should be NA	If UPC or PC: reinforce with stamp												
		If UPE: wait reminder and error correction (repeat task direction)												
		If PE: error correction (repeat task direction)												
		Trial was 60 +/- 30s after previous trial												
		Target was related												
		Target was unrelated												

Correct materials presented		
Correct beginning script		
Correct number of trials		

Play-Based Procedural Fidelity Data Collection Form

Date:			**Session:**			**Delay:**	
Participants:			**Implementer:**			**Data Collector:**	
MTS: Check at least one column at the end of each 10 s interval							
Interval	**Parallel Play**	**Comment**	**Respond**	**TD**	**Reinforce**	**Corrective**	**None**
0-10s							
11-20s							
21-30s							
31-40s							
41-50s							
51-60s							
61-70s							
71-80s							
81-90s							
91-100s							
101-110s							
111-120s							
121-130s							
131-140s							
141-150s							
151-160s							
161-170s							
171-180s							
181-190s							
191-200s							
200-210s							
211-220s							
221-230s							
231-240s							
241-250s							
251-260s							
261-270s							
271-280s							
281-290s							
291-300s							

Baseline Procedural Fidelity Data Collection Form

Date:	**Session #:**	**Time Delay:**
Participants:	**Implementer:**	**Data Collector:**
Indicate the number of trials presented to each participant, and tally the frequency of each item for each trial. Indicate yes or no for whether the correct materials were present throughout the session and whether the implimenter stated the correct		

Participant 1:		Participant 2:	
Number of Targets Presented:		Number of Targets Presented:	
Ensure attending		Ensure attending	
No prompt		No prompt	
Neutral response		Neutral response	
Trial conducted within 10s of previous		Trial conducted within 10s of previous	

Correct beginning script	Yes	No
Play materials not present	Yes	No