

6-11-2018

An Efficiency Evaluation of Procedures to Evoke Vocalizations in Children with Autism

Alejandro Rene Diaz

Florida International University, adiaz087@fiu.edu

DOI: 10.25148/etd.FIDC006864

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

 Part of the [Applied Behavior Analysis Commons](#), and the [Developmental Psychology Commons](#)

Recommended Citation

Diaz, Alejandro Rene, "An Efficiency Evaluation of Procedures to Evoke Vocalizations in Children with Autism" (2018). *FIU Electronic Theses and Dissertations*. 3771.

<https://digitalcommons.fiu.edu/etd/3771>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

AN EFFICIENCY EVALUATION OF PROCEDURES TO EVOKE
VOCALIZATIONS IN CHILDREN WITH AUTISM

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Alejandro Rene Diaz

2018

To: Dean Michael R. Heithaus
College of Arts, Sciences, and Education

This dissertation, written by Alejandro Rene Diaz, and entitled An Efficiency Evaluation of Procedures to Evoke Vocalizations in Children with Autism, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Maricel Cigales

Leslie D. Frazier

Stephen E. Wong

Anibal Gutierrez, Jr., Major Professor

Date of Defense: June 11, 2018

The dissertation of Alejandro Rene Diaz is approved.

Dean Michael R. Heithaus
College of Arts, Sciences, and Education

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2018

DEDICATION

This dissertation is dedicated to my wife, Julie, my parents Nancy and Pedro, and my brothers and sisters, Adrian, Carlos, Patricia, Andrea, and Isabel. Thank you for supporting me in the countless ways you do. Some say we are the average of the people we spend the most time with. I've been blessed to be surrounded by people who push me to do better, who believe in me, and encourage me even when I doubt myself. I could never have come this far without you. I love you. Thank you for believing in me.

ACKNOWLEDGMENTS

I would like to thank my mentor, Dr. Anibal Gutierrez, Jr., whose advice and guidance was paramount throughout this process, as well as his calm and collected demeanor while facing challenges. His guidance has helped me learn what it is to be more than just a researcher, but a leader as well.

I would also like to thank my committee members for all their advice and guidance. Their expertise and feedback helped me turn this project into something I am proud of. Their guidance was invaluable, especially regarding research considerations I overlooked.

Finally, I would also like to express my deep gratitude to the Florida Education Fund for providing me with funding, professional development, inspiration, and giving me the honor to be surrounded by brilliant people who challenge and inspire me.

ABSTRACT OF THE DISSERTATION
AN EFFICIENCY EVALUATION OF PROCEDURES TO EVOKE
VOCALIZATIONS IN CHILDREN WITH AUTISM

by

Alejandro Rene Diaz

Florida International University, 2018

Miami, Florida

Professor Anibal Gutierrez, Jr., Major Professor

Development of vocalizations in early learners with autism is critical to the acquisition of verbal behavior and other important life skills. The purpose of the present studies was to (1) evaluate the efficiency and efficacy of Stimulus-Stimulus Pairing (SSP) and standard Echoic Training (ET) procedures for the development and onset of verbal behavior in early learners with ASD to improve early intervention efficiency and (2) elucidate predictive characteristics or variables for the effective use of SSP. The present studies were comprised of a multiple-baseline (across behaviors) experimental design buttressed within a reversal design, also known more broadly as within-subject controlled experimental designs. It was found that SSP can have a greater treatment efficacy than ET, but any efficacy advantage is transitory. Shifting an SSP treatment to direct reinforcement contingencies once vocalizations are produced are likely the most effective strategy. SSP produces discrepant effects across learners, thus highlighting the need to assess a learner's characteristics and assumed reinforcer effectiveness. It was also found that higher-functioning learners will benefit more greatly from ET as opposed to SSP.

TABLE OF CONTENTS

CHAPTER	PAGE
I. Introduction.....	1
Applied Behavior Analysis.....	3
History & Method.....	3
The Three-Term Contingency.....	6
Consequences.....	8
Behavior.....	9
Antecedents.....	10
Verbal Operants.....	11
Communication Intervention.....	13
Echoic Training.....	16
Echoic Training Intervention.....	16
Stimulus-Stimulus Pairing.....	19
Stimulus-Stimulus Pairing Literature.....	21
Purpose of Study.....	32
Significance of Study.....	33
II. Method.....	34
Assessments.....	34
The Mullen Scales of Early Learning.....	34
Autism Spectrum Rating Scales.....	34
Behavioral Language Assessment Form.....	35
Reinforcement Assessment for Individuals with Severe Disabilities.....	36
Stimulus Preference Assessment.....	36
Participants.....	37
Jake.....	37
Max.....	38
Lane.....	40
Jordan.....	41
Dante.....	42
Setting and Materials.....	43
Design.....	46
Procedures.....	46

	Study 1	46
	Pre-session and post-session observations.....	47
	Baseline.....	47
	Control.....	47
	Treatment.....	48
	Stimulus-stimulus pairing.....	48
	Echoic training.....	49
	Reliability.....	50
	Treatment fidelity.....	51
	Problem behavior.....	53
	Study 2	53
III.	Results.....	54
	Study 1	55
	Visual Analysis	55
	Jake.....	56
	Max.....	56
	Lane.....	59
	Jordan.....	63
	Dante.....	63
	Nonoverlap of all Pairs	65
	Study 2	68
	Visual Analysis	68
	Max.....	68
	Lane.....	71
	Dante.....	73
	Nonoverlap of All Pairs	74
IV.	Discussion.....	75
	Limitations and Future Directions	83
	Conclusions.....	86
	References.....	88
	VITA	94

LIST OF TABLES

TABLE	PAGE
Table 1. Skinner’s (1957) Verbal Operants	13
Table 2. Participant MSEL Results.....	39
Table 3. Autism Spectrum Rating Scale Results	45
Table 4. Study 1 NAP Effect Size Scores.....	67
Table 5. Study 2 NAP Effect Size Scores.....	75

LIST OF FIGURES

FIGURE	PAGE
Figure 1. Pre-session, Treatment, and Post-session Observations.....	50
Figure 2. Jake’s Frequency of Vocalizations (Study 1).....	57
Figure 3. Max’s Frequency of Vocalizations (Study 1).....	58
Figure 4. Lane’s Frequency of Vocalizations (Study 1).....	61
Figure 5. Jordan’s Frequency of Vocalizations (Study 1)	62
Figure 6. Dante’s Frequency of Vocalizations (Study 1)	65
Figure 7. Max’s Frequency of Vocalizations (Study 2).....	70
Figure 8. Lane’s Frequency of Vocalizations (Study 2).....	72
Figure 9. Dante’s Frequency of Vocalizations (Study 2)	74

I. Introduction

According to *The Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM–5; American Psychiatric Association, 2013), Autism spectrum disorder (ASD) is a developmental disorder with persistent social communication and social interaction deficits, across multiple contexts, in the domains of: (1) social-emotional reciprocity, (2) nonverbal communicative behaviors typically used for social interaction, and (3) developing, maintaining, and understanding relationships. Children with ASD are also likely to demonstrate restricted and repetitive patterns of behavior or interests. The current version of the DSM has included prior diagnoses of Asperger’s disorder and pervasive developmental disorder not otherwise specified (PDDNOS) as subsets of ASD, which are also expected to be given an ASD diagnosis. One of the core skill deficits in children with autism is the ability to communicate (American Psychiatric Association, 2013), which contributes a significant amount to problems in social interaction. Studies have found that functional communication and language play major roles in the development of children with ASD (Paul & Cohen, 1984; Paul, 2008). Functional communication and language have also been shown to lead to improved outcomes (Luyster, Qiu, Lopez, & Lord, 2007) and are therefore a logical intervention target. For example, early language skills have been demonstrated to be significantly related to social functioning in adulthood; individuals with the highest scores in friendship and social competence tend to also have the highest scores in verbal IQ (Howlin, Mawhood, & Rutter, 2000). It is therefore unsurprising that many treatments would target language and communication skills. Parents of children with ASD have reported using many

different types of treatments (e.g., medication, alternative diets, detoxification, aroma therapy, hippo therapy, and others) to address these deficits (Bowker, D`Angelo, Hicks, & Wells, 2011). Although there are several treatments available, behavioral interventions using Applied Behavior Analysis (ABA) are one of the most widely empirically supported treatments. Many studies (Bowker et al., 2011; Kasari & Lawton, 2010, Virués-Ortega, 2010), have shown ABA interventions are effective for addressing the communicative deficits associated with ASD.

Applied behavior analysis is the practice of applying tactics derived from the principles of behavioral science, systematically, to improve socially significant behavior to a meaningful degree, and to demonstrate that the interventions implemented are responsible for the improvement (Baer, Wolf, & Risley, 1968; Cooper, Heron, & Heward, 2007). Applied behavior analysis has had a growing empirical base of evidence for its benefits for children with autism since the 1980s (Fenske, Zalenski, Krantz, & McClannahan, 1985; Lovaas, 1987) and the efficacy of ABA is reflected in the continued growth in the number of federal and state mandates requiring insurances to cover ABA as a medically necessary treatment.

The strategies, procedures, and principles of ABA are discussed in greater detail in subsequent sections of the present paper. However, it is important to note that some procedures currently employed by practitioners are yet to be thoroughly investigated regarding optimization. Many procedures and tactics in ABA are effective at producing behavior change. However, even effective tactics can be improved if there are alternative sequences or procedural refinements that produce the desired results more efficiently or

make the process of producing the results more socially acceptable or significant. For instance, sequencing procedures in a specific manner might result in fewer trials to reach a criterion (Love, Carr, Almason, & Petursdottir, 2009; Gutierrez, Hale, O'Brien, Fischer, Durocher, & Alessandri, 2009). As a result, an area of study in behavior analysis that remains open to investigation is procedural optimization of well established procedures for the development of verbal behavioral repertoires. Research in the area of vocal development has the potential to influence practitioner sequencing and evaluation of early intervention effectiveness.

Applied Behavior Analysis

History & Method. The history of ABA is often traced at least as far back as 1913, when John B. Watson published *Psychology as the Behaviorist Views It*. Watson (1913) proclaimed that the science of behavior, as the behaviorist views it, was a purely objective and experimental branch of psychology. Watson advocated that the field of psychology dispense with the analysis of mental states, emotions, and consciousness. The philosophy of behaviorism, according to Watson (1913), emphasized observable events which could be precisely defined and recorded. Essentially, for the field of psychology to advance, psychologists would have to recognize that behavior has universal principles, and that human and animal research was on the same plane, even if it was not yet understood how. Watson argued that this philosophy would lead to better ways to talk about behavior and better methods to examine it. Watson's predictions became prophetic through the work of B.F. Skinner. Skinner discovered the basic principles of behavior, which would become known as operant conditioning. Skinner also played a key role in

the development of single-case research methods (SCM), the methods employed in the present study. Skinner, though, was more influenced by Ivan Pavlov than Watson, with respect to methods. As evidence, Skinner stated “Russell and Watson had given me no glimpse of experimental method, but Pavlov had: Control the environment and you will see order in behavior” (Skinner 1967, p. 399). Skinner began controlling the environment and analyzing the effects on behavior of individual subjects, as opposed to groups of them. The reason for Skinner’s emphasis on single-case methods can be explained through his frustrating experience using the group comparisons approach (e.g., Heron & Skinner, 1939) where he describes it as (a) inflexible; (b) not providing information at the level of the individual organism and instead creating a “nonexistent” average subject; (c) promoting the “smoothing” of data variability by increasing the size of the groups, instead of forcing the experimenter to explain such variability; and (d) not providing useful information for the purpose of identifying functional relationships (Hurtado-Parrado & López-López, 2015).

Skinner’s criticisms of group comparisons over half a century ago have been bolstered by recent controversies regarding Null Hypothesis Significance Testing (NHST) (Cohen, 1994; Krantz, 1999) and the replicability of experiments in psychological research (Open Science Collaboration, 2015). A recent collaboration (Open Science Collaboration, 2015) attempted to replicate 100 experimental and correlational studies published in *Psychological Science (PSCI)*, *Journal of Experimental Psychology: Learning, Memory, and Cognition (JEP: LMC)*, and *Journal of Personality and Social Psychology (JPSP)*. The reproducibility of *P* values, effect sizes, subjective assessments of replication teams, and meta-analysis of effect sizes were evaluated.

Average effect size magnitudes of the replications ($M_R = 0.197$, $SD = 0.257$) were found to be half of the original effects ($M_R = 0.403$, $SD = 0.188$) and of the 97% of original studies that reported significant results ($P < .05$), only 36% of replications had a significant effect. Correlational analysis indicated that the strength of evidence in the original studies were more highly correlated with the ability to replicate findings than were the research teams conducting the replication studies. These findings highlight the need to improve the strength of studies through within-study replication. The findings also call into question the methods and practices of researchers in the field of psychology. A major drawback of NHST is that it frames experimental questions in binary terms. Was there sufficient evidence to reject the null hypothesis? In both situations, when there is or is not, the findings produce simple yes or no answers which do not explain why or generate functional relationships at the individual level. The NHST approach forces researchers into experiments for detecting a predicted difference rather than elucidating properties or features of a relationship between an independent and dependent variable (Johnston & Pennypacker, 1993).

Unfortunately, by using NHST and group design methodology, a researcher may overlook the individual characteristics and variables for each participant in the study. An alternative to this is to ask different questions, such as, “how can procedures be optimally sequenced and what indicators can practitioners use to guide procedural sequencing?” and use methods that allow for closer individual analysis. In addition, incorporating replication to increase confidence in the integrity of the findings is a necessity. Fortunately, as Skinner discovered many decades ago, single-case research methods (SCRM) is extremely well suited to accomplishing these aims. Therefore, the present

study aims at answering the experimental questions using SCRM experimental design. One of Skinner's most influential discoveries with the use of SCRM is what came to be known as the three-term contingency. The three-term contingency would expand the prior two-term contingency, also known as the stimulus-response (i.e., S-R model) model of behavior.

The Three-Term Contingency

One of the most fundamental and basic principles of behavior analysis is that operant behavior is maintained by consequences. Some of the terms used to describe specific cases of consequence-maintained behavior are described with greater detail in subsequent sections. This principle was discovered and made explicit simultaneously with the formal creation of the experimental branch of behavior analysis, formed by B.F. Skinner with the publication of his book, *The Behavior of Organisms* (1938/1966). Skinner's book was a summary of all the behavioral research he had conducted between 1930 and 1937. In effect, Skinner had discovered a new classification of behavior; operant behavior. Prior to Skinner's contribution, reflexive behavior, called respondent behavior, was discovered by Ivan Pavlov (1927/1960) and explained how some behavior was controlled, elicited, or caused by stimuli that immediately preceded behavior. Respondent behavior was also known as the two-term contingency (i.e., S-R). However, it became apparent that respondent behavior could only account for a limited subset of responding. Respondent behavior's limitation likely inspired Skinner. Skinner was interested in giving a scientific account of all behavior (Glenn, Ellis, & Greenspoon, 1992). As one researcher eloquently stated:

He did not deny that physiological variables played a role in determining behavior. He merely felt that this was the domain of other disciplines, and for his part, remained committed to assessing the casual role of the environment. This decision meant looking elsewhere in time. Through painstaking research, Skinner accumulated significant, if counterintuitive, evidence that behavior is changed less by the stimuli that precede it (though context is important) and more by the consequences that immediately follow it (i.e., consequences that are contingent on it). The essential formulation for this notion is S-R-S, otherwise known as the three-term contingency. It did not replace the S-R model—we still salivate, for instance, if we smell food cooking when we are hungry. It did, however, account for how the environment “selects” the great part of learned behavior. With the three-term contingency Skinner gave us a new paradigm. He achieved something no less profound for the study of behavior and learning than Bohr’s model of the atom or Mendel’s model of the gene. (Kimball, 2002, p. 71)

The three-term contingency as the unit of analysis was therefore, in many ways, a conceptual breakthrough (Glenn et al., 1992) for the science of behavior. The three-term contingency unit of analysis allowed for research to focus on the relationship between behavior and environment more astutely than the S-R model and brought clarity to the experimental analysis of behavior. More specifically, the three-term contingency refers to antecedents, behaviors, and consequences (i.e., stimulus-response-stimulus); Antecedents are any stimuli, contexts, or setting events that precede the behavior. Behaviors are the specific target responses one is interested in. Consequences are the stimuli that

immediately follow the behavior. To clarify and expand on each of the three terms in the three-term contingency, the following sections require an individual analysis.

Consequences. Skinner's discovery that consequences were the primary determinant of behavior led to the investigation and development of different taxonomies for environmental outcomes (i.e., variables). Before defining specific types of consequences, some broad definitions should be considered. Foremost, all determinants of behavior occur within some environmental context. Behavior does not occur in a void or vacuum, and though there are different ways to define environments, definitions are arbitrary. It is therefore important to recognize that when any analysis is made, it is usually with the intention to limit the domain of analysis for the sake of building a foundation of knowledge at a specific level of analysis. Stated differently, there are many determinants of behavior that fall outside of the definitions proposed in this section, but for the present, limiting the discussion is intentional. *Environment* has been defined as all the circumstances in which the organism or part of the organism is located, or everything except the moving parts involved in the behavior (Johnston & Pennypacker, 1993). The important implications of this definition are that (1) only real physical events are considered and (2) other parts of the organism can serve as an environment for a behavior. Thus, the skin of an organism is not necessarily the barrier for the definition of environment, and the barrier will change depending on the behavior of interest. A definition for the term *stimulus* is "an energy change that affects an organism through its receptor cells" (Michael, 2004, p. 7). An example of a stimulus would be the sound waves coming from the speaker of a radio which affect an organism through its hearing receptors. In contrast, even though electro-magnetic waves (i.e., radio waves) can be

defined as an energy change, if they do not affect the organism through some receptor, as they are not detectable until transformed into sound or light, they are not considered a stimulus.

There are at least three types of consequential stimuli: *Reinforcers*, *Punishers*, and *neutral* stimuli. There are different variations of the definitions for these terms but in the present paper we will use the following: A reinforcer is a stimulus change, when contingent upon a response, increases or maintains the likelihood of that response in the future, under similar conditions. A punisher is a stimulus change, when contingent upon a response, decreases the likelihood of that response, in the future, under similar conditions. A neutral stimulus is one which has no effect on the likelihood of the behavior. Other types of consequences are beyond the scope of the present paper and are therefore not addressed.

Behavior. Defining behavior in what may be the most conceptually sound and complete definition to date is the following excerpt by Johnston and Pennypacker (1993):

The behavior of an organism is that portion of organism's interaction with its environment that is characterized by detectable displacement in space through time of some part of the organism and that results in measurable change in at least one aspect of the environment. (p. 23)

This definition restricts the subject matter to behavior of organisms, as opposed the "behavior" of objects in outer space, or interactions between other non-living matter. The definition also emphasizes that behavior is measurable movement. Behavior must change some aspect of the environment, as opposed to being changed by it. This limitation

eliminates certain expressions as behavior, such as “laying still” or “getting tired.”

Behavioral definitions are especially important to distinguish since they separate antecedents and consequences. Antecedents come immediately before the behavior.

Antecedents. Antecedents refer to the stimulus changes or conditions that exist prior to the behavior of interest. Behaviors can also function as antecedents for other behaviors. Some of the most commonly referenced antecedent conditions or stimuli are called discriminative stimuli (S^D) and motivating operations. Discriminative stimuli set the occasion for behavior to occur by signaling to the organism that reinforcement is available for responding. An example of this would be the computer screen of a computer, if the screen is on, typing and moving a mouse is very likely to produce a desired reinforcer, such as visual stimuli on the screen. Motivating operations are categorized as establishing operations and abolishing operations. Establishing operations increase the effectiveness of a stimulus, object, or event as a reinforcer, and also alter the current frequency of all behavior that has been reinforced by that stimulus, object, or event (Michael, 1982). An abolishing operation reduces the effectiveness of a reinforcer and the momentary frequency of that behavior. Eating salty food would function as establishing operation for drinking behavior, while consumption of water would function as an abolishing operation. Motivating operations have been described as the fourth-term, adding to the three-term contingency.

Together, antecedents, behaviors, consequences, and motivating operations have given the experimental analysis of behavior a basic model to predict, control, and explain behavior. It is unsurprising then, models for complex human behavior would arise from

these basic principles. One highly complex area of human behavior where the four-term contingency model has come to proliferate is in the area of language, or as Skinner coined it, verbal behavior.

Verbal Operants

Arguably, one of the most important works in the field of ABA is *Verbal Behavior* (1957) by B.F. Skinner. In *Verbal Behavior* (1957), Skinner provides a conceptual framework and taxonomy for verbal behavior. One of the most important and misunderstood arguments Skinner makes in *Verbal Behavior* (1957) is that the term “language” generally refers to the topography of speaker behavior. A significant amount of language research had focused on language and the structure of speaker behavior and the differences between them. In contrast, Skinner chose to focus on the functional relationship between responses of the speaker and the listener. Stated differently, English, Russian, and American Sign Language all differ in the way they sound and/or appear, but they all share and develop the same functions. These functions stem from the four-term contingency and produce the desired reinforcers through a specific and important part of the environment: other individuals. Although all three languages look and sound differently, the spoken response of “water please” in its corresponding form for all languages will likely produce the same exact result (consequence) from another individual that speaks the same language or has a learning history from the same verbal community. Another crucial point that Skinner put forth in *Verbal Behavior* was that verbal behavior, in contrast to other operant behavior, was strictly socially mediated. That is, communication or speaker behavior, can only be reinforced by a listener. Although,

with time, a speaker could also come to serve as his own listener, speaker and listener behavior would have to be developed through a community. Skinner's perspective was a significant departure from previous explanations and led to Skinner's entire verbal operant classification. Skinner identified six verbal operants: mands, tacts, echoics, textuials, transcriptions, and intraverbals. The six verbal operants and their controlling variables are outlined in table 1. Generally speaking, mands are requests for something, such as when a child asks for a new toy, saying "Can I have this toy airplane?" Tacts are essentially the identification of some non-verbal stimulus that is present, such as when one sees an airplane and yells "An airplane!" Echoics are repeated speaker behavior, such as when a model says "airplane" and the listener repeats "airplane." A textual operant is an operant in which a reader reads aloud the written word, such as seeing the written word airplane and saying "airplane." A transcription is the writing of the word, such as hearing a person say "airplane" and then writing it out. An intraverbal is an operant that is controlled by the behavior of another speaker, such as when a speaker says, "I have yet to see an airplane today," and the person responds, "I have seen at least two airplanes today." These verbal operants are often targeted for skill acquisition in communication training protocols of ABA. Although some protocols do not label the target skills using the terms outlined in Skinner's analysis of verbal behavior (SAVB), ABA interventions target the same type of response classes using some other categorical label, such as "expressive identification of objects" (in lieu of tacting objects).

Communication Intervention

There are some common characteristics of ABA interventions for children with ASD, such as; (1) direct measurement of a behavior being targeted, (2) contingency management, which is the management of antecedents and the consequences of a targeted

Table 1
Skinner's (1957) Verbal Operants

Antecedent Controlling Variable	Verbal Operant	Controlling Consequence
Establishing operation	Mand	A specified reinforcer
Verbal Stimulus with point to point correspondence and formal similarity	Echoic	Non-specific reinforcement
Present non-verbal stimulus	Tact	Non-specific reinforcement
Verbal Stimulus without point to point correspondence and formal similarity	Intraverbal	Non-specific reinforcement
Verbal Stimulus with point to point correspondence but without formal similarity	Textual & Transcriptive	Non-specific reinforcement

behavior, (3) precise behavioral language, (4) implementation of procedures or tactics to modify behavior, (5) promotion of skills for independence and socially significant behaviors, (6) pinpointing specific behaviors for modification, and (7) graphing data to guide intervention decisions (Lovitt, 2012). Using these tactics, communication and language skills are targeted to address the communication and social deficits experienced

by children with autism. Previous research (Helt et al., 2008; McEachin, Smith, & Lovaas, 1993; Rogers & Vismara, 2008) has found that behavioral interventions can lead to ideal outcomes in some individuals. Individuals experiencing ideal or optimum outcomes is defined as: (1) Having had a convincing history of ASD, such as diagnosis and the child's development will have been delayed, (2) eventually the learner must be learning and applying a core set of skills with a quality that reaches the trajectory of typical development in at least most skill domains, and (3) the individual will no longer meet criteria for ASD (Helt et al., 2008). In the present paper it is assumed that children and adolescents that have experienced an optimum outcome will be able to function independently in typically developing classroom settings. Adults that have experienced an optimum outcome will be living independently, working full-time, often be married, and have friends.

It can be reasonably assumed that most interventions aim to achieve or contribute to an optimum outcome. As a result, it is important to identify predictors of optimum outcomes. One particularly important predictor of optimum outcomes is the child's age at the time of admission into an intensive behavioral intervention program using ABA. One study (Harris & Handleman, 2000) found a significant relationship between age of admission and the ultimate educational placement, $r(25) = .658, p < .005$, such that those children who were admitted into an ABA program at a young age were more likely to be in regular class room settings at follow-up when compared to older children. These findings indicate a crucial need for early identification and admission of children with ASD into intensive ABA programs. Another important and consistent prognostic indicator of optimal outcomes is early communication and language abilities (e.g.,

Howlin et al., 2000; Paul & Cohen, 1984; Paul, 2008; Ventner, Lord, & Schopler, 1992). For instance, it has been found that communication scores at age 2 and 3 years predict language and other outcomes at age 9 (Luyster et al., 2007), and early language skills have also been found to be significantly related to social functioning in adulthood, with those individuals initially scoring higher in verbal IQ also tending to have higher scores for outcomes in friendship and social competence (Howlin et al., 2000).

In addition to increasing the likelihood of optimal outcomes, communication and language skills have benefits in other areas of functioning. For instance, it has been found that receptive communication skills were associated with advances in daily living skills, social skills, and reduced the frequency of problem behaviors related to social interaction for children with autism (Park, Yelland, Taffe, & Gray, 2012). Optimal outcomes become more evident with early intervention compared to later interventions because the benefits of intervention cascade into other areas of development at an earlier period.

Many curriculum tools (e.g. Vineland, ABLLS-R, VB-MAPP) for ABA interventions are primarily composed of language and communicative behavior skill acquisition programs (e.g. Leaf & McEachin, 1999; Lovaas, 1981; Lovaas, 2003; Sundberg & Partington, 1998). As a result of the profound importance of early diagnosis, early admission into an ABA program, and a focus on communicative behavior, it is unsurprising that a recent survey (Love et al., 2009) of clinical practices indicates that these curriculum tools are widely used by early intensive behavioral interventions (EIBI) and ABA intervention services. Echoic training is one of the primary methods used to

teach verbal behavior once practitioners use these curriculum tools to elucidate a learner's current behavioral repertoire.

Echoic Training. For early learners, targeting echoic verbal operants, or vocal imitation as some practitioners may call it, is one of the first steps in building a verbal behavior repertoire. Echoic training and vocal imitation is generally done using simple reinforcement and shaping procedures. The instructor provides a vocal model and reinforces any similar response approximations by the learner. For instance, if a target vocalization is the word "ball," the practitioner would likely start by presenting the S^D "say bah", reinforcing the "bah" vocalization, then repeating the procedure for the "all" sound, and finally reinforcing an increasing fluency of the two sounds in succession until the learner is fluently saying "ball."

Echoic Training Intervention. Often, an echoic verbal repertoire will be the initial target of a verbal acquisition program since echoic skills are often used as a prerequisite for other verbal operant training procedures, such as mand training. An echoic repertoire can also serve as a prognostic indicator of a learner's ability, as pointed out by Sundberg and Michael (2001):

Information regarding the quality and strength of the echoic repertoire can reveal potential problems in producing response topographies that are essential for other verbal interactions. If the child cannot echo specific sounds, then the probability of those responses occurring in other functional units of verbal behavior is quite low (p. 706).

Unfortunately, echoic training requires for an imitative response to occur for the delivery of a reinforcer, which may not naturally occur during training trials. Therefore, echoic training can prove difficult for some early learners. For instance, Cividini-Motta (2014) found echoic training to be ineffective for one of three participants in a study comparing echoic training (ET) to stimulus-stimulus pairing (SSP), an alternative procedure. As a result, some early learners may not immediately benefit from verbal acquisition training when ET is used exclusively. One possible explanation for this phenomenon is that early learners with autism demonstrate a decreased attending to social stimuli (Chawarska, Macari, & Shic, 2013) which delays or prevents development of imitation skills. As a result, the reinforcement contingencies that typically emerge during development for typically developing infants do not occur for those with autism, such as reinforcer pairing contingencies. A reinforcer pairing contingency is one in which a reinforcer is presented concurrently with a neutral stimulus. With sufficient pairings, the neutral stimulus becomes a conditioned reinforcer. In the case of a developing infant, a mother's affectionate and playful vocalizations "say mama" may become paired with the presentation of food, warmth, touch, and other reinforcers commonly delivered by the mother. Over time and many pairings, the sound "mama" may become a reinforcer. Producing those sounds would therefore be reinforcing for the infant. Skinner (1957) noted that automatic reinforcement, as in this case, could occur to strengthen a variety of behaviors that produce the reinforcing consequence. Stated differently, the infant being able to produce that particular sound, approximations of it, or other sounds produced by the mother would be automatically reinforced just by doing so. However, these processes

that contribute to typical verbal behavior development may be delayed or absent in some children with autism.

Additionally, since vocal imitation and other pivotal prerequisite skills require some prior neuromuscular development, such as the ability to coordinate motor movements (e.g., of the mouth), there are times when interventions that attempt to train vocal imitation immediately in very early learners fall short of doing so, at least initially. If the imitative behavior never occurs, the learner will not experience a reinforcement contingency, thereby eliminating the ability to increase the probability of echoic behavior. Many practitioners may consider this a situation in which prerequisite skills should be targeted and often do. Though targeting pre-requisite skills is a logical step, and many skills can be shaped, crossing the chasm from prerequisite skills to imitative skills can often take a significant amount of time. Furthermore, one of the many benefits of vocal imitation training is that it leads to the neuromuscular development through shaping processes. Learners may not be able to say the word “Ball” initially, but through imitation training, a “Buh” sound can eventually become a “Ball.” Helping the early learner imitate the first “Buh” sound, however, is likely where much time is spent getting that “first response.”

Because of the difficulty with producing vocalizations with the early learner population and to further elucidate how verbal behavior develops, several studies have investigated an alternative method for evoking vocal behavior in early learners which is called stimulus-stimulus pairing (Carrol & Klatt, 2008; Esch, Carr, & Grow, 2009; Esch, Carr, & Michael, 2005; Lepper, Petursdottir, & Esch, 2013; Miguel, Carr, & Michael,

2002; Miliotis, et al., 2008; Normand & Knoll, 2006; Rader, et al., 2014; Stock, Schulze, & Miranda, 2008; Sundberg, Michael, Partington, & Sundberg, 1996; Ward, Osnes, & Partington, 2007; Yoon, 1998; Yoon & Bennett, 2000; Yoon & Feliciano, 2007).

Stimulus-Stimulus Pairing. Stimulus-stimulus pairing (SSP) is a process by which two or more stimuli are presented together. During the process, the paired stimulus may acquire some of the properties of the pairing stimulus. For instance, saying “nice job” (paired stimulus) while delivering a reinforcer (e.g., five dollars, a sticker, food, etc.) would be an instance of pairing. As a result, the phrase “nice job” may function as a reinforcer for speaker or listener behavior in the future, especially with repeated instances of pairing. Pairing can also occur with aversive stimuli, such as when a police officer (paired stimulus) delivers a fine (aversive). As a result, the police officer may become an aversive stimulus.

Researchers have observed that changes in typography and range of vocalizations in infants and young children occur rapidly during early years of development without direct reinforcement (Holland, 1992; Kravitz & Boehm, 1971; Mowrer, 1954; Nakazima, 1962; Thelen, 1979, 1981). These observations have contributed to the belief among psycholinguists that reinforcement is insufficient to explain language acquisition (Yoon & Bennett, 2000). Although reinforcement can be defined as a *procedure* by which others deliver reinforcers for specific behavior, reinforcement can also be defined as a *process* by which behavior products or consequences that automatically occur contingent upon the behavior, increase the likelihood of a behavior. For instance, vocalizations have stimulus properties, such as proprioceptive, exteroceptive, and auditory stimulation.

Similarly to how adults may speak to themselves, sing a song aloud, or hum a familiar tune, infant produced vocalizations and babbling may qualify as stimulation that is automatically reinforcing. Automatic reinforcement is therefore reinforcement that does not require mediation of consequences by another organism. It is also likely that under some conditions, some of these vocalizations will be shaped into more complex vocalizations. Other processes, such as direct reinforcement, and naturally occurring SSP can further strengthen the likelihood of these vocalizations occurring. Skinner (1957) provided an example of how this process can unfold in *Verbal Behavior*:

The young child alone in the nursery may automatically reinforce his own exploratory vocal behavior when he produces sounds which he has heard in the speech of others. The self-reinforcing property may be merely an intonation or some other idiosyncrasy of a given speaker or of the speakers in general. Specific verbal forms arise from the same process. The small child often acquires verbal behavior in the form commendation used by others to reinforce him. The process is important in the automatic shaping up of standard forms of response. (p. 58)

Infants that produce sounds during reflexive behavior, such as crying and coughing, also strengthen their vocal muscles (Bijou & Baer, 1965) and develop the nervous system needed to produce more varied types of vocalizations. It is likely that many of the processes that lead to automatically reinforcing stimuli begin in the womb, where the fetus can hear its mother's voice long before being born. The implications of automatic reinforcement as the primary driver for the development of early verbal behavior is

profound. It leads to many new questions, such as: When is it more effective to focus on automatic reinforcement versus direct reinforcement procedures?

Since it has been found that children with ASD often do not orient to social stimuli (Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998), it is likely that many social stimuli are not functioning as reinforcers or are not sufficiently salient (Shillingsburg et al., 2015). As a result, many of the naturally occurring SSP and direct reinforcement contingencies which contribute to the verbal development of typically developing children, do not come to bear on those with autism, which may naturally lead to delays in many areas of development that require verbal repertoires.

Stimulus-Stimulus Pairing Literature. To determine whether SSP of speech sounds could lead to increases in vocalizations, several studies have examined the effects of SSP under a variety of experimental conditions.

The first known study to examine SSP's effect on vocalizations in humans was published a relatively short time ago (Sundberg, Michael, Partington, & Sundberg, 1996). The study by Sundberg et al., (1996) examined whether new vocal responses could be established in the babbling repertoires of five children by using SSP. The study included five children between the ages of two and four. Four of the participants were described as having severe to moderate language delays. The fifth participant was described as typically developing. Two of the participants were diagnosed with autism (ages 4-years-old and 3-years-old), one with mental retardation (age 4-years-old), and the final developmentally delayed participant (age 2-years-old) did not have a diagnosis but was described as born to a crack cocaine addicted mother and raised by a developmentally

disabled Aunt. The typically developing participant was 2-years-old and lived at home with her family. During the study, the participant's vocal responses were recorded and categorized by Skinner's verbal operants or as 'other' vocalizations (e.g. automatically reinforced vocal play, random, vocalizations, reflexive vocalizations). Words that were known to be novel or that did not occur in the pre-pairing condition were selected as target words. For those participants with high rates of vocalizations, only the targeted sound or phrase was recorded throughout the experiment. The study employed an AB design that compared each subject's performance during baseline and after pairing. During pairing, an established form of reinforcement such as tickles, praise, clapping, or bouncing was used with approximately 15 pairings per minute.

In the Sundberg et al. (1996) study, it was found that for all participants, the pairing of a sound, word, or phrase (e.g., dee dum) by the experimenter with an established reinforcer (e.g. tickles) resulted in the unprompted emission of that response after the pairing condition. There were some occasions where the pairing did not produce an increase in vocal behavior, but overall novel vocalizations increased across all participants. Finding SSP could produce vocalizations was significant in that it demonstrated new responses can be acquired by a learner without direct reinforcement, direct verbal training, or prompts to respond. Another finding was that the pairing led to mands for particular reinforcers. For instance, one participant began to request head shakes with an established mand frame ("Baba, do ___") and the target vocalization of "Dee dum" after SSP. The developmental significance of mand emergence is profound and adds evidence for the hypothesis that direct reinforcement is not needed for the emergence of language. Rather, direct-reinforcement strengthens verbal behavior that

may emerge through naturally occurring SSP. The findings of Sundberg et al. (1996) are also significant in that they occurred for developmentally heterogeneous individuals. Another finding of note was that for at least one of the participants, the emotional state of the participant affected the rate of vocalization. For instance, if the participant was quiet and sullen, pairing did not produce an increase in vocalizations. The researchers also noted that during a second experiment, which was comprised of only one of the participants, they were able to replicate rates of responding. The unprompted vocalization rates returned to a pre-pairing level after only a few minutes. Essentially, Sundberg et al. (1996) found that pairing effects are immediate and robust, but temporary.

Another study (Yoon & Bennett, 2000) soon followed and tried to use SSP to condition vocal sounds as reinforcers with tickles. They also compared the effects of SSP to echoic training (ET). There were three participants between the ages of three and four-years-old. The three participants were described as having severe developmental delays, no speaker behavior, and limited listener skills. One participant could imitate large gross motor movements without any physical prompts. The other two participants required prompting to imitate large gross motor movements. None of the participants had oral motor or vocal verbal imitation skills. Regarding vocal play, only one participant seemed to engage in any vocal play. All target sounds were one-syllable words or target sound approximations. Similar to the Sundberg et al. (1996) study, a single-subject AB design with a pre-pairing and post-pairing phase was used. The researchers did approximately 12 pairings per minute for a total of three minutes of pairing.

Data from the first experiment in the Yoon & Bennett (2000) study were consistent with previous findings in that the SSP pairing procedure was effective in increasing the rate of target vocalizations. The target vocalizations all began at a baseline rate of zero and increased significantly after the pairing procedure. The researchers noticed that the occurrence of target sounds was more significant for the participant that was described as already having vocal play skills. It was suggested that the reinforcement history through pairing may have had an influence on the effect size of these procedures. Reinforcement history through pairing is an important consideration when determining developmental indicators of efficacy.

In a second experiment (Yoon & Bennett, 2000), vocalizations were compared for SSP and ET. The second experiment was similar in design to the first experiment but added an echoic and post-echoic phase, however these phases were not counter-balanced. During the echoic phase, the experimenter emitted the target vocalization and would have delivered the reinforcer if the participants had emitted the target sound. None of the participants emitted the target vocalization during the echoic phase. Results from the second experiment demonstrated that SSP was effective at evoking target behaviors whereas the ET procedures were unsuccessful at evoking the behavior. These target vocalizations were novel in that they had not been observed before. The investigators indicated that an important area for investigation is to determine whether the automatic reinforcement associated with SSP could lead to faster acquisition rates of echoic or manding behavior. However, two possible confounds of the results of the Yoon and Bennett (2000) study were that there was no counter-balancing of the phases or control for adventitious (i.e., accidental) reinforcement. Therefore, it is possible that the

significant increase in vocalizations was the result of direct reinforcement instead of automatic reinforcement of SSP. It is also important to note that for both studies (Sundberg et al., 1996; Yoon & Bennett, 2000) discussed thus far, all data were collected within a single session. Most ABA interventions occur across several sessions or days when targeting a new skill.

To address the limitations and extend the findings of these previous studies, another study (Miguel et al., 2002) was conducted to evaluate SSP. A multiple baseline design across vocal behaviors with a reversal design was used to assess the SSP procedure's effects on one-syllable vocalizations of three participants diagnosed with autism. Two of the participants were five-years-old and the third participant was three-years-old. The participants were described as being able to emit a few sounds but could not exhibit more meaningful verbal behavior like mands and tacts. A participant named Leo was very cooperative and had good receptive (listener) and matching-to-sample skills. However, he did not have a generalized echoic repertoire. Leo would spontaneously vocalize certain sounds. Another participant, Rob, also did not have an echoic repertoire. The third and final participant, Dave, was described as very cooperative and had a generalized motor and vocal imitation repertoire as well as excellent receptive and matching-to-sample skills. Dave was also able to mand for five different items without prompting.

The targeted sounds for each participant were the lowest frequency one-syllable vocalizations produced during baseline. Close approximations were also recorded as the target sound. Similar to the previous studies, there were pre-session and post-session

observations to collect data. Phases consisted of baseline (A), control (\hat{A}), and pairing (B) conditions with a return to baseline. The baseline phase was identical to pre-session and post-session observations. In the control phase, the experimenter presented the vocalization five times and would wait approximately 20 s to deliver a reinforcer. The delay was used to ensure that the utterance or presentation of the target vocalization by the experimenter was not sufficient to produce an increase in vocalizations, but rather the simultaneous pairing procedure itself was responsible. Vocalizations by the participant would cause a delay of 20 s before the experimenter would deliver the reinforcer during this phase and also in the SSP phase to prevent adventitious reinforcement. Each session of pairing consisted of 20 trials.

The results of the Miguel, Carr, and Michael (2002) study partially replicated the those of the previous studies (Sundberg et al., 1996; Yoon & Bennett, 2000). For Leo, SSP produced an immediate replicable increase in both target sounds across sessions. For Rob, the pairing produced an immediate increase in his first target sound, but only resulted in more consistent occurrences not exceeding baseline levels in the second target vocalization. For Dave, SSP turned out to be ineffective in increasing target sounds. The researchers noted a similar effect as those noted by Sundberg et al. (1996) where in some sessions the participants vocalized less often. These findings are consistent with prior studies and according to Miguel, Carr, and Michael (2002), are also consistent with an unpublished study by Bennett and Yoon (2000) that found that the more advanced a child's verbal repertoire was, the less likely he or she was to be responsive to SSP. Verbal repertoire was defined by Bennett and Yoon (2000) as the number of functional response forms (i.e., echoics, mands, tacts, intraverbals) and the rate of vocalizations produced by

the child. It was hypothesized (Miguel et al., 2002) that for children with advanced verbal repertoires, competing responses (e.g. mands, tacts, and intraverbals) were more likely to produce reinforcement and thus more likely to occur.

The authors (Miguel et al., 2002) also noted that only vocalizations in the repertoire were targeted to ensure that responses were possible by the participants. However, future research may want to investigate differential effects of SSP on novel vocalizations versus those already in the repertoire. It was also reported that similar to the Sundberg et al. (1996) study, one participant (Dave) began to mand using one of the paired sounds, adding evidence to the hypothesis that mand verbal operants emerge from naturally occurring SSP. Overall, the results from these studies support the notion that SSP can be used to supplement direct reinforcement procedures in early learners with language delays.

To examine the effects of SSP and potentially address the transient (i.e., short lived) effects of SSP as a clinically relevant tool, Esch, Carr, and Michael (2005) conducted a follow-up study. The study consisted of three participants diagnosed with ASD. The participants Alexa, David, and Jodi were 6 years 10 months of age, 6 years 11 months of age, and 8 years 2 months of age, respectively. Participants were described as having age-equivalent scores below 2 years of age on the Kaufman Speech Praxis Test (KSPT; Kaufman, 1995). Echoic skills for the participants were also evaluated using the Behavioral Language Assessment (BLA: Sundberg & Partington, 1998) and resulted in a score of 1 from informants for the vocal imitation section, indicating that participants were unable to repeat any sounds or words. A Receptive-Expressive Emergent Language

Test, Third Edition (REEL; Bzoch, League, & Brown, 2003) was also completed and indicated that all participants were developmentally below 12 months of age for receptive and expressive language measures.

Since SSP had been reported to produce only temporary effects (Esch et al., 2005) in prior studies, it was hypothesized that this was likely a result of extinction (i.e., unpairing) as a result of the newly emitted vocalizations not coming into contact with reinforcement or no longer being paired with reinforcers. Therefore, it was assumed that a decline in those vocalizations could be prevented through direct reinforcement (Esch et al., 2005). The rate of vocalizations would strengthen and thereby be more likely to come into contact with naturally occurring reinforcement contingencies. Thus, the experimenters (Esch et al., 2005) set out to bring newly acquired vocal responses via SSP under echoic control (i.e., direct reinforcement). However, the experimenters were unsuccessful in bringing paired vocalizations under echoic stimulus control with three participants and three target vocalizations per participant. The experimenters then carried out a second experiment to replicate the positive effects of SSP reported by previous studies. Experiment 2 yielded null effects for SSP with the three participants, suggesting that SSP was not an effective intervention for increasing the frequency of post-pairing free-operant vocalizations. The null results were a partial replication of a previous study (Miguel et al., 2002) in which one of the participants did not demonstrate an increase in vocalizations following SSP. The experimenters suggested that children with weak verbal skills did not benefit from a pairing procedure. However, other variables responsible for procedural effectiveness (e.g., reinforcer effectiveness) had not been evaluated and it was also likely that other unidentified variables could be responsible for suppressing the

effectiveness of the SSP procedure. A third experiment conducted by the researchers (Esch et al., 2005) to evaluate the degree to which specific vocal responses were sensitive to reinforcement via a simple shaping procedure. It was found that for one of the two participants, the reinforcement procedure was effective at producing an increase over baseline levels, indicating that the reinforcer was at least somewhat effective. However, for the second participant, baseline and treatment did not differ. These results underscored the importance of evaluating reinforcer effectiveness but also the need to uncover other variables that appear to affect the effectiveness of SSP.

Other studies (e.g., Normand & Knoll, 2006; Stock et al., 2008; Yoon & Feliciano, 2007) have also reported absent or discrepant SSP effects. Some determinants potentially affecting effect sizes could include; reinforcer effectiveness, preexisting language skills, and measurement systems not sensitive enough to detect effects of SSP (Esch et al., 2009). Because of the many conflicting results, it is likely that many SSP studies have not been optimally arranged to produce effects. A recent literature review (Shillingsburg et al., 2015) examined and summarized 13 experiments related to SPP that were published between 1996 and 2014. Across the studies reviewed, there was a significant amount of procedural variability, which makes drawing conclusions a difficult process. Despite this, the authors of the literature review were able to provide a systematic quantitative analysis and quantify effectiveness of the variables in eight studies with 19 participants that used SSP to increase vocalizations in children with language delays using nonoverlap of all pairs (NAP). Nonoverlap of all pairs (NAP; Parker & Vannest, 2009) is a nonparametric effect size calculation that can be used with SCRM design to determine intervention effectiveness and is further discussed in the

subsequent analysis section. The average age of participants across the studies included in the literature review was three years and seven months ($M = 43$ months, range 1-8 years), with most studies evaluating preschool children. The majority of participants were males, with about two males for every female. Most of the participants were diagnosed with autism (69.2%), with other diagnoses consisting of educational delay (15.4%), developmental delay (12.8%), and intellectual disability with visual impairment (2.6%). There was not a consistent language assessment conducted across studies, though 15 out of the 19 participants used in the NAP analysis were described as having no functional language. Functional language was defined as being able to mand, though the extent of the repertoire varied across studies, where some participants also had hundreds of mands, tacts, and intraverbals. The nonfunctional language group was described as either having echoics or vocalizations in their repertoire. In addition to participant variability, studies also varied across: (1) type of target sound (novel or in repertoire), (2) number of experimenter-emitted sounds per pairing, (3) types of pairing procedure, (4) number of pairings per minute, (5) control for adventitious reinforcement, and (6) types of preferred item pair.

Since the studies often targeted more than one sound per participant, effect sizes were calculated for 35 targeted sounds. Effect sizes are described as *small/weak* (0-0.65), *medium/moderate* (0.66-0.92), and *large/strong* (.93-1.0). Average effect size was moderate (0.72, $SD = 0.20$; 95% CI [.64-76]) across studies. Across targeted sounds, there were a total of 12 weak effects (34%), 17 moderate effects (49%), and 6 strong effects (17%).

According to Shillingsburg et al. (2015), a higher percentage of children who were five-years-old or less demonstrated moderate to strong effects compared to those older than five-years-old. Participants who had no functional communication also showed a higher percentage of moderate to strong effects compared to those who had functional language. None of the participants with functional language showed strong effects of SSP. These findings imply that younger children without functional communication may be more likely to benefit from SSP and is a major consideration of the present study.

Other variables that Shillingsburg and colleagues (2015) reported as having the highest percentages of moderate to strong effect sizes included those for whom only edible reinforcers were used, participants for whom control of adventitious reinforcement were used, and those who received delay conditioning during SSP. Delay conditioning is described as presenting the paired stimulus (e.g., the sound) followed immediately, and possibly overlapping, with the preferred item (i.e., reinforcer) (Miliotis et al., 2012; Shillingsburg et al., 2015). It was also noted that controlling for adventitious reinforcement led to a higher percentage of moderate to strong effect but was also likely an artifact of coincidentally having been primarily employed with younger participants. Thus, highlighting the difficulty with drawing conclusions.

Drawing conclusions from the present literature is premature given the variability and overlap of procedures across studies. Nonetheless, these preliminary data can aid in the comparison of current treatment options by allowing the future SSP treatments to be intentionally designed. Shillingsburg et al. (2015) also made several recommendations:

(1) include comprehensive characterization of participants, (2) include high quality measures for assessing and diagnosing participants, (3) consider using a diagnostic assessment battery to confirm or rule out an ASD diagnosis, (4) assess sounds that are both novel and presently in the repertoire, (5) assess the number of experimenter-emitted sounds per pairing as a treatment variable, (6) specify the type of pairing procedure used, (7) specify the rationale for its use, (8) consider assessing the type of pairing being used, (9) record and report the number of pairings per minute conducted, (10) conduct a brief preference assessment immediately prior to SSP sessions to increase the effectiveness of pairing, and (11) include information regarding the quality of sound production during pairing trials. Many of these recommendations were considered in the design of the present studies.

Purpose of Study

The primary goal of the studies was to evaluate the efficiency and efficacy of SSP and ET procedures, relative to each other, in terms of their effect on producing vocalizations across sessions for the development and onset of verbal behavior in early learners with ASD. The purpose of evaluating the efficacy of these procedures is to improve the overall efficiency of early interventions. If clinicians can quickly identify learner characteristics that predict treatment efficacy for a set of procedures, then selecting the best intervention based on characteristics will lead to efficient sequencing of procedural options. Therefore, a secondary objective was to identify participant traits that may predict the ideal treatment and strategies to be implemented for a particular learner, given that the efficacy of a procedure may vary depending on the development of the

individual. Identifying predictive participant traits for the best treatment was done by examining the individual treatment effects for each participant relative to descriptive assessments conducted prior to intervention. A third and final objective of the proposed studies was to develop standards for determining how to begin interventions and when to pivot from one procedure to a more efficient one.

Significance of Study

A substantial amount of research indicates that early intervention in the development of language has profound implications for learners. In general, the earlier the intervention, the better the outcomes (Harris & Handleman, 2000). However, the quality of the intervention and the strategies implemented during these early interventions are likely to influence the degree or size of effects. Much of these early intervention strategies have yet to be investigated thoroughly. As a result, practitioners often default to known methods and procedures which yield desired results at some point in development but have not been examined in the context of developmental appropriateness and how there might be more appropriate alternatives. The present studies fill in research gaps and identify differences in strategy or sequence with the best known probability of producing strong effects. Given the fact that 1 in 68 children is currently diagnosed (Center for Disease Control, 2014) with ASD, the significance of the findings have the potential to impact a significant population and have very desirable long-term effects for their families and communities. Some of these long-term impacts could include improved communication, independence, and overall functioning of those receiving improved treatments.

II. Method

Assessments

A total of five assessments (described in the following section) were administered during the initial course of the studies. Three assessments (MSEL, ASRS, & BLAF) were intended to describe the participant characteristics or assess their current repertoire skill set. Describing and assessing participants were recommended (Shillingsburg et al., 2015) steps to improve an understanding of intervention effects. The fourth and fifth assessment (RAISD & SPA) were used to identify potential reinforcers to use throughout the course of the studies.

The Mullen Scales of Early Learning. The *Mullen Scales of Early Learning* (MSEL; Mullen, 1995) assessment was used to assess each participant's level of functioning in different skill set areas. The MSEL evaluates readiness for school from infants up to the age of 68 months as well as their developmental progress. The MSEL provides a general measure of development and skill based on five scales: Gross Motor, Visual Reception (i.e. visual discrimination), Fine Motor, Expressive Language, and Receptive Language. The five domains can be used to describe a T-score, percentile, age equivalent score, and an overall Early Learning Composite score.

Autism Spectrum Rating Scales. Another assessment used in the present studies to describe participants and confirm the likelihood of ASD was the *Autism Spectrum Rating Scales* (ASRS; Goldstein & Naglieri, 2010). The ASRS is a parent or teacher completed assessment comprised of Likert-rating scales. The ASRS is intended to identify symptoms, behaviors, and associated characteristics of ASD. It is a standardized

norm-referenced assessment for individuals 2 to 18 years old, that can be used to guide treatment decisions and treatment effectiveness. The ASRS evaluates peer socialization, adult socialization, social and emotional reciprocity, atypical language, stereotypy, behavioral rigidity, sensory sensitivity, attention and self-regulation, and attention. There are two versions of the ASRS. One version is for children 2 to 5 and one for children 6 to 18 years of age. The version used in the present study was the one intended for children 2 to 5. Furthermore, the ASRS includes a prorated version for individuals who do not speak or speak infrequently. Since all the participants in the present studies did not speak or spoke infrequently, the prorated scoring method was used. The ASRS Social/Communication domain rating indicates the extent to which the child uses verbal and non-verbal communication appropriately to initiate, engage in, and maintain social contact. The Unusual Behaviors domain rating indicates the child's level of tolerance for changes in routine, engagement in apparently purposeless and stereotypical behaviors, and overreaction to certain sensory experiences. The child's ratings on the Total Score scale indicate the extent to which the child's behavioral characteristics are similar to the behaviors of children diagnosed with ASD. The child's ratings on the DSM-V Scale indicate how closely the child's symptoms match the DSM-V criteria for ASD (Goldstein & Naglieri, 2010). It is important to note that the ASRS was completed by the parents who often have difficulty identifying red flags for autism or atypical behavior for very young children.

Behavioral Language Assessment Form. The *Behavioral Languages Assessment Form* (BLAF; Sundberg & Partington, 1998) is a questionnaire form used to describe participants across 12 verbal behavior skills. The skill domains addressed in the

questionnaire include: cooperation with adults, requests (mands), motor imitation, vocal play, vocal imitation, matching to sample, receptive (i.e. listener skills), labeling (tacts), receptive skills by function, feature, and class, conversation skills (intraverbals), letters and numbers, and social interaction. The BLAF was completed by treating therapists that had several weeks of experience working with the participants.

Reinforcement Assessment for Individuals with Severe Disabilities. A

structured reinforcer assessment survey, *Reinforcement Assessment for Individuals with Severe Disabilities* (RAISD; Fisher, Piazza, Bowman, & Amari, 1996), was also conducted prior to beginning the study to identify potential edible reinforcers. At least three potential edible reinforcers were identified for subsequent use.

Stimulus Preference Assessment. Prior to beginning any session, a *stimulus preference assessment* (Higbee, Carr, & Harrison, 2000) was also conducted. The stimulus preference assessment consists of a single array of three to five items, identified in reinforcer assessment survey, being placed in front of the participant. The first item the participant reached or pointed to was selected as the stimulus to be used in the SSP or ET procedure. However, there was one participant (Jordan) which did not receive edible reinforcers throughout the course of the study. Jordan did not show preference for any edible reinforcers consistently. When Jordan did show preference for an edible reinforcer, it was often only for one to three trials and he would no longer consume any edible. Therapists working with Jordan throughout the course of the day also reported Jordan's reinforcers generally did not last more than a few trials. On the basis of previous studies,

consumption times, and therapist recommendations, tickles were used exclusively as a reinforcer for Jordan throughout all sessions.

Participants

Participants in the first study consisted of five early learners, between the ages of 26 months to 55 months. All the participants had a community diagnosis of ASD. All participants had limited or no verbal behavior repertoires. Participants were recruited through the University of Miami Center for Autism and Related Disabilities (UM-CARD) Intensive Behavior Intervention Services (IBIS) clinic. Three of the participants that completed the first study also participated in the second study.

Jake. Jake was a 2-year-old boy diagnosed with autism spectrum disorder. He was participating in UM's Intensive Behavioral Intervention Services program when he was recruited to participate in the study. Jake was described as cooperative with adults for only one brief and easy response for a powerful reinforcer. With requests and mands, he would typically pull people, point, or stand by reinforcing items. He was unable to imitate anybody's motor movements. With vocal play, Jake was described as making a few speech sounds at a low rate. In vocal imitation, he would repeat a few specific sounds or words. With matching to sample, the participate could match one or two objects or pictures to a sample. Receptively, he would follow a few instructions related to daily routines. Jake could not identify any items or actions and could not identify items based on information about them. He could not identify letters, numbers or written words. Finally, he could not fill-in missing words or parts of songs and would not initiate interactions with others.

Jake's MSEL results demonstrated that although he was chronologically over two years of age, his repertoire of skills was delayed significantly. His receptive and expressive language domain scores indicated developmental age equivalence of less than a year old for both language domains. Jake's MSEL results are summarized in table 2. Jake's ASRS results are summarized in table 2. Jake's ASRS scores indicated a very elevated difficulty with communicating appropriately to initiate, engage in, and maintain social contact. The ASRS ratings did not indicate any difficulties with unusual behavior. Jake's ratings in the Total domain indicated a slightly elevated extent to which the child's behavioral characteristics are similar to the behaviors of children diagnosed with ASD and the DSM-V domain results indicated an elevated symptom match to the DSM-V criteria for ASD.

Max. Max was a 2-year-old boy diagnosed with autism spectrum disorder. He was enrolled in UM's Intensive Behavioral Intervention Services program when he was recruited to participate in the study. Max was described as cooperative with adults for only one brief and easy response for a powerful reinforcer. With requests and mands, he typically used one to five words, signs or pictures to ask for reinforcers. He was described as able to imitate a few gross motor movements modeled by others. With vocal play, Max was described as making a few speech sounds at a low rate. In vocal imitation, he would repeat a few specific sounds or words. Receptively, he would follow a few instructions to do actions or touch items. Max could label only one to five items but could not identify items based on information about them. He could not identify letters, numbers or written words. Finally, he could fill-in a few missing words and provides animal sounds but would not initiate interactions with others.

Table 2

Participant MSEL Results

Participant	Chronological Age (months)	Domain	Age Equivalence	Descriptive Category	Percentile Rank
Jake	26	Visual Reception	16 months	Very Low	1
		Fine Motor	13 months	Very Low	1
		Receptive Language	5 months	Very Low	1
		Expressive Language	10 months	Very Low	1
Max	31	Visual Reception	20 months	Very Low	1
		Fine Motor	22 months	Very Low	1
		Receptive Language	14 months	Very Low	1
		Expressive Language	9 months	Very Low	1
Lane	28	Visual Reception	18 months	Very Low	1
		Fine Motor	18 months	Very Low	1
		Receptive Language	7 months	Very Low	1
		Expressive Language	7 months	Very Low	1
Jordan	34	Visual Reception	5 months	Very Low	1
		Fine Motor	16 months	Very Low	1
		Receptive Language	10 months	Very Low	1
		Expressive Language	3 months	Very Low	1
Dante	55	Visual Reception	21 months	Very Low	1
		Fine Motor	22 months	Very Low	1
		Receptive Language	14 months	Very Low	1
		Expressive Language	12 months	Very Low	1

Max's MSEL results demonstrated that although he was chronologically over two years of age, his repertoire of skills was delayed significantly. His receptive and expressive language domain scores indicated developmental age equivalence of less than 15 months old for both language domains. Max's MSEL results are summarized in table 2. Max's ASRS results are summarized in table 3. Max's ASRS scores indicated an elevated difficulty with communicating appropriately to initiate, engage in, and maintain social contact. The ASRS ratings also indicated very elevated difficulties with unusual behavior. Max's ratings in the Total domain indicated a very elevated extent to which the Max's behavioral characteristics are similar to the behaviors of children diagnosed with ASD and the DSM-V domain results indicated a very elevated symptom match to the DSM-V criteria for ASD.

Lane. Lane was a 2-year-old boy diagnosed with autism spectrum disorder. He was recruited to participate in the study from UM's Intensive Behavioral Intervention Services program. Lane was described as always uncooperative with adults, avoided work and engaged in negative behavior. With requests and mands, he typically pulled people, pointed or stood by reinforcing items. He could imitate several gross motor movements on request. With vocal play, the participant vocalized frequently with varied intonation and said a few words. In vocal imitation, he would repeat or closely imitate several sounds or words. With matching to sample, the participant can match one or two objects or pictures to a sample. Receptively, he would follow a few instructions to do actions or touch items. Lane could also label only one to five items but could not identify items based on information about them. He could not identify letters, numbers or written

words. He could fill-in a few missing words and provide animal sounds. He would also physically approach others to initiate interactions.

Lane's MSEL results demonstrated that although he was chronologically over two years of age, but his repertoire of skills was delayed significantly. His receptive and expressive language domain scores indicated developmental age equivalence of less than a year old for both language domains. Lane's MSEL results are summarized in table 2. Lane's ASRS results are summarized in table 3. Lane's ASRS scores indicated a slightly elevated difficulty with communicating appropriately to initiate, engage in, and maintain social contact. The ASRS ratings did not indicate any difficulties with unusual behavior. Lane's ratings in the Total and DSM-V domain indicated Lane did not have behavioral characteristics similar to the behaviors of children diagnosed with ASD and did not match the DSM-V criteria for ASD.

Jordan. Jordan was a 2-year-old boy diagnosed with autism spectrum disorder. He was recruited to participate in the study from UM's Intensive Behavioral Intervention Services program. Jordan was described as cooperative with adults for only one brief and easy response for a powerful reinforcer. With requests and mands, he typically pulled people, pointed or stood by reinforcing items. He was described as able to imitate few gross motor movements modeled by others. With vocal play, Jordan was described as making a few speech sounds at a low rate. In vocal imitation, he was described as unable to repeat any sounds or words. Jordan could not match any objects or pictures to a sample. Receptively, he would follow a few instructions related to daily routines. Jordan could not identify any items or actions and could not identify items based on information

about them. He could not identify letters, numbers or written words. Finally, he could not fill-in missing words or parts of songs and would not initiate interactions with others.

Jordan's MSEL results demonstrated that although he was chronologically over two years of age, his repertoire of skills was delayed significantly. His receptive and expressive language domain scores indicated developmental age equivalence of less than a year old for both language domains. Jordan's MSEL results are summarized in table 2. Jordan's ASRS results are summarized in table 3. Jordan's ASRS scores indicated a very elevated difficulty with communicating appropriately to initiate, engage in, and maintain social contact. The ASRS ratings also indicated very elevated difficulties with unusual behavior. Jordan's ratings in the Total domain indicated a very elevated extent to which the Jordan's behavioral characteristics are similar to the behaviors of children diagnosed with ASD and the DSM-V domain results indicated a very elevated symptom match to the DSM-V criteria for ASD.

Dante. Dante was a 4-year-old boy who was also diagnosed with autism spectrum disorder. He was enrolled in UM's Intensive Behavioral Intervention Services program when he was recruited to participate in the study. Dante was described as always uncooperative with adults, avoided work and engaged in negative behavior. With requests and mands, he typically pulled people, pointed or stood by reinforcing items. He was described as able to imitate few gross motor movements modeled by others. With vocal play, Dante was described as making a few speech sounds at a low rate. In vocal imitation, he was described as unable to repeat any sounds or words. With matching to sample, Dante could match one or two objects or pictures to a sample. Receptively, he would follow a few instructions related to daily routines. Dante could not identify any

items or actions and could not identify items based on information about them. He could not identify letters, numbers or written words. Finally, he could not fill-in missing words or parts of songs and would not initiate interactions with others.

Dante's MSEL results demonstrated that although he was chronologically over four years of age, his repertoire of skills was delayed significantly. His receptive and expressive language domain scores indicated developmental age equivalence of about a one year old for both language domains. Dante's MSEL results are summarized in table 2. Dante's ASRS results are summarized in table 3. Dante's ASRS scores indicated a very elevated difficulty with communicating appropriately to initiate, engage in, and maintain social contact. The ASRS ratings did not indicate any difficulties with unusual behavior. Dante's ratings in the Total domain indicated a very elevated extent to which the child's behavioral characteristics are similar to the behaviors of children diagnosed with ASD and the DSM-V domain results indicated an elevated symptom match to the DSM-V criteria for ASD.

Setting and Materials

The studies were conducted at the University of Miami's Intensive Behavioral Intervention Services clinic. The room was rectangular with eight work stations located throughout the room. One station was a social play area with a large carpet and five toy shelves located adjacent to the station. Seven other stations located throughout the room were composed of a small child-sized chair and at least one toy shelf with different toys placed throughout the shelves. All sessions were completed at a specific station with a table and one toy shelf. The experimenter sat adjacent to the table and to the right of the

participant on a small cube chair. A large iPad with a large timer was placed on the table, out of reach from the child. The timer did not make any sounds and served only to cue the experimenter's presentation of pairings or echoic training trials. A second experimenter sat behind the child to prevent the child from leaving the station and to collect fidelity and IOA data. The participant was given free access to any toys located at the adjacent shelf. The parents of the participant were not generally present during the sessions but had the option to observe from behind a two-way mirror as was typical during normal operations. Sessions were completed during the normal class-room schedule. Chosen reinforcers were placed at the back of the table or in the experimenter's lap, out of reach from the participant. It is important to note that the term reinforcer is used throughout the entire study since a stimulus that functions as a reinforcer for behavior during certain conditions will likely serve as a reinforcer during other conditions. It is assumed that since the stimulus has functioned as a reinforcer for a particular behavior, it will serve to increase the value of a vocalization through the use of a pairing procedure. Despite its actual function in a given moment or procedure, the term reinforcer is used for the sake of clarity.

Table 3

Autism Spectrum Rating Scale Results

Participant	Percentile Rank			Classification		
	Social Communication	Unusual Behaviors	DSM-V	Social Communication	Unusual Behaviors	DSM-V
Jake	99	42	96	Very Elevated	Average	Elevated
Max	96	98	99	Elevated	Very Elevated	Very Elevated
Lane	90	46	82	Slightly Elevated	Average	Average
Jordan	98	98	99	Very Elevated	Very Elevated	Very Elevated
Dante	98	73	93	Very Elevated	Average	Elevated

Design

There were two studies conducted. Both studies were a within-subject controlled ABAB designs buttressed within a multiple baseline across behaviors design (Richards, Taylor, Ramasamy, & Richards, 1999). The design was used to demonstrate replicability within and across participants. Study 1 and 2 both compared two treatments across two behaviors per participant to determine their relative efficacy in regard to the frequency of vocalizations produced across sessions.

Procedures

Study 1

The independent variables (treatments) were the stimulus-stimulus pairing procedure (SSP) and echoic training procedure (ET). Participants were exposed to an equal number of SSP trials as ET trials (i.e., 1:1 ratio). The dependent variable was the frequency of vocalizations, of the target vocalization being trained, that occurred prior to (pre-session), during (in-treatment), and following (post-session) treatment. There were three phases during the course of the study: baseline (A), control (A'), and treatment (B). All three phases were composed of pre-session and post-session observations. Two sessions per day were conducted, with sessions occurring Monday through Thursday. At least 35% of all sessions had an independent observer collect inter-observer agreement (IOA) and fidelity data to ensure treatment fidelity. Participants completed two sessions per day unless there were logistical reasons that prevented the experimenters from completing sessions (e.g. a child being checked out of the program early). Sessions were

split by Treatment type (SSP vs Echoic) and were randomly assigned as the first or second session for each participant to control for any carry-over effects. In addition, a minimum of 30 minutes between sessions was also integrated into the treatment. In some cases, due to logistical constraints, such as some of the participants attending the program for only half the day, these breaks between sessions could not be increased.

Pre-session and post-session observations. Observations spanned five minutes. Each observation occurred immediately before beginning each session (pre-session) regardless of the phase and immediately following each session (post-session) regardless of the phase. During these observations, participants were given free access to toys at an adjacent shelf but were required to stay seated in front of a small table. Interaction between experimenter and participant was kept at a minimum. The experimenter would only interact with the participants to keep them seated in the work space. Session times varied but were generally less than 20 minutes in length, with exception of the control phase which generally required 23 to 25 minutes.

Baseline. Baseline sessions were identical to pre-session and post-session observations. The purpose of this condition was to document the frequency of vocalizations prior to any treatment intervention.

Control. To control for adventitious reinforcement, as was done in Miguel, Carr, and Michael (2002), two controls were implemented. First, a control phase (A`) followed the baseline phase (A). During the control phase, the experimenter repeated the target sound three times, waited 20 seconds, and then delivered a reinforcer. The target sound emitted by the experimenter had a slightly exaggerated prosodic pattern (motherese; Falk,

2004). If the target vocalization was emitted by the participant at any time, the delivery of the reinforcer was delayed 15 to 20 seconds. After the participant was given 20 seconds to consume the reinforcer, a new trial was presented. This control condition was designed to determine whether modeling and/or emission of sounds with the delivery of preferred items separated in time was sufficient to increase vocal behavior. The control phase consisted of 20 reinforcer deliveries. The second control implementation occurred strictly during the SSP treatment, which involved withholding reinforcement in the event of a vocalization during treatment.

Treatment. Both treatments, SSP and ET, had a total of 20 trials presented per session for each target vocalization. Each treatment target vocalization was compared to the other corresponding treatment vocalization. Target sounds were equitable with regard to difficulty. For instance, target sounds had an equal number of syllables and were chosen from a list of targets. Target sounds that both the parent and treating therapist had not observed were selected as targets. Target sounds were also determined by commensuration of the learner's age and typical speech-sound development outlined by Shriberg (1993). Shriberg described the Early-8 (e.g., m, b, y, n, w, d, p, h) sounds as developing from 18 months to 36 months of age, whereas the Middle-8 (e.g., t, ng, k, g, f, v, ch, j) develop from two to six-years of age. Depending on overall performance for the Early-8 that had previously been observed, Middle-8 targets were considered. One target sound per treatment was selected.

Stimulus-stimulus pairing. During the SSP phase, SSP sessions began immediately after pre-session observations. During a pairing trial, the experimenter

would repeat the target vocalization three times over a period of three to four seconds and pair the third vocalization with a reinforcer by delivering them simultaneously. The experimenter emitted slightly exaggerated prosodic sound patterns to increase the likelihood that vocalizations during pairing were more salient than vocalizations outside of treatment. There were approximately two to three pairings per minute in the SSP treatment. Pairings were presented every 20 seconds. To control for adventitious reinforcement, if the target vocalization occurred following the experimenter's emission, the experimenter withheld delivery of the reinforcer or pairing trial for 15 to 20 seconds.

Echoic training. During the ET phase, ET sessions began immediately after a pre-session observation. The experimenter presented the discriminating stimulus “say” and then the target vocalization three times (e.g. “say we, we, we”). The target vocalization was repeated three times to control for the number of experimenter emitted vocalizations and make the ET treatment commensurate with the SSP treatment. Like the control phase and SSP treatment, during ET the target vocalizations emitted by the experimenter had a slightly exaggerated prosodic pattern to increase its salience relative to speech outside of the treatment. There were approximately two to three training trials per minute in the ET treatment. During echoic training, reinforcers were delivered only when the participant repeated the target sound presented by the experimenter within four seconds of presentation of the discriminating stimulus. If no target sound was repeated by the participant, reinforcers were withheld and a discriminating stimulus was not presented until the next trial. Trials were presented every 20 seconds. See Figure 1 for a visual depiction of the session observation times.



Figure 1. Pre-session, Treatment, and Post-session observations.

Reliability. For behavioral data, reliability was calculated by determining interobserver agreement (IOA; Cooper et al., 2007). To ensure the accuracy of data collection, interobserver agreement was calculated for each session by having two observers record the number of occurrences of the target vocalization per 1-min interval during treatment and during each of the 5-min observations. The number of intervals with exact agreement was then divided by the number of total intervals and multiplied by 100 to give a percentage of agreement. The average IOA across sessions is reported as the final IOA per participant. For study 1, IOA was calculated for 46.0% of Jake’s sessions, 51.6% of Max’s sessions, 79.7% of Lane’s sessions, 83.1% of Jordan’s sessions, and 75.4% of Dante’s sessions. Both observers independently recorded their observations during the sessions. Calculations were made after sessions ended. Disagreements greater than 90% resulted in re-training of the observer, clarification of inconsistencies, or changes to observation conditions. For instance, some observations resulted in errors due to a loud treatment environment. In subsequent sessions, observers were instructed to sit closer to the participant or position themselves more ideally. Mean interobserver agreement for Jake and Jordan was 100% for both. Mean interobserver agreement for Max was 98.7% (range, 82.3-100.0). Mean interobserver agreement for Lane was 93.1%

(range, 66.6-100.0). Lastly, mean interobserver agreement for Dante was 97.7% (range, 82.3-100.0). For study 2, IOA was calculated for 85.0% of Max's sessions, 94.8% of Dante's sessions, and 97.4% of Lane's sessions. Mean interobserver agreement was 96.7% for Max's sessions, 97.8% of Lane's sessions, and 99.0% for Dante's sessions.

Treatment fidelity. Treatment fidelity refers to the degree to which an experimenter correctly implements an experimental procedure. For all treatment conditions, the primary experimenter completed a fidelity checklist. An independent observer recorded whether each step was carried out correctly. For the control condition, observers checked off whether the experimenter correctly (1) emitted the target sound after the session start and every appropriate 20 second interval (2) delivered the reinforcer within four seconds after the 20 second reinforcer delay had elapsed, and (3) if the target vocalization was emitted by the participant, the reinforcer was not delivered for at least 15 to 20 seconds since the target vocalization's last occurrence. For SSP, the observer indicated whether the experimenter correctly (1) presented the target vocal sound within four seconds of every 20 seconds of elapsed time, (2) paired the reinforcer (i.e. delivered within two seconds) with the experimenter's third vocalization of the target sound, and (3) withheld a pairing trial for 20 seconds if the participant emitted the target response. For ET, the observer recorded whether the experimenter correctly (1) presented the correct ET discriminating stimulus within four seconds of every 20 seconds of elapsed time, (2) reinforced the participant's emission of a target vocalization within 4 seconds, and (3) presented a new S^D at the subsequent and appropriate 20 second interval. For all three conditions, if there were any reasons to skip an interval (e.g. distracted

participant), the experimenter and observer would both have recorded having done so and agreement per interval was compared after the session was completed.

For the control, SSP, and ET conditions, each of the three steps is considered one interval. If all the steps were completed correctly, the interval is scored as correct. If even one step was performed incorrectly, the interval was marked as incorrect. Checklists between the experimenter and observer were compared following the session and a treatment fidelity percentage was calculated. The fidelity percentage was calculated by totaling the number of agreements and dividing by the total number of intervals and multiplying by 100. Treatment fidelity was calculated above 50% of all sessions, across all phases, of each participant.

In study 1, Jake's treatment fidelity was assessed for 74.4% of all treatment sessions and averaged 99.6% (range, 90.0-100.0). For Max, treatment fidelity was assessed for 83.3% of treatment sessions and averaged 99.7% (range, 95.2-100.0). For Lane, treatment fidelity was assessed for 100% of treatment sessions and averaged 98.8%. For Jordan, treatment fidelity was assessed for 92.3% of treatment sessions and averaged 99.4% (range, 95.2-100.0). Lastly, Dante's treatment fidelity was assessed for 95.1% of treatment sessions averaging 98.9% (range, 80.0-100.0). For study 2, Max's treatment fidelity was assessed for 100.0% of treatment sessions and averaged 99.3% (range, 95.2-100.0). For Lane, treatment fidelity was assessed for 95.0% of treatment sessions and averaged 99.2% (range, 95.0-100). Dante's treatment fidelity was assessed for 95.0% of treatment sessions averaging 94.1% (range, 0-100.0).

Problem behavior. Prior to a session, if the participant was engaging in any problem behavior, the session was delayed until such behavior subsided. During experimental sessions, if the participant attempted to leave the work area, the participant was not allowed to leave and was redirected to his seat. If the behavior was much more disruptive (e.g., tantrums or potty accidents) to the experimental session, the experimenter temporarily suspend the session. To increase the probability that the participant was attending to the experimenter's vocalizations, the experimenter did not present any trials in which the participant was not facing the experimenter. The delay procedure was only needed once during the course of the study for participant Lane.

Study 2

The purpose of Study 2 was to determine whether a combination of the two treatments was more effective, with regards to frequency of vocalizations across sessions, for each learner, compared to echoic training alone. Since Study 1 controlled for adventitious reinforcement, such an analysis was not possible in Study 1. For reinforcement to occur, a behavior must occur, followed by a reinforcing consequence. When controlling for adventitious reinforcement during an SSP procedure, one ensures that an increase in response frequency is due to the effect of SSP and not reinforcement following the response. Therefore, Study 2 was necessary to determine whether a combination of treatments could produce more desirable effects versus just using echoic training alone. In applied settings, practitioners are less concerned with elucidating variables that affect treatment efficacy than they are with maximizing treatment effects. Stated differently, they are less concerned with comparing treatments than they are with

modifying an intervention for maximum effect. Clinicians often do not have the time or resources to compare treatment variable effects. Therefore, the goal of Study 2 was to determine if clinicians should directly reinforce vocalizations evoked via SSP, rather than attempting to use SSP in isolation. If SSP with direct reinforcement was shown to be more effective than SSP alone, clinicians could use SSP to “jump-start” the vocalization process and then strengthen it with direct reinforcement.

There were two major differences between Study 1 and Study 2. First, the SSP procedure in Study 2 did not control for adventitious reinforcement. Instead, Study 2 was intended to evaluate the combined effects of an SSP procedure with typical reinforcement procedures. If the target vocalization occurred at any time during the SSP treatment, the experimenter would immediately deliver a reinforcer contingent on the response. In addition to the modification of the SSP procedure, there was no control phase during Study 2 since control levels of responding were demonstrated during Study 1. For Study 2, new target vocalizations were targeted. Like Study 1, two target vocalizations underwent treatment. All other procedures remain unchanged from Study 1.

III. Results

The studies aimed to: (a) compare treatment outcomes; (b) identify participant traits that predict ideal outcomes; (c) identify unique developmental patterns of each individual that may contribute to outcomes; and (d) provide recommendations and standards for determining appropriate intervention as to when to pivot to other procedures if necessary. To evaluate treatment effectiveness and compare treatment outcomes, two types of analyses were used. As is customary in SCRM, visual analysis was used.

Second, Non-overlap of All Pairs (NAP; Parker & Vannest, 2009), a non-parametric overlap index of data between phases in SCRM was used to evaluate treatment effect sizes. These two analyses were used in combination with participant assessment and evaluation data to achieve the remaining aims of the studies.

Study 1

Visual Analysis

Visual analysis of data using criteria such as differences in level and variations in trend or slope are generally carried out when continuous data are gathered, data are graphically represented, and the researcher makes formative and summative analyses of a study's outcome (Hurtado-Parrado & López-López, 2015; Richards et. al., 1999). Since the study's aims were to compare treatment outcomes and discern predictors through developmental factors such as participant traits and context, in addition to developing standards for treatment optimization, a visual analysis and corresponding interpretations were necessary. Generally, when using visual analysis, two overall aspects of the data are analyzed: level and trend (Richards et. al., 1998). To analyze these aspects, certain prior steps were completed. First, legends, axes, and all phases were clearly labeled. Second, the scaling of the y-axes were adjusted to an appropriate range in order for changes in performance to appear to commensurate with their significance. For instance, a life-threatening behavior occurring once should appear significant on a graph despite only one occurrence. A change in performance on a math test of only a few percentage points in accuracy should not represent such a significant change (Richards et. al., 1999). Given

that the present study evaluated participant's performance of vocalizations based on two treatments, the axes remained equal for both treatments for each participant. Furthermore, the number of data points within a phase had to be sufficient to make a reasonable determination of the data path or level for a given treatment. When there is very little variability in performance (e.g. flat path or a clear increasing or decreasing trend) fewer data points are necessary. For instance, Jake and Jordan demonstrated zero or near zero vocalizations of their respective target sounds (i.e. zero occurrences of the response). As a result, fewer sessions were necessary compared to Dante which had a variable level of responding initially for the control phase. However, with Dante there was a sudden increase after treatment, the intervention clearly depicts a functional relationship, and demonstration of that functional relationship was strengthened since there was replication in reversal phases to baseline and intervention.

Jake. Jake's results for study 1 are depicted in Figure 2. In Figure 2, for target vocalization Chip and Hop, baseline, control, and treatment were all at a frequency of zero for both treatments of SSP and ET. There was only one session in which there were 3 occurrences of the vocalization Chip. Throughout the study, there were never any occurrences of the word Hop.

Max. Max's results for study 1 are depicted in Figure 3. In Figure 3, for target vocalization Up (i.e. SSP treatment), initial baseline and control were all at a frequency of zero. The first transition into SSP treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total

of 36 occurrences (range, 0-11 per session) across eight sessions. During pre-session and post-observations, frequency of vocalizations remained at baseline levels. One notable exception was session 11 which resulted in three vocalizations during pre-observation.

Jake's Frequency of Vocalizations (Study 1)

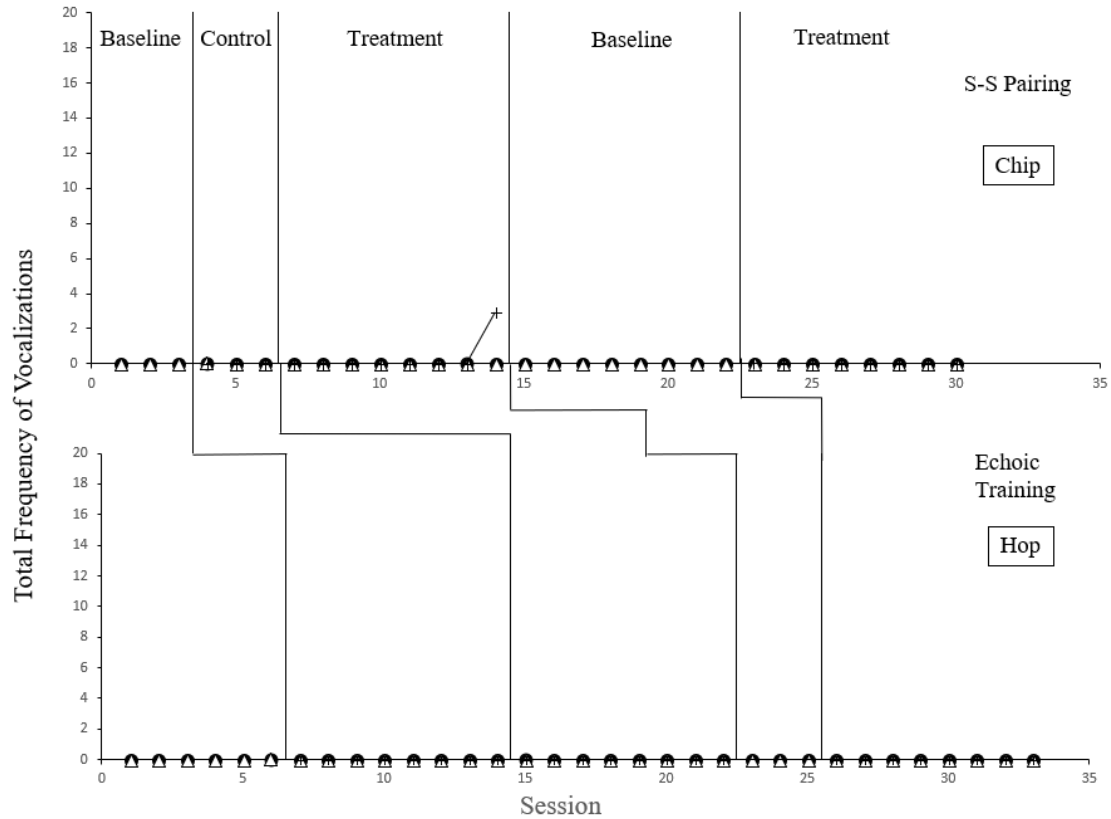


Figure 2. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Frequency of vocalizations for Up were in an upward trend when a reversal to baseline occurred. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for SSP which resulted in a sudden initial increase in the frequency of vocalizations but eventually returned and ended on baseline levels of frequency.

Max's Frequency of Vocalizations (Study 1)

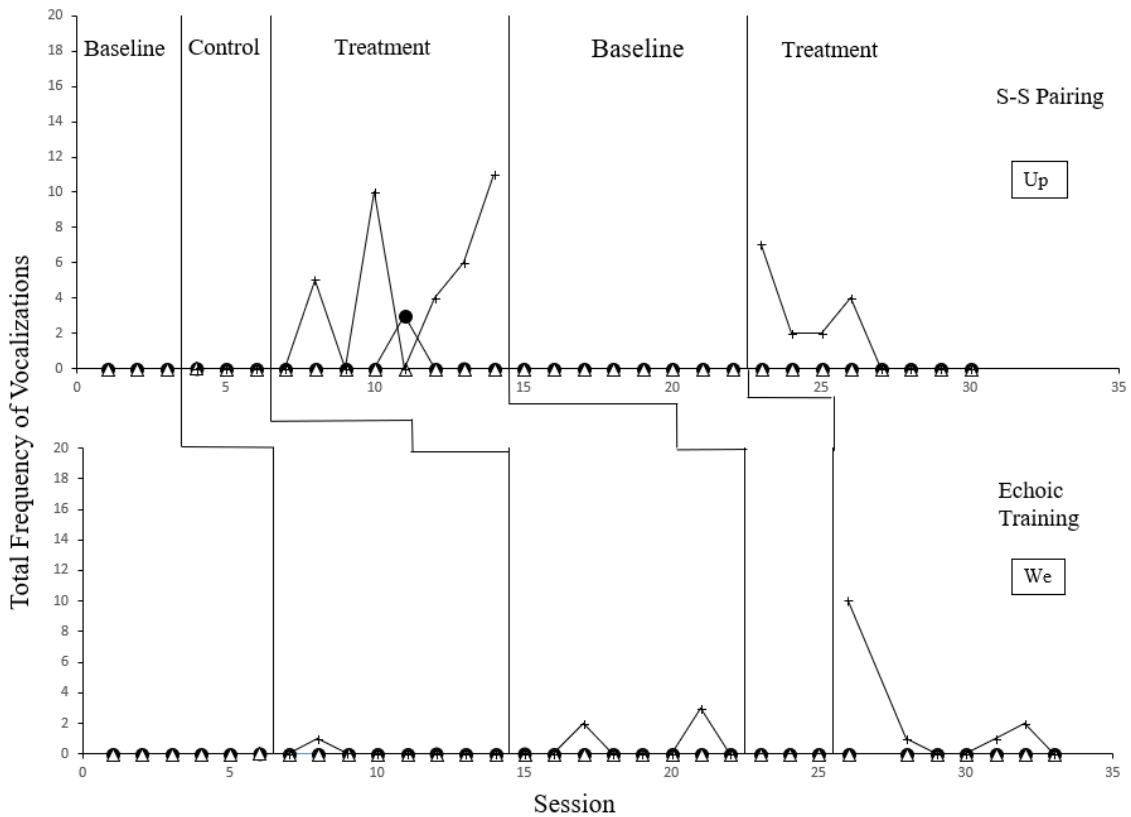


Figure 3. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

The total frequency of vocalizations during the reversal to SSP treatment was 15 (range, 0-7 per session) across eight sessions.

As shown in Figure 3, for target vocalization “We” (i.e. ET treatment), initial baseline and control were all at a frequency of zero, with exception to session 8, which had one occurrence during treatment. The first transition into ET treatment resulted in a minimal increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of five occurrences (range, 0-3 per session) across eight sessions. Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout treatment. Frequency of vocalizations for We were at a level of zero when a reversal to baseline occurred. Frequency of vocalizations maintained at levels of zero frequency across sessions during the reversal. Next, a reversal to treatment was implemented for ET which resulted in a sudden initial increase in the frequency of vocalizations but eventually returned and ended on baseline levels of frequency. The total frequency of vocalizations during the reversal to ET treatment was 14 (range, 0-10 per session) across seven sessions. Overall, the SSP treatment produced vocalizations above zero level one session earlier than ET and on 10 out of 14 sessions, whereas ET produced vocalizations on 6 out of 14 sessions.

Lane. Lane’s results for study 1 are depicted in Figure 4. In Figure 4, for target vocalization Potty (i.e. ET treatment), initial baseline was at a frequency of zero during pre-session and post-session observation. The implementation of the control phase resulted in a sudden increase in the frequency of vocalizing Potty during treatment but

had no effect on pre-session or post-session levels. Total vocalizations during the control phase was 53 (range 3-27 per session) across four sessions. The control phase was in a downward trend when ET treatment was implemented. The first transition into ET treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of 129 occurrences (range, 6-22 per session) across eight sessions. During pre-session and post-observations, frequency of vocalizations remained at baseline levels. Frequency of vocalizations for Potty were in an upward trend when a reversal to baseline occurred. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for ET which resulted in a sudden initial increase in the frequency of vocalizations during treatment but had no effect on pre-session or post-session observation levels. ET treatment reversal remained level with prior treatment levels. The total frequency of vocalizations during the reversal to ET treatment was 76 (range, 8-20 per session) across five sessions. As shown in Figure 4, for target vocalization Taco (i.e. SSP treatment), initial baseline was at a frequency of zero. A control phase was implemented after baseline. During treatment, frequency of vocalizations increased significantly for a total of 121 occurrences (range, 2-36 per session) across eight sessions. Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout the control phase. Frequency of vocalizations for Taco were in a downward trend when treatment was implemented.

Lane's Frequency of Vocalizations (Study 1)

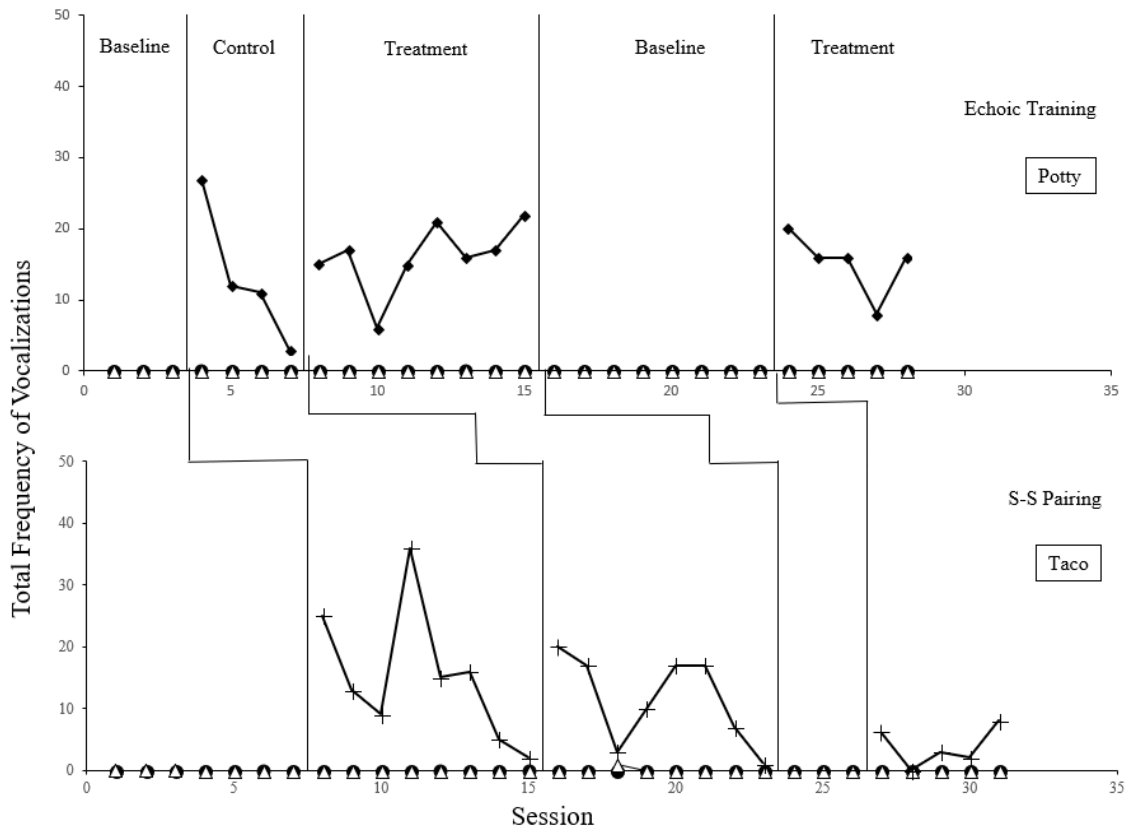


Figure 4. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (diamond and plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

In the treatment phase, frequency of vocalizations maintained at levels of zero frequency during pre-session and post-session observations. However, during treatment, frequency levels recovered. During treatment, frequency of vocalizations totaled 92 occurrences (range, 1-20 per session) across eight sessions. Next, a reversal to baseline was implemented for SSP. No vocalizations occurred during reversal to baseline which resulted in a sudden initial increase in the frequency of vocalizations but eventually

returned and ended on baseline levels of frequency. Lastly, a reversal to the treatment phase was implemented. The total frequency of vocalizations during the reversal to SSP treatment was 19 (range, 0-8 per session) across seven sessions. Overall, the ET treatment produced vocalizations above zero level on the first session of both ET and SSP. For ET vocalizations were produced in 13 out of 13 sessions, whereas SSP produced vocalizations on 12 out of 13 sessions.

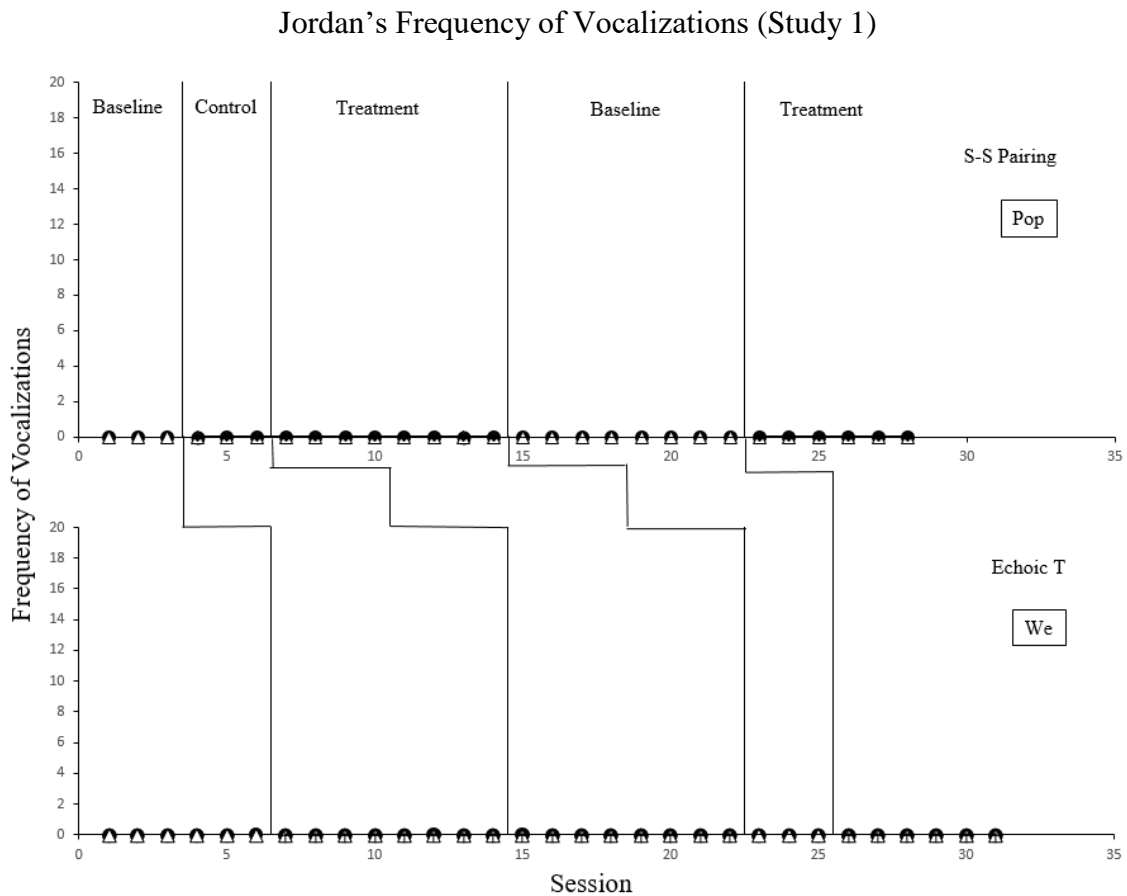


Figure 5. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (diamond and plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Jordan. Jordan's results for study 1 are depicted in Figure 5. In Figure 5, for target vocalization Pop and We, baseline, control, and treatment were all at a frequency of zero for both treatments of SSP and ET. There were no sessions in which any vocalizations occurred for either pre-session, post-session, or during treatment.

Dante. Dante's results for study 1 are depicted in Figure 6. In Figure 6, for target vocalization We (i.e. ET treatment), initial baseline was at a frequency of zero during pre-session and post-session observation. The implementation of the control phase resulted in a sudden increase in the frequency of vocalizing We during treatment but had no effect on pre-session or post-session levels. Total vocalizations during the control phase was 111 (range 6-26 per session) across seven sessions. The control phase frequency of vocalizations was at a stable level when ET treatment was implemented. The first transition into ET treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of 229 occurrences (range, 9-44 per session) across eight sessions. During pre-session and post-observations, frequency of vocalizations remained at baseline levels. Frequency of vocalizations for We were stable when a reversal to baseline occurred. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for ET which resulted in a sudden initial increase in the frequency of vocalizations during treatment. Overall the treatment had no effect on pre-session or post-session observation levels. However, there were two sessions in which frequency of post-session observation occurred at a frequency of three. ET treatment reversal increased from prior treatment levels. The total frequency

of vocalizations during the reversal to ET treatment was 234 (range, 26-56 per session) across five sessions.

As shown in Figure 6, for target vocalization Pop (i.e. SSP treatment), initial baseline was at a frequency of zero. A control phase was implemented after baseline. During treatment, frequency of vocalizations increased slightly for a total of 19 occurrences (range, 0-10 per session) across eight sessions. Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout the control phase except for one session that had a total of two occurrences during post-session observation. Frequency of vocalizations for Pop remained at a stable level when treatment was implemented. In the treatment phase, frequency of vocalizations maintained at levels of zero frequency during pre-session and post-session observations. During treatment, frequency of vocalizations remained lower than the control phase and only totaled 2 occurrences (range, 0-2 per session) across eight sessions. Next, a reversal to baseline was implemented for SSP. No vocalizations occurred during reversal to baseline. Lastly, a reversal to the treatment phase was implemented. The total frequency of vocalizations during the reversal to SSP treatment remained at zero (range, 0-0 per session) across five sessions. Overall, the ET treatment produced vocalizations above zero level four sessions sooner than SSP. For ET vocalizations were produced in 13 out of 13 sessions, whereas SSP produced vocalizations on 1 out of 13 sessions.

Dante's Frequency of Vocalizations (Study 1)

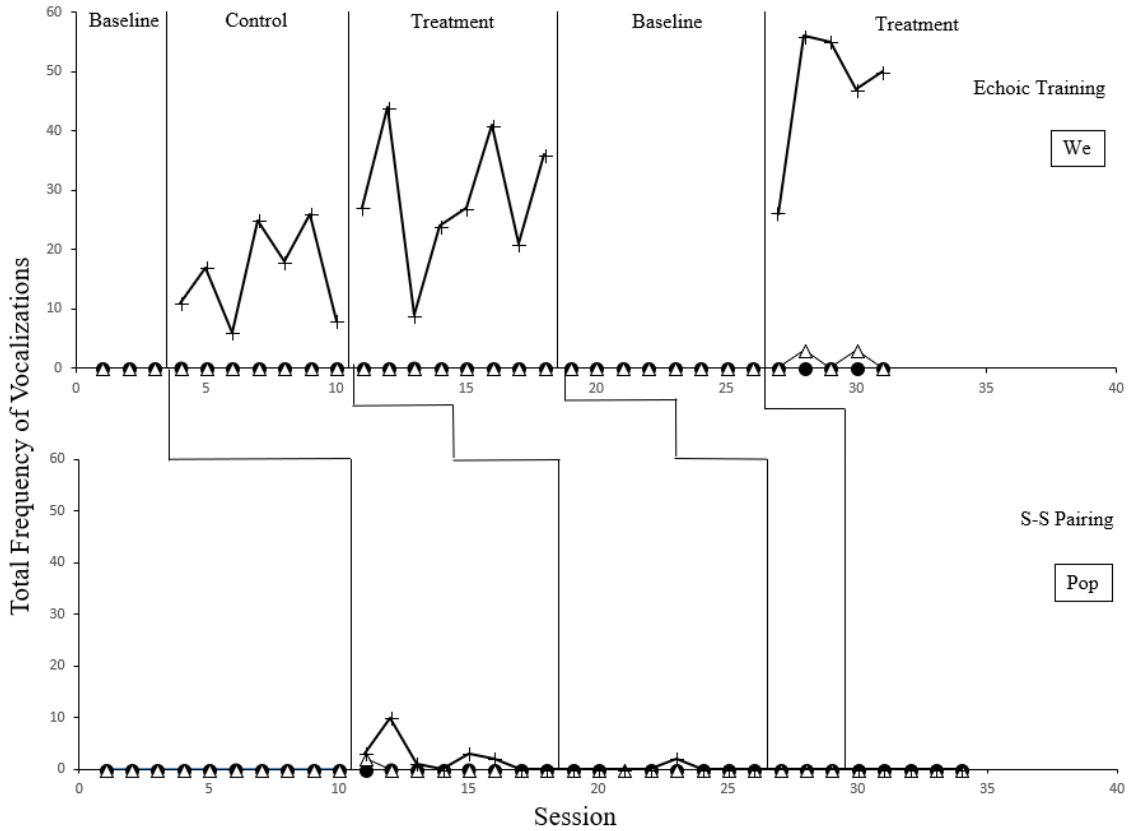


Figure 6. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Nonoverlap of all Pairs

To quantify and evaluate treatment efficacy, Non-overlap of All Pairs (NAP) was used in conjunction with visual analysis. NAP is a nonparametric index of data overlap between two phases in SCRM design experiments that was developed to improve on existing overlap-based effect sizes for SCRM (Parker & Vannest, 2009). The primary advantage of NAP over parametric analyses (e.g. *t*-tests, analyses of variance, ordinary

least squares regression) is that unlike parametric tests, NAP does not require assumptions of normality, and constant variance of residual scores (Parker, 2006). This is especially important since SCRM commonly fail to meet these assumptions. Furthermore, parametric effect sizes are limited by the fact that they are disproportionately influenced by extreme outliers, which are common in SCRM (Wilcox, 1998).

NAP can be defined as “the probability that a score drawn at random from a treatment phase will exceed (overlap) that of a score drawn at random from a baseline phase” or “the percent of non-overlapping data between baseline and treatment phase” (Parker & Vannest, 2009, p. 359). There are several ways to calculate NAP: (1) hand-calculation from graphs; (2) obtaining the percentage directly via Area Under Curve (AUC) percent from a Receiver Operator Characteristics (ROC) analysis; or (3) the intermediate output of the Wilcoxon Rank-Sum Test which is the larger U value (U_L) for phase B, divided by the total number of data comparisons ($N_A \times N_B$) (Parker & Vannest, 2009).

In NAP, overlap pairs are counted as one point, ties are counted as half a point. The total number of points is then subtracted from the total possible pairs and then divided by the total possible pairs. The total number of possible pairs is calculated by multiplying the number of data points from baseline (A phase) and treatment (B phase), $N_A \times N_B$. In the case of the present studies, NAP scores were calculated using the hand-calculation method. Calculations were completed to compare effect sizes between baseline and treatment as well as control and treatment for each participant. Parker and

Vannest (2009) describe the following ranges for effect sizes: weak effects: 0-0.65; medium effects: .66-.92; large or strong effects: .93-1.0. Table 9 summarizes the NAP scores of the participants.

For Jake, the SSP treatment NAP effect size was .53 from baseline and was .50 from control. The ET treatment NAP effect size was .50 from baseline and was .50 from control. Based on Parker and Vannest (2009), the treatment effect from both baseline and control was weak for both SSP and ET treatments. For Max, the SSP treatment NAP effect size was .81 from baseline and was .81 from control. The ET NAP treatment effect size was .70 from baseline and was .65 from control. Therefore, the treatment effect for SSP was moderate compared to baseline and control. For ET treatment, the effect size was moderate compared to baseline but weak from control. For Lane, the SSP treatment

Table 4

Study 1 NAP Effect Size Scores

	Baseline to	Control to	Baseline to	Control to
Learner	SSP	SSP	ET	ET
Jake	0.53	0.50	0.50	0.50
Max	0.81	0.81	0.70	0.65
Lane	0.96	0.34	1.00	0.67
Jordan	0.50	0.50	0.50	0.50
Dante	0.54	0.22	1.00	0.89

NAP effect size was .96 from baseline and was .34 from control. The ET treatment NAP effect size was 1.00 from baseline and was .67 from control. Therefore, the treatment effect from SSP was strong compared to baseline but weak compared to control. For ET treatment, the effect size was strong compared to baseline but moderate from control. For Jordan, the SSP treatment NAP effect size was .53 from baseline and was .50 from control. The ET treatment NAP effect size was .50 from baseline and was .50 from control. The treatment effect when compared to both baseline and control was weak for both SSP and ET treatments. For Dante, the SSP treatment NAP effect size was .54 from baseline and was .22 from control. The ET treatment NAP effect size was 1.00 from baseline and was .89 from control. Therefore, the treatment effect from SSP was weak compared to baseline and control. For ET treatment, the effect size was strong compared to baseline but moderate from control.

Overall, the SSP intervention produced at least moderate effects, when compared to baseline, for 2 out of 5 participants. SSP produced moderate to strong effects for 1 out of 5 participants when compared to control. Overall, the ET intervention produced at least moderate effects for 3 out of 5 participants when compared to baseline. ET produced at least moderate effects for 2 out of 5 participants when compared to control.

Study 2

Visual Analysis

Max. Max's results for study 2 are depicted in Figure 7. For target vocalization Boo (i.e. SSP treatment with direct reinforcement), baseline and reversal to baseline were all at a

frequency of zero. The first transition into SSP treatment with direct reinforcement resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. Total number of vocalizations that occurred during initial treatment was 17 occurrences (range, 0-14) across five sessions. During pre-session and post-observations, frequency of vocalizations remained at baseline levels. Frequency of vocalizations for Boo remained at a low level when a reversal to baseline occurred. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for SSP which resulted in a sudden initial increase in the frequency of vocalizations but eventually returned to low levels of frequency; five or less. The total frequency of vocalizations during the reversal to SSP treatment was 49 (range, 0-26) across six sessions.

As shown in Figure 7, for target vocalization Knee (i.e. ET treatment), initial baseline frequency was at zero. The first transition into ET treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of 29 occurrences (range, 0-29) across five sessions.

Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout treatment with only one exception in which a pre-session observation resulted in seven occurrences. Frequency of vocalizations for Knee reduced to a level of zero when a reversal to baseline occurred. Frequency of vocalizations maintained at levels of zero frequency across sessions during the baseline reversal. Next, a reversal to treatment was implemented for ET which

resulted in a sudden initial increase in the frequency of vocalizations. The total frequency of vocalizations during the reversal to ET treatment was 43 (range, 0-29) across five sessions.

Max's Frequency of Vocalizations (Study 2)

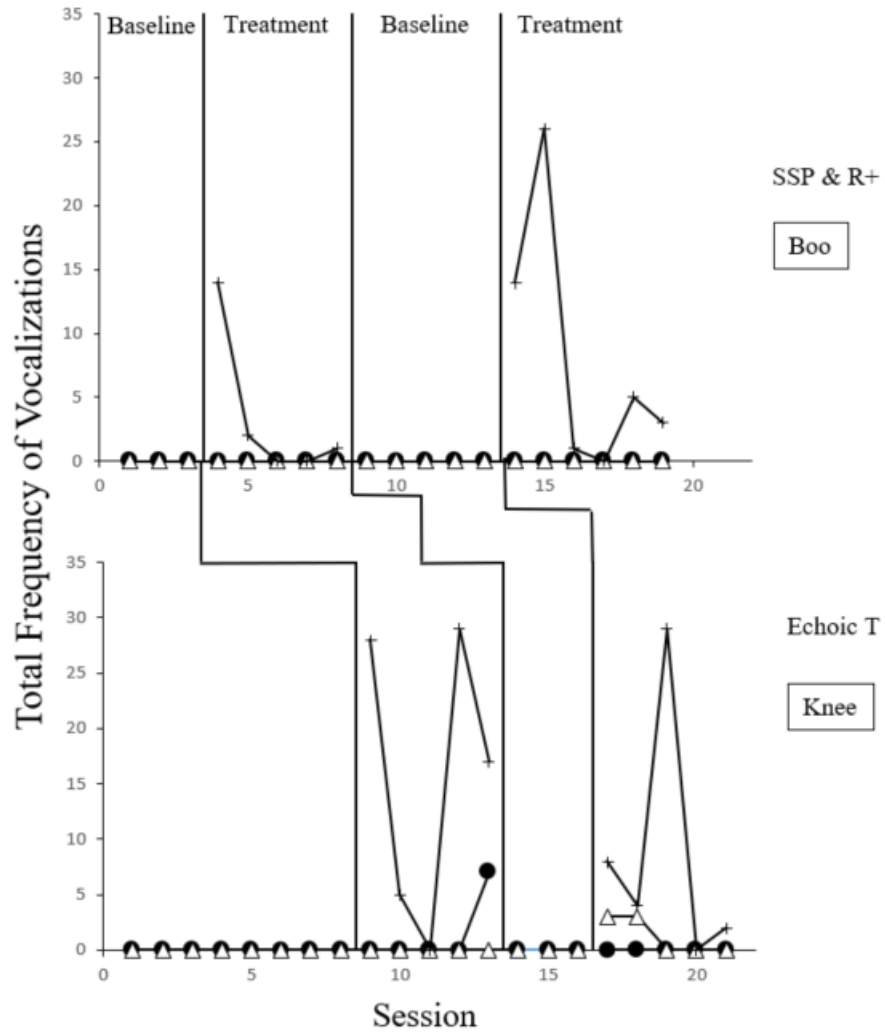


Figure 7. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Lane. Lane's results for study 2 are depicted in Figure 8. For target vocalization Window (i.e. SSP treatment with direct reinforcement), baseline and reversal to baseline were all at a frequency of zero. The first transition into SSP treatment with direct reinforcement resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. Total number of vocalizations that occurred during initial treatment was 52 occurrences (range, 0-16) across five sessions. During pre-session and post-observations, frequency of vocalizations remained at baseline levels. Frequency of vocalizations for Funny dropped to a 0 level when a reversal to baseline occurred. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for SSP and direct reinforcement which resulted in a sudden initial increase in the frequency of vocalizations but eventually returned to baseline levels of frequency (i.e. 0 occurrences). The total frequency of vocalizations during the reversal to SSP treatment was 24 (range, 0-11) across five sessions.

As shown in Figure 8, for target vocalization Funny (i.e. ET treatment), initial baseline frequency was at zero. The first transition into ET treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of 137 occurrences (range, 0-47) across five sessions. Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout treatment. Frequency of vocalizations for target Knee reduced to a level of zero when a reversal to baseline occurred and maintained at levels of zero across sessions during baseline reversal. Next, a reversal to treatment was implemented

for ET which resulted in a sudden initial increase in the frequency of vocalizations. The total frequency of vocalizations during the reversal to ET treatment was 85 (range, 0-19) across five sessions.

Lane's Frequency of Vocalizations (Study 2)

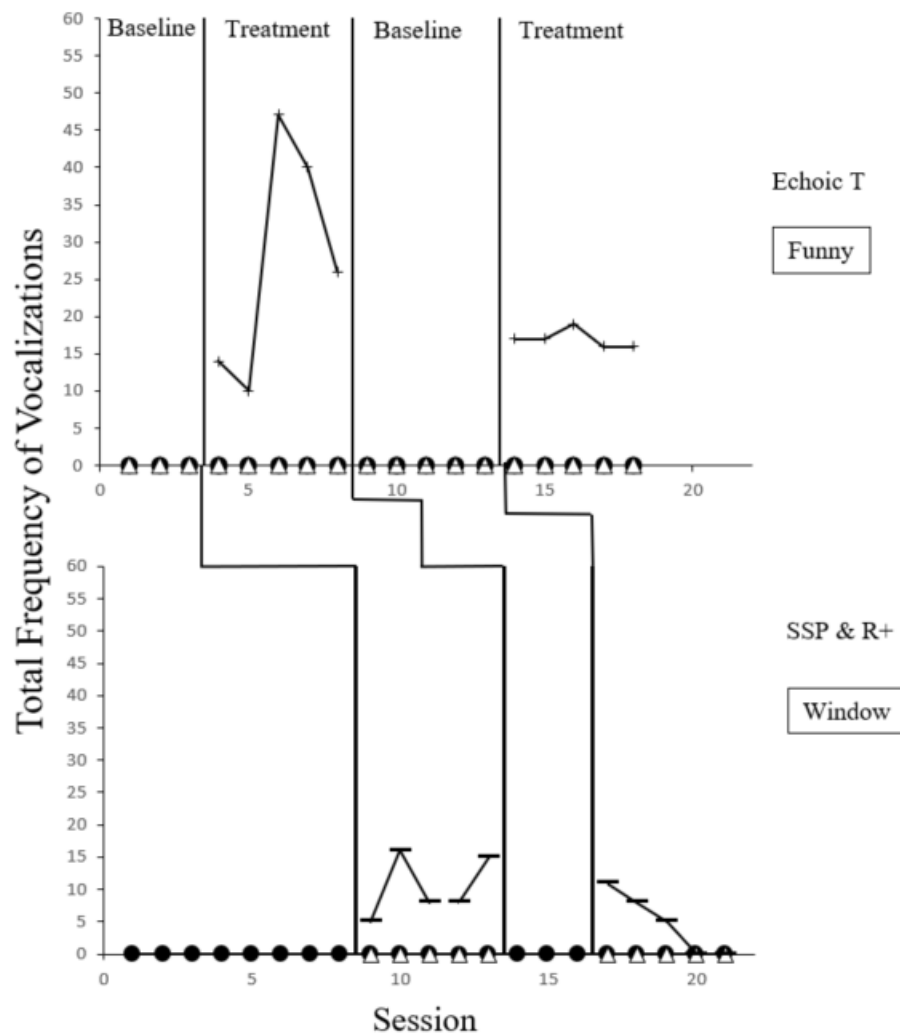


Figure 8. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path & minus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Dante. Dante's results for study 2 are depicted in Figure 9. For target vocalization Puppy (i.e. SSP treatment with direct reinforcement), baseline and reversal to baseline were all at a frequency of zero. The first transition into SSP treatment with direct reinforcement resulted no change in frequency for any pre-session, during treatment, or post-session observations. Total number of vocalizations that occurred during initial treatment was 0 occurrences across five sessions. No vocalizations occurred during reversal to baseline. Next, a reversal to treatment was implemented for SSP and direct reinforcement which did not lead to any change from baseline frequency. The total frequency of vocalizations during the reversal to SSP treatment was 0 across five sessions. As shown in Figure 9, for target vocalization Bunny (i.e. ET treatment), initial baseline frequency was at zero.

The first transition into ET treatment resulted in a significant increase in vocalizations during treatment observations but not for pre-session or post-session observations. The largest number of vocalizations occurred during treatment for a total of 91 occurrences (range, 0-25) across five sessions.

Frequency of vocalizations during pre-session and post-session observations maintained at baseline levels of zero throughout treatment. Frequency of vocalizations for target Bunny reduced to a level of zero when a reversal to baseline occurred and maintained at levels of zero across sessions during baseline reversal. Next, a reversal to treatment was implemented for ET which resulted in a sudden initial increase in the frequency of vocalizations. The total frequency of vocalizations during the reversal to ET treatment was 66 (range, 0-17) across five sessions.

Dante's Frequency of Vocalizations (Study 2)

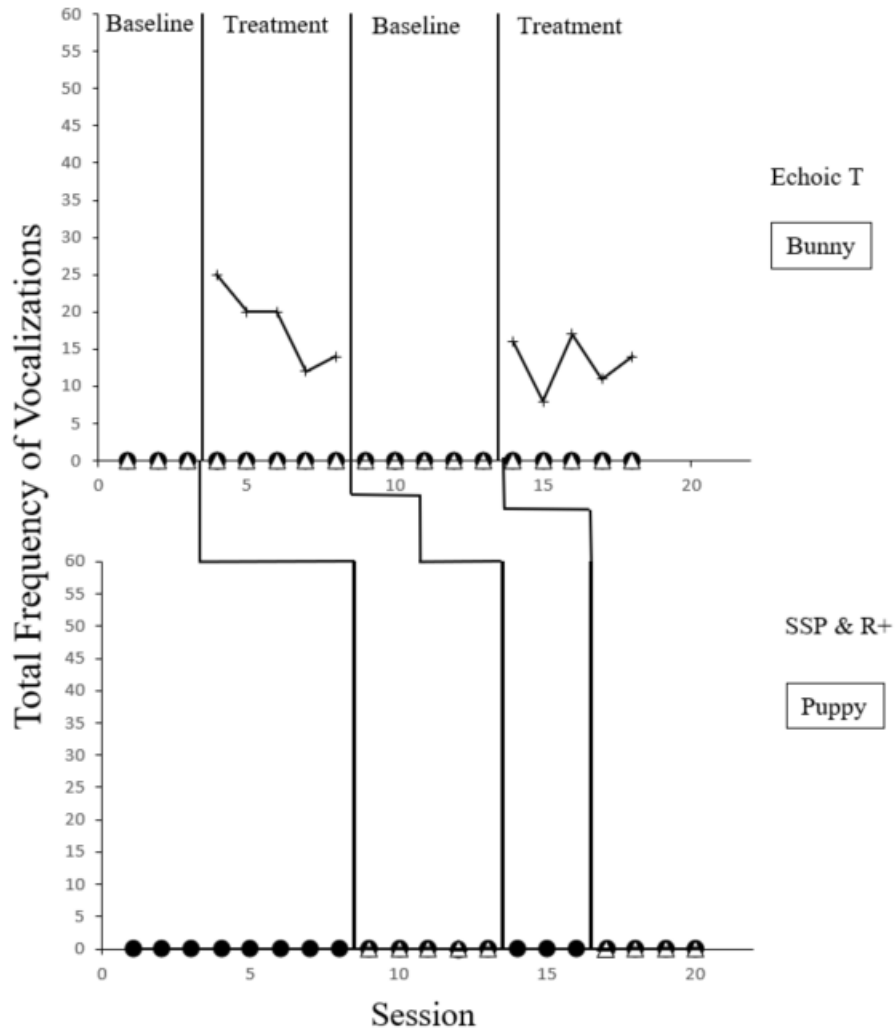


Figure 9. Total frequency of vocalizations in Pre-session observations (circle data path), during treatment (plus-sign data path & minus-sign data path), and post-session observations (triangle data path) for SSP and ET.

Nonoverlap of All Pairs

Refer to table 10 for a summary of the NAP scores. For Max, the SSP with direct reinforcement treatment effect size was .86 from baseline. The ET treatment effect size was .90 from baseline. Based on Parker and Vannest (2009), the treatment effect from

baseline was moderate for both SSP and ET treatments. For Lane, the SSP treatment effect size was .90 from baseline. The ET treatment effect size was 1.00 from baseline. Therefore, the treatment effect from SSP was strong compared to baseline. For ET treatment, the effect size was strong compared to baseline. For Dante, the SSP treatment effect size was .50 from baseline. The ET treatment effect size was 1.00 from baseline. Therefore, the treatment effect from SSP was weak compared to baseline. For ET treatment, the effect size was strong compared to baseline.

Table 5

Study 2 NAP Effect Size Scores

Learner	Baseline to SSP + R ⁺	Baseline to ET
Max	0.86	0.90
Lane	0.90	1.00
Dante	0.50	1.00

Overall, the SSP intervention with reinforcement produced weak effects for 1 out of 3 participants and moderate effects for 2 out of 3 participants. The ET intervention produced at moderate effects for 1 out of 3 participants and strong effects for 2 out of 3 participants.

IV. Discussion

In Study 1, SSP and ET interventions were compared with respect to each treatment's ability to produce vocalizations across sessions for five participants. A control condition was also included to account for the possibility that presentation of

vocalizations or reinforcers would be sufficient to produce treatment effects. For two of the participants, Jake and Jordan, the control condition and both treatments resulted in no change or low levels similar to baseline level vocalization frequency. Visual analysis indicated little to no difference between conditions. Similarly, NAP analysis indicated weak effects for both Jake and Jordan regardless of the condition (e.g. control or treatment). Jake was described as having little to no vocal play at the onset of the study. His receptive language age equivalence score for the MSEL was only 5 months despite being 26 months old chronologically. Expressive language scores were slightly higher at a 10-month age equivalence. Jake's results did not appear to add evidence for the first hypothesis that younger learners would benefit more from SSP than ET since neither treatment appeared more effective. Jake was the youngest participant in the study. However, an important observation was made by the experimenter during the study that could potentially explain Jake's results. Although Jake only emitted the targeted vocalizations in one session during the entire study, Jake did sporadically vocalize difficult words, such as 'bicycle', 'tricycle,' and 'classical' towards the end of the study. The experimenter questioned Jake's mother regarding the likely source of such vocalizations and discovered that the words were likely being repeated from a highly preferred cartoon Jake had been watching for some time prior to treatment sessions. Although this anecdotal evidence is limited, it indicates that the reinforcers being paired with the target vocalization for the SSP treatment were likely not powerful enough to produce the response. If other stimuli being paired with the words bicycle, tricycle, and classical were sufficiently powerful enough to produce these difficult vocalizations a considerable time later, then it can be reasonably assumed that the reinforcers used during

the intervention were simply too weak to produce a treatment effect. In addition, Jake's emission of difficult vocalizations such as 'tricycle' indicate that neuromuscular development and capability was not the limiting factor in production of speech sounds. Jake's results therefore highlight the importance of evaluating effective reinforcers for early learners as a critical step in the design of SSP based treatments. Lastly, an alternative possibility is that the vocalizations targeted during the study were occurring some time outside of the observation period. This possibility cannot be ruled out since this phenomenon was occurring for the other vocalizations.

For the next participant, Jordan, although the SSP intervention resulted in no effects for the target vocalization, there were other notable observations during Study 1. The SSP procedure did result in a frequency increase for non-target behaviors. Although formal data collection did not occur for non-target responses, the experimenter and observers noticed an increase from a rate of 0 non-target responses. For instance, Jordan began to make frequent eye contact and engage in other social behaviors such as laughing and smiling. Most unexpectedly, Jordan also began to request the reinforcer (tickles) by grabbing the experimenter's hand and bringing it closer to himself. These behaviors indicated that tickles were indeed functioning as a reinforcer for Jordan. The target vocalizations most likely did not increase due to an inability to produce the sounds or imitate the sounds. Throughout the entire course of the study, Jordan was never observed making any vocalizations. Occasionally, only some humming or stereotypy was produced by Jordan. Jordan's Expressive Language age equivalence on the MSEL was only 3 months despite being 34 months old. Jordan also appeared to be the most impacted of the five participants given that his scores were 98th percentile or higher across the four

domains evaluated in the ASRS. Jordan's results indicate that for early learners who are severely impacted and do not engage in vocal play, SSP is unlikely to produce any vocal responding but may aid in the production of other important behaviors. Since Jack and Jordan did not benefit from any treatment effects, they did not participate in study 2.

For the remaining three participants, the treatment effects varied. For Dante, visual analysis and NAP scores were very clear overall in that SSP was not an effective procedure in Study 1 or Study 2. The ET treatment was effective in both Study 1 and 2 and was able to produce vocalizations four sessions sooner than SSP. This efficiency advantage appeared to be tied directly to ET's effectiveness. During Study 1, there was some responding during the control phase for both treatments, indicating some evocative effect of the experimenter just presenting the target vocalizations. However, overall level of responding increased significantly in the ET treatment relative to control and baseline. NAP analysis indicated strong ET treatment effects relative to baseline and moderate, approaching strong, treatment effects relative to control. The NAP scores indicated a weak treatment effect for SSP relative to baseline and control for Study 1. For Study 2, SSP also produced a weak NAP score. Dante was the oldest participant chronologically in the present studies. Dante also had the highest MSEL age equivalence scores relative to his peers across all four skill domains of the MSEL. Combined with his overall chronological age, these results indicated Dante was likely the participant with the most advanced skill set. Dante's results added support to the hypothesis that older children with more developed verbal repertoires were less likely to benefit from SSP and more likely to benefit for ET. However, one unexpected finding was that SSP combined with direct reinforcement was less effective than ET alone. Dante would frequently accept the

delivery of a reinforcer during Study 2's SSP treatment without ever vocalizing after the delivery of the reinforcer. Essentially, Dante had discovered that under the ET treatment, responding was required for the delivery of reinforcers, and although during study 2, the SSP combined with direct reinforcement treatment reinforcement ratios could be larger, this had no effect on his rate of responding. The possibility that Dante had discriminated the difference in contingencies from Study 1 and that those discrimination effects had carried over to Study 2, despite targeting new responses, cannot be ruled out.

For Lane, Study 1 visual analysis indicated a clear treatment effect relative to baseline for both SSP and ET. Both treatments produced vocalizations immediately during their first session, indicating no advantage for either treatment regarding efficiency. NAP analysis also indicated strong treatment effects relative to baseline for both SSP and ET treatments. However, relative to control, SSP produced weak effects and ET produced moderate effects. This indicated that much of the SSP treatment effect relative to baseline was simply from the repetitive presentations of the vocalization by the experimenter having some evocative effect. This also indicated that SSP was not an effective treatment during Study 1 for Lane. During Study 2, SSP with direct reinforcement produced a moderate effect size. However, this was likely due to the direct reinforcement component of the treatment. This comparison provides evidence that SSP was actually producing a suppressive effect since NAP scores for ET treatment resulted in a strong effect size in comparison to SSP with direct reinforcement producing a moderate effect size. Since the main difference between SSP with direct reinforcement and ET is the stimulus-stimulus pairing component, it was likely SSP was inhibiting Lane's responding. This finding added evidence to previously reported findings that SSP

might produce suppressive effects (e.g. Esch et al., 2005) on frequency of responding. Lane's chronological age and MSEL age equivalence scores were lower across all domains relative to both Max and Dante. Given this difference, one might expect to find that SSP would be an effective treatment for Lane. This finding added conflicting evidence for the hypothesis that the SSP would be a superior treatment for younger learners with less developmentally advanced skill repertoires.

For Max, SSP produced vocalizations on the first session of intervention, whereas ET required an additional session, indicating a slight advantage for SSP with regards to efficiency. Visual and NAP analysis revealed a moderate effect size for SSP relative to baseline and control. The ET treatment effect was moderate relative to baseline but weak relative to the control condition. A weak effect size relative to the control condition indicates that simply presenting the target word was likely having some evocative effect. In contrast to SSP, the ET treatment did not appear to be an effective procedure for Max. Initially, there was a sudden increase in vocalizations during SSP treatment. However, visual analysis also revealed a decrease in treatment effect towards the very end of the second SSP treatment phase for Max, with the final three sessions returning to baseline levels. In Study 2, two new targets began treatment. It was hypothesized that SSP in combination with direct reinforcement would result in larger treatment effects than either treatment alone. Though NAP analysis did reveal a slight larger effect size for SSP with direct reinforcement relative to study 1's effect size (.86 vs .81), SSP with direct reinforcement resulted in a lower effect size relative to ET (.86 vs .90) in Study 2. This finding adds contradictory evidence for our hypothesis that SSP combined with direct reinforcement would result in higher rates of responding. Visual analysis also revealed

that the ET treatment had a significantly higher average number of vocalizations across sessions in Study 2. The shift in treatment efficacy is likely due to a diminishing treatment effect of the SSP treatment. The temporary treatment effect is consistent with previous studies reporting temporary efficacy or no treatment efficacy at all (e.g. Esch et al., 2005; Esch et al., 2009; Stock et al., 2008). This diminishing effect can be seen at the end of Max's Study 1 results. It appears that direct reinforcement in Study 2 maintained the SSP effect size at comparable levels to Study 1. It is probable that as a learner begins to discriminate that reinforcers are delivered regardless of responding, rate of responding begins to decrease for stimulus-stimulus pairings. In the case of Max, SSP was initially effective during Study 1, but as imitative behaviors came into contact with reinforcement, rate of responding increased where responding was necessary to contact reinforcement as is the case in echoic training. Since the SSP with direct reinforcement resulted in reinforcer delivery regardless of responding, but also resulted in additional reinforcement contingent on a vocalization, one might expect a higher ratio of reinforcer delivery to result in higher rates of responding. However, the results demonstrated this was not the case as mentioned previously. For Max, it appears SSP began to lose its efficacy in Study 1 and that diminishing efficacy carried over into Study 2. In contrast, visual analysis and NAP reveal that ET became a much more effective treatment relative to Study 1. Essentially, when one visually analyzes the data continuously from Study 1 to Study 2 for Max, one can see an initial period of time where SSP was effective and then that efficacy begins to decrease while the ET treatment efficacy begins to increase.

As previous studies have noted, SSP often produces mixed and inconsistent results across participants. What could explain these findings? It appears that the present

studies captured a shift in treatment efficacy for Max. This shift in efficacy could be conceptualized as a developmental or behavioral cusp (Rosales-Ruiz & Baer, 1997). Abilities (e.g. attending to relevant stimuli, vocal-motor skills, visual/auditory acuity) combined with other experiences, such as a history sufficient in quantity and quality of contingencies necessary to discriminate social contingencies (i.e. socially mediated patterns of reinforcement), interact to move a learner past a developmental cusp. Once a learner's repertoire reaches a skill set comprised of these experiences and skills, previous processes or procedures (e.g. SSP) can become ineffective in controlling behavior without more powerful aspects included (e.g. direct reinforcement). Developmental cusps could be conceptualized as comparable to behavioral cusps (Rosales-Ruiz & Baer, 1997) in that a new behavioral skill or set of skills provide access to novel contingencies of reinforcement but also explain developmentally how stimuli, processes, or procedures lose their effectiveness in bringing about behavioral responding. Stated differently, once a learner acquires a developmental history of a particular quality, it is no longer possible or extremely unlikely for the learner to respond the same to naturally present stimuli and contingencies. For instance, a learner that acquires the ability to read phonetically is unlikely to respond to written words in the environment as they once did. Similarly, responding to social stimuli through new skills and newly acquired learner history can result in similar outcomes. It is therefore possible that SSP is a process or procedure that is highly sensitive to a developmental cusp and could therefore explain why the findings have been inconsistent across many studies.

Overall, the findings are the following: (1) SSP was more efficient and had greater efficacy for 1 out of 5 participants in study 1; (2) ET was more efficient and had

greater efficacy for 1 out of 5 participants in study 1; (3) Neither treatment was more efficient for 3 out of 5 participants in study 1; (4) ET was more effective for 2 out of 5 participants in study 1; (5) ET was more effective at producing vocalizations for 3 out of 3 participants in study 2 when compared to SSP combined with direct reinforcement and was also more efficient for one participant. These findings, when considered with the current literature, highlight the need for additional research and that predictive participant characteristics for effective use of SSP require a very nuanced examination.

Limitations and Future Directions

The present studies and their findings were limited by a variety of factors. First, the sample size was limited to five participants and two of those participants did not demonstrate any treatment effect for either procedure. This result limited the strength of the evidence supporting any of the hypotheses. Future studies should evaluate the strength of a stimulus as a reinforcer by testing its ability to increase a response and function as a reinforcer. For instance, if a snack was suspected to function as a reinforcer, it should be tested in its ability to reinforce a motor imitation target, alternative vocalization, or complying with a simple request. Documenting that a stimulus can increase the probability of some other behavior indicates it will likely function as a reinforcer for a target vocalization. Testing stimuli reinforcer effectiveness will likely also be required if SSP is adopted in the future as a potential treatment option. Second, although the present study produced results in a natural classroom learning environment, it is possible that under more controlled and a less distracting environments the results might have had a stronger contrasting effect. Throughout the course of the study, for all

participants, other activities were occurring in the class room that could have shifted the learner's attention away. Maximizing a learner's attention to the SSP procedures will likely increase the salience of the pairing process. Furthermore, since some learner's may engage in the targeted responses outside of the observation windows, it is highly recommended that caregivers or those spending significant time with the learners be given some data collection tools, such as a simple form, that can be collected at the end of future studies to potentially capture any responding happening outside of the experimenter's observations. Another recommendation is that future studies control for carry-over effects between studies such as by using new participants. For instance, in the present studies, Dante did not respond to the SSP treatment in Study 2, potentially because of carry-over effects from the previous study. It's possible that SSP combined with direct reinforcement may be more effective than ET alone, but the learner's history in the present studies eliminated the possibility of a vocal response occurring after SSP since they had learned to discriminate when responding was necessary for contact with reinforcement. Using new participants would eliminate this possibility and allow an analysis of novel contingencies concurrently, thus providing more evidence for or against the combination of SSP and direct reinforcement as an effective treatment option.

Another important finding that the present studies revealed were unexpected benefits resulting from SSP for socially significant behaviors. SSP demonstrated the ability to produce other forms of desirable responding for non-targeted behaviors. Even though SSP might not produce specifically targeted vocal responses, it should be investigated in its ability or utility to produce other forms of responding such as eye contact, non-vocal mands, identification of reinforcer effectiveness, and conditioning of

reinforcers. Since SSP was able to produce non-vocal manding (e.g. hand pulling), eye contact, and smiling from the most impacted participant in the group, it is likely SSP can be used to evaluate the effectiveness of reinforcers in evoking social responses or increase the probability of other developmentally appropriate target responses that are perhaps easier to evoke. SSP should also be investigated for its ability to increase the probability of compliance for other tasks since it has the potential to increase socially important behaviors.

The findings from the present study also indicate that vocal babbling behaviors are likely a minimum necessary milestone needed by learners to benefit from SSP and that learners who are capable of echoic operants are also unlikely to benefit from SSP as a method of evoking or eliciting vocalizations. Investigations focused on identifying specific milestones or skills for predicting treatment outcome will need to center around examining development in between these major milestones.

Lastly, the present studies incorporated a NAP analysis (Parker & Vannest, 2009) to quantify effect sizes. Although Parker and Vannest (2009) suggest 0.0 to 0.65 as a “weak effect,” the present studies demonstrated that even at a .5 effect size, intervention did not differ from baseline and likely indicates no effect. It is therefore recommended that effect sizes interpretations be adjusted regarding NAP or that scoring be adjusted for two use cases. When (1) the target is an acceleration target (i.e. targeted for increase in rate) any sessions in which the target behavior does not occur should be scored as an overlap (e.g. 0). Similarly, when (2) the target is a deceleration target, any data point that is equivalent with the highest frequency data point in baseline should also be scored as an

overlap (e.g. 0). A baseline of zero represents a floor for responding and the highest frequency of an observed behavior could potentially be a naturally occurring ceiling for responding. These adjustments in scoring would allow NAP values of .5 and below to become more meaningful and NAP values overall to be better representations of their suggested interpretive ranges.

Conclusions

The current study has contributed to the literature by evaluating SSP as an alternative treatment option relative to ET for very early learners. The findings of the current study replicate and extend upon the findings of past studies evaluating SSP (Carrol & Klatt, 2008; Esch et al., 2009; Esch et al., 2005; Lepper, Petursdottir, & Esch, 2013; Miguel et al., 2002; Miliotis, et al., 2008; Normand & Knoll, 2006; Rader et al., 2014; Shillingsburg et al., 2015; Stock et al., 2008; Sundberg et al., 1996; Ward et al., 2007; Yoon, 1998; Yoon & Bennett, 2000; Yoon & Feliciano, 2007) for vocalizations, showing that SSP has the potential to work for some early learners briefly. Despite there being inconsistent findings across learners, the present studies indicate that SSP could prove more efficient and effective for some learners. If (a) clinicians struggle to evoke vocalizations using ET or (b) can identify participant characteristics or reinforcers that indicate SSP will be effective, then (1) SSP could be used to produce initial responding, (2) clinicians can reinforce imitative behaviors, and lastly (3) fade out SSP procedures to transfer control to imitation S^D s (discriminative stimuli). These steps would essentially be a SSP to Echoic stimulus control transfer procedure. Correctly implementing SSP in this manner could improve overall treatment efficiency rather than assuming ET should be the

starting point for production of vocalizations across all learners. Although SSP treatment efficacy could be short lived, SSP should serve as a facilitating procedure until vocalizations can be brought under control of more robust direct reinforcement procedures. The studies also indicate that SSP should be faded out immediately after direct reinforcement procedures have been successfully implemented, to prevent the suppressive effects of SSP that eventually occur. For example, participants Max and Lane both demonstrated suppressive effects of SSP in study 2 and their results indicated direct reinforcement should quickly be used in isolation once responding is under stimulus control. Overall, additional research must be conducted on SSP based procedures and their ability to produce desired outcomes. The present studies added some evidence that higher functioning learners with echoic skills are unlikely to benefit from SSP and that those with no vocal babbling are also unlikely to benefit. SSP shows potential to be an effective procedure in producing both vocalizations and other non-vocal but socially significant behaviors for the development an onset of other skills. Although age and developmental skillsets appear to be correlated with SSPs effectiveness, further investigation is required for clear delineations of intervention recommendations.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.
- Bennett, G., & Yoon, S. (2000, May) *The Role of a pairing procedure on subsequent mand acquisition and relations between children's baseline vocal play or vocal verbal repertoires and the effectiveness of the pairing procedure*. Paper presented at the 26th annual convention of the Association for Behavior Analysis, Washington, DC.
- Bijou, S W., & Baer, D. M. (1965). *Child Development II: Universal stage of infancy*. New York: Appleton-Century-Crofts.
- Bzoch, K. R., League, R., & Brown, V. L. (2003). *Receptive-expressive emergent language test* (3rd ed.). Austin, TX: Pro-Ed.
- Bowker, A., D'Angelo, N. M., Hicks, R., & Wells, K. (2011). Treatments for autism: parental choices and perceptions of change. *Journal of Autism and Developmental Disorders*, 41(10), 1373. doi: 10.1007/s10803-010-1164-y.
- Carroll, R. A., & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach vocal verbal behavior to young children with autism. *The Analysis of Verbal Behavior*, 24(1), 135–146.
- Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network. MMWR 2014;63(No. SS-#6302).
- Chawarska, K., Macari, S., & Shic, F. (2013). Decreased Spontaneous Attention to Social Scenes in 6-Month-Old Infants Later Diagnosed with Autism Spectrum Disorders. *Biological Psychiatry*, (3), 195. doi:10.1016/j.biopsych.2012.11.022
- Cividini-Motta, C. (2014). *Teaching verbal behavior: A systematic comparison of commonly used echoic teaching procedures* (Order No. 3629358). Available from ProQuest Dissertations & Theses A&I; ProQuest Dissertations & Theses Global. (1562018091).
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Dawson, G., Meltzoff, A. N., Osterling, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479-485. doi: 10.1023/A:1026043926488
- Esch, B. E., Carr, J. E., & Grow, L. L. (2009). Evaluation of an enhance stimulus–stimulus pairing procedure to increase early vocalizations of children with autism.

Journal of Applied Behavior Analysis, 42(2), 225–241.
<http://doi.org/10.1901/jaba.2009.42-225>

- Esch, B. E., Carr, J. E., & Michael, J. (2005). Evaluating stimulus-stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with Autism. *The Analysis of Verbal Behavior*, 21(1), 43–58.
- Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, 27(04), 491–503.
<http://doi.org/10.1017/S0140525X04000111>
- Fisher, W. W., Piazza, C. C., Bowman, L. G., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment. *American Journal on Mental Retardation*, 101, 15–25.
- Glenn, S.S., Ellis, J., & Greenspoon, J. (1992). On the revolutionary nature of the operant as a unit of behavioral selection. *American Psychologist*, 47, 1329-1336.
- Gutierrez, A., Hale, M. N., O'Brien, H. A., Fischer, A. J., Durocher, J. S., & Alessandri, M. (2009). Evaluating the effectiveness of two commonly used discrete trial procedures for teaching receptive discrimination to young children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 3, 630-638.
- Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders*, 30, 137–142.
- Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., et al. (2008). Can children with autism recover? If so, how? *Neuropsychology Review*, 18, 339-366.
- Heron, W. T., & Skinner, B. F. (1939). An apparatus for the study of animal behavior. *Psychological Record*, 3, 166–176.
- Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities*, 21(1), 61–73.
- Holland, J. G. (1992). Language and the continuity of species. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 197-209). Reno, NV: Context Press.
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and Developmental Receptive Language Disorder- a Follow-up Comparison in Early Adult Life. II: Social, Behavioural, and Psychiatric Outcomes. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(05), 561–578.

- Hurtado-Parrado, C., & López-López, W. (2015). Single-case research methods: History and suitability for a psychological science in need of alternatives. *Integrative Psychological and Behavioral Science, 49*(3), 323–349. <http://doi.org/10.1007/s12124-014-9290-2>
- Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and Tactics of Behavioral Research* (Second Ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Kaufman, N. R. (1995). *Kaufman Speech Praxis Test*. Detroit: Wayne State University Press.
- Kasari, C., & Lawton, K (2010) New directions in behavioral treatment of autism spectrum disorders. *Current Opinion in Neurology, 23*(2), 137-143. doi: 10.1097/WCO.0b013e32833775cd.
- Kimball, J.W. (2002). Behavior-analytic instruction for children with autism: Philosophy matters. *Focus on Autism and Other Developmental Disabilities, 17*(2), 66-75.
- Krantz, D. H. (1999). The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association, 94*(448), 1372–1381.
- Kravitz, H. & Boehm, J. (1971). Rhythmic habit patterns in infancy: Their sequences, age of onset, and frequency. *Child Development, 42*, 399-413.
- Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of operant discrimination training on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 46*(3), 656–61.
- Lord, C., Rutter, M., DiLavore, P., and Risi, S., (2000). *Autism Diagnostic Observation Schedule*. Western Psychological Services. n.d.
- Lovaas, O. I. (1981). *Teaching developmentally disabled children: The me book*. Austin, TX: PRO-ED.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3–9.
- Lovaas, O. I. (2003). *Teaching individuals with developmental delays. Basic intervention techniques*. Austin, TX: Pro-Ed.
- Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive behavioral intervention for autism: A survey of clinical practices. *Research in Autism Spectrum Disorders, 3*, 421-428.

- Lovitt, T. C. (2012). Applied behavior analysis: A method that languished but should be restored. *Intervention in School and Clinic, 47*(4), 252-256. doi: 10.1177/1053451211424598
- Luyster, R., Qiu, S., Lopez, K., & Lord, C. (2007). Predicting outcomes of children referred for autism using the MacArthur–Bates communication inventory. *Journal of Speech, Language, and Hearing Research, 50*, 667–681.
- McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intervention behavioral treatment. *American Journal on Mental Retardation, 97*(4), 359-372.
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37*, 149-155.
- Michael, J. (2004). *Concepts and principles of behavior analysis* (rev. ed.) Kalamazoo, MI: Society for the Advancement of Behavior Analysis.
- Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus—stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior, 18*, 3–13.
- Miliotis, A., Sidener, T. M., Reeve, K. F., Carbone, V., Sidener, D. W., Rader, L., & Delmolino, L. (2012). An evaluation of the number of presentations of target sounds during stimulus-stimulus pairing trials. *Journal of Applied Behavior Analysis, 45*(4), 809–13.
- Mowrer, O. G. (1954). The psychologist looks at language. *American Psychology, 9*, 660-694.
- Mullen, E. M. (1995). *Mullen Scales of Early Learning*(AGS ed.). Circle Pines, MN: American Guidance Service Inc.
- Nakazima, S. (1962). A comparative study of the speech developments of Japanese and American English in childhood. *Study of Phonology, 2*, 27-29.
- Normand, M. P., & Knoll, M. L. (2006). The effects of a stimulus-stimulus pairing procedure on the unprompted vocalizations of a young child diagnosed with Autism. *The Analysis of Verbal Behavior, 22*(1), 81–85.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*, 943-aac4716.
- Parker, R. I. (2006). Increased Reliability for Single-Case Research Results: Is the Bootstrap the Answer? *Behavior Therapy, 37*(4), 326–338.

- Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012). The relationship between language skills, Adaptive behavior, and Emotional and Behavior problems in pre-schoolers with autism. *Journal of Autism and Developmental Disorders*, *42*(12), 2761+.
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*, 357-367. doi: 10.1016/j.beth.2008.10.006
- Paul, R. (2008). Interventions to Improve Communication. *Child and Adolescent Psychiatric Clinics of North America*, *17*(4), 835-856. doi:10.1016/j.chc.2008.06.011
- Paul, D. R., & Cohen, D. J. (1984). Outcomes of severe disorders of language acquisition. *Journal of Autism and Developmental Disorders*, *14*(4), 405–421. doi:10.1007/BF02409831
- Pavlov, I. P. (1927/1960). *Conditioned Reflexes* (G. V. Anrep, Trans.) New York: Dover.
- Rader, L., Sidener, T. M., Reeve, K. F., Sidener, D. W., Delmolino, L., Miliotis, A., & Carbone, V. (2014). Stimulus-stimulus pairing of vocalizations: A systematic replication. *The Analysis of Verbal Behavior*, *30*(1), 69-74.
- Richards, S. B., Taylor, R. L., Ramasamy, R., & Richards, R. (1999). *Single Subject Research: Applications in Educational and Clinical Settings*. Singular Publishing Group.
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, *37*, 8-38.
- Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: a developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, *30*(3), 533–544. <http://doi.org.ezproxy.fiu.edu/10.1901/jaba.1997.30-533>
- Shillingsburg, M. A., Hollander, D. L., Yosick, R. N., Bowen, C., & Muskat, L. R. (2015). Stimulus-Stimulus pairing to increase vocalizations in children with language delays: A review. *The Analysis of Verbal Behavior*, *31*(2), 215–235. <http://doi.org/10.1007/s40616-015-0042-2>
- Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies. *Journal of Speech & Hearing Research*, *36*(1), 105.
- Skinner, B.F. (1938/1966). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century. (Copyright renewed in 1966 by the B.F. Skinner Foundation, Cambridge, M.A.)
- Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley.

- Skinner, B. F. (1967). B. F. Skinner. In E. G. Boring & G. Lindzey (Eds.), *A history of psychology in autobiography* (Vol. 5, pp. 385–413). New York: Appleton.
- Stock, R. A., Schulze, K. A., & Mirenda, P. (2008). A comparison of stimulus-stimulus pairing, standard echoic training, and control procedures on the vocal behavior of children with autism. *The Analysis of Verbal Behavior*, 24(1), 123.
- Sundberg, M.L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, 13, 21-37.
- Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner’s analysis of verbal behavior for children with autism. *Behavior Modification*, 25(5), 698–724.
- Sundberg, M. L., & Partington, J.W. (1998). *Teaching language to children with autism or other developmental disabilities*. Danville, CA: Behavior Analysts, Inc.
- Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behavior*, 27, 699-715.
- Thelen, E. (1981). Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology*, 17, 237-257.
- Ventner, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry*, 33, 489–507
- Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review*, 30(4), 387-399. doi: <http://dx.doi.org/10.1016/j.cpr.2010.01.008>
- Ward, S. J., Osnes, P. J., & Partington, J. W. (2007). The effects of a delay of noncontingent reinforcement during a pairing procedure in the development of stimulus control of automatically reinforced vocalizations. *The Analysis of Verbal Behavior*, 23(1), 103–111.
- Wilcox, R. R. (1998). How many discoveries have been lost by ignoring modern statistical methods? *American Psychologist*, 53, 300-314.
- Yoon, S. (1998). *Effects of an adult’s vocal sound paired with a reinforcing event on the subsequent acquisition of mand functions*. Unpublished doctoral dissertation, Columbia University, New York.
- Yoon, S.-Y., & Bennett, G. M. (2000). Effects of a stimulus—stimulus pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior*, 17, 75–88.

Yoon, S., & Feliciano, G. M. (2007). Stimulus-Stimulus pairing and subsequent mand acquisition of children with various levels of verbal repertoires. *The Analysis of Verbal Behavior*, 23(1), 3–16.

VITA

ALEJANDRO RENE DIAZ

Born, Elizabeth, New Jersey

- 2004-2007 A.A., Psychology
Miami-Dade College
Miami, Florida
- 2008-2010 B.A., Psychology
Florida International University
Miami, Florida
- 2012-2015 M.S., Psychology
Florida International University
Miami, Florida
- 2015 -2018 Doctoral Candidate
Florida International University
Miami, Florida
- Teaching Assistant
Florida International University
Miami, Florida

PRESENTATIONS

Diaz, A.R., Patin-Baratz, S., Correa, D., & Alvarez, A. (2017, October). *Stimulus-Stimulus Pairing vs. Echoic Training: An examination of developmental appropriateness*. Florida Association for Behavior Analysis, Annual Convention, Daytona, FL.

Diaz, A.R. (2017, October). *Ignite! How to Marry a Behavior Analyst.....and Survive*. Florida Association for Behavior Analysis, Annual Convention, Daytona, FL.

Diaz, A.R., & Gutierrez, A. (2014, September). *The sequence effects of receptive and expressive training*. Florida Association for Behavior Analysis, Annual Convention, Bonita Springs, FL.

Diaz, A.R., & Gutierrez, A. (2014, May). *The sequence effects of two types of training on verbal acquisition rates in children with autism*. Association for Behavior Analysis International, Annual Convention, Chicago, IL.